Lecture Notes
in Economics and
Mathematical Systems

Managing Editors: M. Beckmann and W. Krelle

273

Large-Scale Modelling and Interactive Decision Analysis

Proceedings, Eisenach, GDR, 1985

Edited by
G. Fandel, M. Grauer, A. Kurzhanski and A. P. Wierzbicki

Springer-Verlag

Lecture Notes
in Economics and
Mathematical Systems

Managing Editors: M. Beckmann and W. Krelle

273

Large-Scale Modelling and Interactive Decision Analysis

Proceedings of a Workshop sponsored by IIASA
(International Institute for Applied Systems Analysis)
and the Institute for Informatics of the Academy of Sciences
of the GDR
Held at the Wartburg Castle, Eisenach, GDR,
November 18–21, 1985

Edited by
G. Fandel, M. Grauer, A. Kurzhanski and A.P. Wierzbicki

Springer-Verlag
Berlin Heidelberg New York London Paris Tokyo

Managing Editors

Prof. Dr. M. Beckmann
Brown University
Providence, RI 02912, USA

Prof. Dr. W. Krelle
Institut für Gesellschafts- und Wirtschaftswissenschaften
der Universität Bonn
Adenauerallee 24–42, D-5300 Bonn, FRG

Editors

Prof. Dr. Günter Fandel
Fernuniversität Hagen
Fachbereich Wirtschaftswissenschaft
Feithstr. 140, D-5800 Hagen 1, FRG

Prof. Dr. Manfred Grauer
IIASA (International Institute for Applied Systems Analysis)
Schlossplatz 1, A-2361 Laxenburg, Austria and
Institute for Informatics, Academy of Sciences of the GDR
Rudower Chaussee 5, 1140 Berlin, GDR

Prof. Dr. Alexander Kurzhanski
IIASA and
Institute of Mathematics and Mechanics of the
Ural Scientific Center of the Academy of Sciences of the USSR
Sverdlovsk, USSR

Prof. Dr. Andrzej Piotr Wierzbicki
IIASA and
Institute of Automatic Control, Technical University of Warsaw
Nowowiejska 15/19, 00-665 Warsaw, Poland

ISBN 3-540-16785-4 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-16785-4 Springer-Verlag New York Berlin Heidelberg

Printed in Germany

Printing and binding: Beltz Offsetdruck, Hemsbach/Bergstr.
2142/3140-543210

Foreword

These Proceedings report the scientific results of an International Workshop on *Large-Scale Modelling and Interactive Decision Analysis* organized jointly by the System and Decision Sciences Program of the International Institute for Applied Systems Analysis (IIASA, located in Laxenburg, Austria), and the Institute for Informatics of the Academy of Sciences of the GDR (located in Berlin, GDR). The Workshop was held at a historically well-known place - the Wartburg Castle - near Eisenach (GDR). (Here Martin Luther translated the Bible into German.)

More than fifty scientists representing thirteen countries participated. This Workshop is one of a series of meetings organized by or in collaboration with IIASA about which two of the Lecture Notes in Economics and Mathematical Systems have already reported (Vol. 229 and Vol. 248).

This time the aim of the meeting was to discuss methodological and practical problems associated with the modelling of large-scale systems and new approaches in interactive decision analysis based on advanced information processing systems.

The meetings on multicriteria, interactive decision analysis that have been organized and supported by IIASA have established a tradition of an outstanding level of scientific discussions. This supplements the International Conferences on MCDM (multiple criteria decision making) in creating a unique forum for exchanging new research results between scientists from East and West that provides valuable ideas for future research, consistent with the goals of IIASA. Especially important, besides the broad representation of research results from planned economy countries, is the growing interest of participants from advanced industrial countries, e.g. Japan, the USA and many countries of Western Europe: this might be an indication of increasing interest in an organized exchange of scientific results between East and West. In this sense, the Wartburg Workshop provided also a convenient forum for a discussion and consensus-building on future activities in this scientific area.

With regard to the scientific results of the meeting, the following conclusions could be stressed:

- In the development of methodology of solving and analyzing multicriteria decision problems, one can observe consolidation tendencies, mature examples of comparison of existing methods and solution principles, as well as interesting methods of comparing alternative decisions. Further attention should be given to solving problems under uncertainty, with several decision makers and to corresponding problems in the social sciences and mathematical psychology.
- The fast development of information and knowledge processing technology provides a basis for the integration of many stages of systems analysis, starting from modelling through model simulation, analysis and optimization up to decision analysis and structuring of decision processes. These directions of integrated processing of numerical, symbolic and graphical information should be intensively developed further by intensifying the research on software modules and robust optimization algorithms for the analysis of decision problems.
- Many cases of applications presented at the Workshop have the mature character of concrete investigations of substantive problems from several disciplines, e.g. economics, energy, transportation, environmental studies or health services and have thus gone far beyond the early stage academic examples.

Based on these conclusions the contributions of the Workshop are structured in the Proceedings in three parts: (I) Theory and Methodology, (II) Interaction Principles and Computational Aspects and (III) Applications.

Part I contains papers dealing with utility and game theory, multicriteria optimization theory and interactive procedures, dynamic models/systems and concepts of multicriteria analysis. In Part II are papers combined dealing with the user-machine interface, intelligent (user-friendly) decision support and problems of computational aspects. Contributions with applications are mainly concentrated in Part III but can also be found in several papers of the other parts. Use of the term "large-scale" in the title of the Proceedings was especially substantiated by contributions dealing with modelling and decision analysis problems of the size of a whole national economy like structuring the carbochemical industry, the energy system or even natural gas trade in Europe.

The editors would like to take the opportunity to express their thanks to the sponsors, the System and Decision Sciences Program of IIASA and the Institute for Informatics of the Academy of Sciences of the GDR, for the successful organization and support of the Workshop as well as for the friendly reception in the historical rooms of the Wartburg Castle. We also wish to thank the authors for permission to publish their contributions in this volume and Elfriede Herbst for preparing the Proceedings.

February, 1986

G. Fandel
M. Grauer
A. Kurzhanski
A.P. Wierzbicki

Table of Contents

Part I. Theory and methodology

1. Utility and game theory

On measurable multiattribute value functions based on finite-order independence of structural difference 1
Hiroyuki Tamura and Shiro Hikita (Japan)

Preference separability and value functions 12
P.L. Yu (USA)

Two-person finite games with messages on the nature choices 22
B.S. Metev and Georgi G. Slavov (Bulgaria)

2. Multicriteria optimization theory and interactive procedures

A methodological approach to comparing parametric characterizations of efficient solutions 27
Andrzej P. Wierzbicki (Poland)

On scalarizing method in vector optimization 46
Dinh T. Luc (Vietnam)

Multicriterial duality, examples and advances 52
Alfred Göpfert (GDR)

Concepts of efficiency and fuzzy aggregation rules 59
Jochen Ester (GDR)

Multiregional economic policy-modelling 67
Marc Despontin (Belgium)

Interactive methods for multi-objective integer linear programming 75
J. Teghem Jr. and P.L. Kunsch (Belgium)

Interactive decision making for multiobjective linear programming problems with fuzzy parameters 88
Masatoshi Sakawa and Hitoshi Yano (Japan)

3. Dynamic models and systems

Multiobjective decisions for growth processes of time series 97
Manfred Peschel, Werner Mende (GDR) and Felix Breitenecker (Austria)

Interactive decision analysis and simulation of dynamic systems 107
Felix Breitenecker (Austria)

Pattern formation by natural selection with applications to decision support 117
Hans-Michael Voigt (GDR)

4. Concepts of multicriteria analysis

Rank preservation and propagation of fuzziness in pairwise-comparison methods for multi-criteria decision analysis 127
Freerk Lootsma (The Netherlands)

Learning in decision making 138
Robert Genser (Austria)

Interactive decision analysis for strategic management 148
Ivan Stantchev (Bulgaria)

Multi-objective decision analysis - the right objectives? 155
Eberhard Bischoff (UK)

Part II. Interaction principles and computational aspects

1. User-machine interface and intelligent decision support

Man-machine dialogue interfaces in decision support systems 161
Andrzej Lewandowski (Poland)

Solving discrete multiple criteria problems by using visual interaction 176
Pekka Korhonen (Finland)

Advanced computer applications for large-scale systems analysis 186
Kurt Fedra (Austria)

On an expert system for the design of chemical process systems 200
Klaus Hartmann and Michael Wagenknecht (GDR)

The object-oriented problem solving environment ARK-concept and first experience 215
Werner Gabrisch, Manfred Grauer and Lothar Pollmer (GDR)

2. Computational aspects of decision support

An interactive (optimization-based) decision support system for multi-criteria (nonconcave) control problems 222
R. Straubel and A. Wittmüss (GDR)

Toward the consolidation of interactive multiple objective programming procedures 232
Ralph E. Steuer and Alan W. Whisman (USA)

The interactive multiobjective package IMPROVE with application to strategic planning of carbochemical industry 242
Anne B. Barnikow, Ralph Gollmer and Manfred Grauer (GDR)

Part III. Applications

Experiences with interactive multi-objective linear programming methods in energy planning 251
Matthijs Kok (The Netherlands)

Decision-support model system for long-term water management in open-pit lignite mining areas 267
Stefan Kaden (GDR)

An approach for integrated energy-economy decision analysis: The case of Austria 277
Manfred Strubegger (Austria)

Natural gas trade in Europe and interactive decision analysis 291
Sabine Messner (Austria)

An application of satisficing trade-off method to a blending problem of industrial materials 303
H. Nakayama, J. Nomura, K. Sawada and R. Nakajima (Japan)

Multiobjective systems analysis for impact evaluation of the Hokuriku Shinkansen (Super-express railroad) 314
F. Seo, M. Sakawa, K. Yamane and T. Fukuchi (Japan)

A queueing theoretic model for planning diagnosis systems in hospitals 326
Günter Fandel and Holger Hegemann (FRG)

Decision support model for pest management: a preliminary report 340
Gerald L. Jubb, Subhash C. Narula and John F. Wellington (USA)

Decision support for development strategy in the area of interaction of the industrial and agricultural sectors 347
T. Ryś, M. Skocz, M. Zebrowski and W. Ziembla (Poland)

APPENDIX

List of Participants 358

On Measurable Multiattribute Value Functions Based on Finite-Order Independence of Structural Difference

Hiroyuki TAMURA and Shiro HIKITA*
Faculty of Engineering, Osaka University
2-1 Yamada-oka, Suita, Osaka 565, Japan

1. INTRODUCTION

Mathematical modeling of preferences has been widely studied in multiattribute decision analysis. Measurable value functions are based on the concept of a "difference in the strength-of-preference" (Fishburn, 1970) between alternatives. These functions provide an interval scale of measurement for riskless preferences. However, it is practically too difficult to directly identify a measurable multiattribute value function. Therefore, it is necessary to develop conditions that reduce the dimensionality of the functions that are required to identify. These conditions restrict the form of a measurable multiattribute value function in a decomposition theorem. Dyer and Sarin (1979) presented conditions for additive and multiplicative forms of the measurable multiattribute value function. These conditions are called "difference independence" and "weak difference independence". These conditions correspond to additive independence and utility independence, respectively, in multiattribute utility theory (Keeney and Raiffa, 1976).

In this paper we extend the condition of weak difference independence and propose a new concept of "finite-order independence of structural difference" for constructing measurable multiattribute value functions under certainty. This concept corresponds to "convex dependence" (Tamura and Nakamura, 1983) among multiple attributes for constructing multiattribute utility functions under risk. The essential idea of the concept of finite-order independence of structural difference is that we consider the change of decision maker's conditional strength-of-preference on one attribute depending upon the given conditional level of the other attributes. We describe decompositions of measurable multiattribute value functions based on this concept. These decompositions include Dyer-Sarin's additive/multiplicative decompositions as special cases.

---

* Presently with the Central Research Laboratory, Mitsubishi Electric Corporation, Amagasaki, Japan.

2. WEAK DIFFERENCE INDEPENDENCE

Let X be the set of all consequences in a decision problem. In the multiattribute problem X is described as $X = X_1 \times X_2 \times \ldots \times X_n$ where X_i denotes the set of possible consequences for the i-th attribute. Let $x^1, x^2, x^3, x^4 \in X$ where $x^k = (x_1^k, x_2^k, \ldots, x_n^k)$, k=1,2,3,4. Define X* as a non-empty subset of X×X and $\succsim^*$ as a weak order on X*. We describe

$$x^1x^2 \succsim^* x^3x^4$$

to mean that the difference of the strength-of-preference for x^1 over x^2 is greater than or equal to the difference of the strength-of-preference for x^3 over x^4. If we assume that $(X, X^*, \succsim^*)$ denotes a positive difference structure (Krantz, et al. 1971), there exists a real valued function v on X such that, for all $x^1, x^2, x^3, x^4 \in X$, if x^1 is preferred to x^2 and x^3 to x^4 then

$$x^1x^2 \succsim^* x^3x^4 \iff v(x^1) - v(x^2) \geqq v(x^3) - v(x^4) \tag{1}$$

Furthermore, since v is unique up to positive linear transformation, it is a cardinal function, and v provides interval scale of measurement.

We define the binary relation $\succsim$ on X by

$$x^1x^3 \succsim^* x^2x^3 \iff x^1 \succsim x^2, \tag{2}$$

then

$$x^1 \succsim x^2 \iff v(x^1) \geqq v(x^2). \tag{3}$$

Thus, v provides a measurable value function on X.

For $I \subset \{1,2,\ldots,n\}$ we partition the attributes into two sets X_I and $X_{\bar{I}}$ where X_I denotes the attribute sets with indices that are elements of I, and $X_{\bar{I}}$ denotes the attribute sets with indices that are elements of the complement of I. For $x_I \in X_I$, $x_{\bar{I}} \in X_{\bar{I}}$ we will write $x = (x_I, x_{\bar{I}})$.

DEFINITION 1.(Dyer and Sarin, 1979) The attribute X_i is difference independent of $X_{\bar{i}}$, denoted $X_i(DI)X_{\bar{i}}$, if, for all $x_i^1, x_i^2 \in X_i$ such that $(x_i^1, x_{\bar{i}}) \succsim (x_i^2, x_{\bar{i}})$ for some $x_{\bar{i}} \in X_{\bar{i}}$,

$$(x_i^1, x_{\bar{i}})(x_i^2, x_{\bar{i}}) \sim^* (x_i^1, x_{\bar{i}}')(x_i^2, x_{\bar{i}}') \tag{4}$$

for all $x_{\bar{i}}' \in X_{\bar{i}}$.

This Definition says that if $X_i(DI)X_{\bar{i}}$ the difference in the strength-of-preference between $(x_i^1, x_{\bar{i}})$ and $(x_i^2, x_{\bar{i}})$ is not affected by $x_{\bar{i}} \in X_{\bar{i}}$.

Dyer and Sarin (1979) introduced a weaker condition than difference independence, which is called weak difference independence. This condition plays a similar role to the utility independence condition in multiattribute utility theory (Keeney and Raiffa, 1976).

DEFINITION 2. (Dyer and Sarin, 1979) X_I is weak difference independent of $X_{\bar{I}}$, denoted $X_I(\mathrm{WDI})X_{\bar{I}}$, if, for given $x_I^1, x_I^2, x_I^3, x_I^4 \in X_I$ and some $x_{\bar{I}} \in X_{\bar{I}}$,

$$(x_I^1, x_{\bar{I}})(x_I^2, x_{\bar{I}}) \succsim^* (x_I^3, x_{\bar{I}})(x_I^4, x_{\bar{I}}), \tag{5}$$

then

$$(x_I^1, x_{\bar{I}}')(x_I^2, x_{\bar{I}}') \succsim^* (x_I^3, x_{\bar{I}}')(x_I^4, x_{\bar{I}}') \tag{6}$$

for all $x_{\bar{I}}' \in X_{\bar{I}}$.

This Definition says that if $X_I(\mathrm{WDI})X_{\bar{I}}$ the ordering of difference in the strength-of-preference depends only on the values of the attribute X_I and not on the fixed values of $X_{\bar{I}}$. Equations (5) and (6) imply that

$$v(x_I, x_{\bar{I}}) = \alpha(x_I) v(x_I, x_{\bar{I}}') + \beta(x_I), \qquad \alpha(x_I) > 0. \tag{7}$$

The property of the weak difference independence can be stated more clearly by using Normalized Conditional Value Function (NCVF) defined as follows:

DEFINITION 3. Given an arbitrary $x_{\bar{I}} \in X_{\bar{I}}$, we define an NCVF $v_I(x_I|x_{\bar{I}})$ on X_I and a Preference Structural Difference Function (PSDF) $d_I(x_I|x_{\bar{I}})$ as

$$v_I(x_I|x_{\bar{I}}) \triangleq [v(x_I, x_{\bar{I}}) - v(x_I^o, x_{\bar{I}})]/[v(x_I^*, x_{\bar{I}}) - v(x_I^o, x_{\bar{I}})] \tag{8a}$$

$$d_I(x_I|x_{\bar{I}}) \triangleq v_I(x_I|x_{\bar{I}}) - v_I(x_I) \tag{8b}$$

where

$$v(x_I^*, x_{\bar{I}}) > v(x_I^o, x_{\bar{I}}), \qquad v_I(x_I) \triangleq v_I(x_I|x_{\bar{I}}^o), \tag{9}$$

and $x^* \in X$ and $x^o \in X$ denote the best and the worst consequenses, respectively.

NCVF $v_I(x_I|x_{\bar{I}})$ denotes the ordering of preference on X_I, which is called preference structure here, under the given conditional level $x_{\bar{I}} \in X_{\bar{I}}$. PSDF $d_I(x_I|x_{\bar{I}})$ denotes the difference of preference structure between NCVF under the conditional level $x_{\bar{I}}$ and NCVF under the worst conditional level $x_{\bar{I}}^o$.

From Definition 3 we obtain

$$v_I(x_I^*|x_{\bar{I}}) = 1, \qquad v_I(x_I^o|x_{\bar{I}}) = 0. \tag{10}$$

From Definitions 2 and 3 the following equations hold, if $X_I(\mathrm{WDI})X_{\bar{I}}$.

$$v_I(x_I|x_{\bar{I}}) = v_I(x_I), \quad d_I(x_I|x_{\bar{I}}) = 0 \quad \text{for all} \quad x_{\bar{I}} \in X_{\bar{I}}. \tag{11}$$

In other words weak difference independence implies that NCVF does not depend on the given conditional level, and hence preference structure does not depend on the given conditional level.

We call that attributes $X_1, X_2, \ldots, X_n$ are mutually weak difference independent if, for every $I \subset \{1,2,\ldots,n\}$, $X_I(\mathrm{WDI})X_{\bar{I}}$. We now state the basic decomposition theorem of the measurable additive/multiplicative value functions.

THEOREM 1. If there exists a measurable value function v on X and if $X_1, X_2, \ldots, X_n$ are mutually weak difference independent, then either

$$1 + \lambda v(x) = \Pi_{i=1}^{n}[1 + \lambda\lambda_i v_i(x_i)] \quad \text{if} \quad \Sigma_{i=1}^{n}\lambda_i \neq 1 \tag{12}$$

or

$$v(x) = \Sigma_{i=1}^{n}\lambda_i v_i(x_i) \quad \text{if} \quad \Sigma_{i=1}^{n}\lambda_i = 1 \tag{13}$$

where

$$v(x^*)=1, \quad v(x^o)=0, \quad v_i(x_i^*)=1, \quad v_i(x_i^o)=0, \quad \lambda_i=v(x_i^*, x_{\bar{i}}^o), \tag{14}$$

and λ denotes a scaling constant such that

$$\lambda > -1, \quad \lambda \neq 0, \quad 1 + \lambda = \Pi_{i=1}^{n}(1 + \lambda\lambda_i).$$

Additive decomposition (13) can be obtained when every attribute is difference independent of the other attributes.

Dyer and Sarin (1979) stated this Theorem under the conditions of mutual preferential independence plus one weak difference independence instead of using the condition of mutual weak difference independence. For practical application it is easier to assess mutual preferential independence than to assess mutual weak difference independence.

3. FINITE-ORDER INDEPENDENCE OF STRUCTURAL DIFFERENCE

In this section we propose a new condition for the cases where eqn.(11) does not hold,

$$v_I(x_I | x_{\bar{I}}) \neq v_I(x_I), \quad d_I(x_I | x_{\bar{I}}) \neq 0, \quad \text{for some } x_{\bar{I}} \in X_{\bar{I}}. \tag{15}$$

that is, weak difference independence does not hold between X_I and $X_{\bar{I}}$.

DEFINITION 4. X_I is m-th order structural difference independent of $X_{\bar{I}}$, denoted $X_I(SDI_m)X_{\bar{I}}$, if, for given $x_I^1, x_I^2, x_I^3, x_I^4 \in X_I$ and some $x_{\bar{I}} \in X_{\bar{I}}$ such that

$$(x_I^1, x_{\bar{I}})(x_I^2, x_{\bar{I}}) \succsim^* (x_I^3, x_{\bar{I}})(x_I^4, x_{\bar{I}}), \tag{16}$$

there exist $x_{\bar{I}}^k \in X_{\bar{I}}$, $\theta_k(x_{\bar{I}})$, $k = 0,1,\ldots,m$, $x_{\bar{I}}^0 \neq x_{\bar{I}}^1 \neq \ldots \neq x_{\bar{I}}^m$ such that

$$\Sigma_{k=0}^{m}\theta_k(x_{\bar{I}})(x_I^1, x_{\bar{I}}^k) \; \Sigma_{k=0}^{m}\theta_k(x_{\bar{I}})(x_I^2, x_{\bar{I}}^k)$$
$$\succsim^* \; \Sigma_{k=0}^{m}\theta_k(x_{\bar{I}})(x_I^3, x_{\bar{I}}^k) \; \Sigma_{k=0}^{m}\theta_k(x_{\bar{I}})(x_I^4, x_{\bar{I}}^k) \tag{17}$$

This Definition represents the ordering of difference in the strength-of-preference between the linear combinations of consequences on X_I with (m+1) different conditional levels. If m = 0 in eqn.(17), we obtain eqn.(6), and hence

$$X_I(SDI_0)X_{\bar{I}} \iff X_I(WDI)X_{\bar{I}} \tag{18}$$

Equations (16) and (17) show that

$$v(x_I, x_{\bar{I}}) = \alpha(x_I)\ \Sigma_{k=0}^{m}\theta_k(x_{\bar{I}})v(x_I, x_{\bar{I}}^k) + \beta(x_I), \quad \alpha(x_I) > 0 \tag{19}$$

If $m = 0$ in eqn.(19), we obtain eqn.(7), and hence we can also find eqn.(18) from this result. Equations (17)-(19) show that the concept of "finite-order independence of structural difference" offers a natural extension of "weak difference independence".

If we define

$$\mu_k(x_{\bar{I}}) \overset{\Delta}{=} \theta_k(x_{\bar{I}})[v(x_I{*}, x_{\bar{I}}^k) - v(x_I^o, x_{\bar{I}}^k)] / \ \Sigma_{j=0}^{m}\theta_j(x_{\bar{I}})[v(x_I{*}, x_{\bar{I}}^j) - v(x_I^o, x_{\bar{I}}^j)] \tag{20}$$

$$k = 0, 1, \ldots, m$$

we will obtain the following Proposition.

<u>PROPOSITION</u>. If $X_I(SDI_m)X_{\bar{I}}$, then

$$d_I(x_I | x_{\bar{I}}) = \ \Sigma_{k=1}^{m}\mu_k(x_{\bar{I}})d_I(x_I | x_{\bar{I}}^k). \tag{21}$$

Proof.

If we rewrite eqn.(8a) using eqns.(19) and (20), we obtain

$$v_I(x_I | x_{\bar{I}}) = \Sigma_{k=0}^{m}\frac{\theta_k(x_{\bar{I}})[v(x_I{*}, x_{\bar{I}}^k) - v(x_I^o, x_{\bar{I}}^k)]}{\Sigma_{k=0}^{m}\theta_j(x_{\bar{I}})[v(x_I{*}, x_{\bar{I}}^j) - v(x_I^o, x_{\bar{I}}^j)]}\ v_I(x_I | x_{\bar{I}}^k)$$

$$= \Sigma_{k=0}^{m}\mu_k(x_{\bar{I}})v_I(x_I | x_{\bar{I}}^k). \tag{22}$$

If we sum up eqn.(20), we obtain

$$\Sigma_{k=0}^{m}\mu_k(x_{\bar{I}}) = 1. \tag{23}$$

If we subtract $v_I(x_I)$ from eqn.(22), we obtain

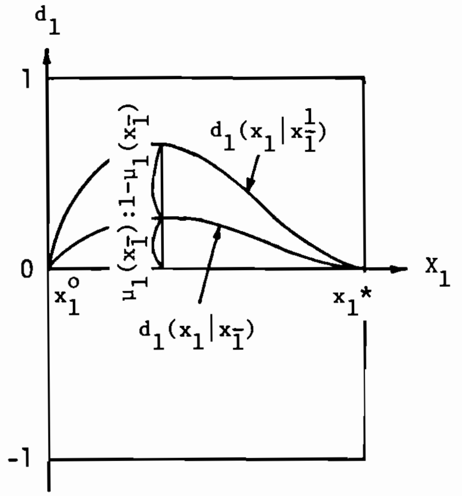
$$\begin{aligned} v_I(x_I | x_{\bar{I}}) - v_I(x_I) &= \Sigma_{k=0}^{m}\mu_k(x_{\bar{I}})v_I(x_I | x_{\bar{I}}^k) - v_I(x_I) \\ &= \Sigma_{k=0}^{m}\mu_k(x_{\bar{I}})v_I(x_I | x_{\bar{I}}^k) - v_I(x_I)\Sigma_{k=0}^{m}\mu_k(x_{\bar{I}}) \\ &= \Sigma_{k=1}^{m}\mu_k(x_{\bar{I}})[v_I(x_I | x_{\bar{I}}^k) - v_I(x_I)] \end{aligned}$$

Therefore, we obtain

$$d_I(x_I | x_{\bar{I}}) = \Sigma_{k=1}^{m}\mu_k(x_{\bar{I}})d_I(x_I | x_{\bar{I}}^k). \tag{21}$$

(Q.E.D.)

Equations (22) and (23) corresponds to the definition of m-th order convex dependence in multiattribute utility theory (Tamura and Nakamura, 1983). Equation (21) implies that if $X_I(SDI_m)X_{\bar{I}}$, PSDF on X_I depends on the given conditional level, and is written by a linear



$X_1(SDI_1)X_{\bar{1}}$

$d_1(x_1|x_{\bar{1}}) = \mu_1(x_{\bar{1}})d_1(x_1|x_{\bar{1}}^1)$

FIGURE 1 The relations among structural difference functions when the independence of structural difference property holds.

combination of m PSDFs with different conditional levels. Geometric illustration of this Proposition is shown in Fig. 1 where I = {1} and m = 1.

Since m-th order structural difference independence in this measurable multiattribute value theory corresponds to m-th order convex dependence in multiattribute utility theory, the decomposition theorems described in Tamura and Nakamura (1983) are valid if the expressions "utility function" are replaced by the expressions "measurable value function". For notational convenience we describe decomposition theorems only for two attribute cases here.

THEOREM 2. For $m = 0,1,\ldots,$ $X_1(SDI_m)X_2$ if and only if

$$v(x_1,x_2) = \lambda_1 v_1(x_1) + \lambda_2 v_2(x_2) + v_1(x_1)f(x_1{}^*,x_2)$$
$$+ \frac{v(x_1{}^*,x_2)-v(x_1^o,x_2)}{|V_1^m|} \Sigma_{k=1}^{m*} \Sigma_{j=1}^{m} \tilde{V}_{1,jk}^{m} d_1(x_1|x_2^k)d_1(x_1^j|x_2) \tag{24}$$

where

$$\lambda_1 = v(x_1{}^*,x_2^o), \qquad \lambda_2 = v(x_1^o,x_2{}^*) \tag{25a}$$

$$f(x_1,x_2) = v(x_1,x_2)-v(x_1^o,x_2)-v(x_1,x_2^o) \tag{25b}$$

$$V_1^m = \begin{bmatrix} d_1(x_1^1|x_2^1) & \ldots & d_1(x_1^1|x_2^{m-1}) & d_1(x_1^1|x_2{}^*) \\ \vdots & & \vdots & \vdots \\ d_1(x_1^m|x_2^1) & \ldots & d_1(x_1^m|x_2^{m-1}) & d_1(x_1^m|x_2{}^*) \end{bmatrix} \tag{25c}$$

$$\tilde{V}_{1,jk}^{m} = (-1)^{j+k}|V_{1,jk}^{m}|, \qquad \tilde{V}_{1,jk}^{1} = 1. \tag{25d}$$

$V_{1,jk}^{m}$ denotes $(m-1)\times(m-1)$ matrix obtained from V_1^m by deleting the j-th row and k-th column, and summation k=1 to m* means k=1,2,...,m-1,*.

THEOREM 3. For $m_1 = 0,1,\ldots,\ m_2 = 0,1,\ldots,\ X_1(SDI_{m_1})X_2$ and $X_2(SDI_{m_2})X_1$ if and only if

$$\begin{aligned} v(x_1,x_2) &= \lambda_1 v_1(x_1) + \lambda_2 v_2(x_2) + (1-\lambda_1-\lambda_2)v_1(x_1)v_2(x_2) \\ &+ (1-\lambda_2)v_2(x_2)\ \Sigma_{k=1}^{m_1} V_{1,k}^{m_1}(x_1)d_1(x_1^k|x_2^*) \\ &+ (1-\lambda_1)v_1(x_1)\ \Sigma_{j=1}^{m_2} V_{2,j}^{m_2}(x_2)d_2(x_2^j|x_1^*) \\ &+ \Sigma_{k=1}^{m_1}\Sigma_{j=1}^{m_2} V_{1,k}^{m_1}(x_1)V_{2,j}^{m_2}(x_2)D(x_1^k,x_2^j) \end{aligned} \tag{26}$$

where

$$V_{1,k}^{m_1}(x_1) = \frac{1}{|V_1^{m_1}|}\ \Sigma_{\ell=1}^{m_1*}\tilde{V}_{1,k\ell}^{m_1} d_1(x_1|x_2^k) \tag{27a}$$

$$V_{2,j}^{m_2}(x_2) = \frac{1}{|V_2^{m_2}|}\ \Sigma_{\ell=1}^{m_2*}\tilde{V}_{2,j\ell}^{m_2} d_2(x_2|x_1^j) \tag{27b}$$

$$\begin{aligned} D(x_1,x_2) &= f(x_1,x_2) - v_1(x_1)f(x_1^*,x_2) - v_2(x_2)f(x_1,x_2^*) \\ &+ v_1(x_1)v_2(x_2)f(x_1^*,x_2^*) \end{aligned} \tag{27c}$$

$$|m_1 - m_2| \leqq 1. \tag{27d}$$

Proof of Theorems 2 and 3 and the decomposition forms for more than two attribute cases can be obtained from Tamura and Nakamura (1983). Equation (27d) shows that if $X_1(WDI)X_2$, then we obtain either $X_2(WDI)X_1$ or $X_2(SDI_1)X_1$. Theorems 2 and 3 and the decomposition theorems for more than two attribute cases can construct a wide variety of measurable multiattribute value functions depending upon the order of structural difference independence.

4. INTERACTIVE ALGORITHM FOR IDENTIFYING MULTIATTRIBUTE MEASURABLE VALUE FUNCTIONS

As seen from the decomposition forms in Theorems 2 and 3, multiattribute measurable value functions can be identified if we know how to obtain

i) the single attribute value functions, since all the NCVFs included in the decomposition forms are single attribute measurable value functions,
ii) the order of structural difference independence, and
iii) the scaling coefficients appeared in the decomposition forms.

4.1 Single Attribute Measurable Value Functions

For identifying single attribute measurable value functions we use equal exchange method based on the concept of equal difference point (Dyer and Sarin, 1979).

DEFINITION 5. (Dyer and Sarin, 1979) For each attribute X_i if there exists $x_i^{0.5} \varepsilon X_i$ such that

$$(x_i^{0.5}, x_{\bar{i}})(x_i^o, x_{\bar{i}}) \sim^* (x_i^*, x_{\bar{i}})(x_i^{0.5}, x_{\bar{i}}) \tag{28}$$

for any given $x_{\bar{i}} \varepsilon X_{\bar{i}}$, then $x_i^{0.5}$ is the equal difference point for $[x_i^o, x_i^*] \subset X_i$.

From eqn.(28) we obtain

$$v_i(x_i^* | x_{\bar{i}}) - v_i(x_i^{0.5} | x_{\bar{i}}) = v_i(x_i^{0.5} | x_{\bar{i}}) - v_i(x_i^o | x_{\bar{i}}). \tag{29}$$

Since in eqn.(29)

$$v_i(x_i^* | x_{\bar{i}}) = 1, \qquad v_i(x_i^o | x_{\bar{i}}) = 0 \tag{30}$$

we obtain

$$v_i(x_i^{0.5} | x_{\bar{i}}) = 0.5. \tag{31}$$

Let $x_i^{0.25}$ and $x_i^{0.75}$ be the equal difference points for $[x_i^o, x_i^{0.5}]$ and $[x_i^{0.5}, x_i^*]$, respectively. Then we obtain

$$v_i(x_i^{0.25} | x_{\bar{i}}) = 0.25, \qquad v_i(x_i^{0.75} | x_{\bar{i}}) = 0.75 \tag{32}$$

If we plot the five points of eqns.(30)-(32) in x_i-v_i plane, the diagram like Fig. 2 is obtained. By some curve fitting techniques, say a

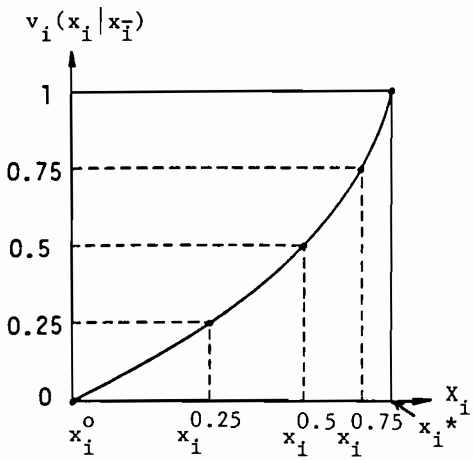


FIGURE 2 Curve fitting to equal difference points.

least square method, a single attribute value function $v_i(x_i|x_{\bar{i}})$ can be identified.

4.2 Order of Structural Difference Independence

For identifying a multiattribute measurable value function under the finite-order independence of structural difference (including 0-th order), we need to find the order m_i of structural difference independence between the attribute i and its complement. We define the matrix $V_i^{m_i}$ and the vectors $\underline{\theta}_i^{m_i}$ and $\underline{d}_i^{m_i}$ as follows:

$$V_i^{m_i} = \begin{bmatrix} d_i(x_i^1|x_{\bar{i}}^1) & \cdots & d_i(x_i^1|x_{\bar{i}}^{m_i-1}) & d_i(x_i^1|x_i*) \\ \vdots & & \vdots & \vdots \\ d_i(x_i^{m_i}|x_{\bar{i}}^1) & \cdots & d_i(x_i^{m_i}|x_{\bar{i}}^{m_i-1}) & d_i(x_i^{m_i}|x_{\bar{i}}*) \end{bmatrix} \tag{33a}$$

$$\underline{\theta}_i^{m_i} = [\theta_i^1 \ \cdots \ \theta_i^{m_i-1} \ \ \theta_i*] \tag{33b}$$

$$\underline{d}_i^{m_i} = [d_i(x_i^1|x_i^{m_i}) \ d_i(x_i^2|x_{\bar{i}}^{m_i}) \ \cdots \ d_i(x_i^{m_i}|x_{\bar{i}}^{m_i})] \tag{33c}$$

The algorithm for obtaining the order m_i is as follows:

Step 0. NCVFs $v_i(x_i)$ and $v_i(x_i|x_{\bar{i}}*)$ are assessed, and we draw a graph of $d_i(x_i|x_i*)$. If we can regard that

$$d_i(x_i|x_{\bar{i}}*) = 0,$$

then we decide that $X_i(SDI_0)X_{\bar{i}}$, and we set $m_i = 0$. If not, we set $m_i = 1$, and we go to Step 1.

Step 1. NCVFs and PSDFs are assessed for obtaining $V_i^{m_i}$ and $\underline{d}_i^{m_i}$.

Step 2. Linear equation

$$V_i^{m_i}\underline{\theta}_i^{m_i} = \underline{d}_i^{m_i} \tag{34}$$

is solved with respect to $\underline{\theta}_i^{m_i}$.

Step 3. We draw the graph of

$$f_i(x_i) = \Sigma_{k=1}^{m_i*} \theta_i^k d_i(x_i|x_{\bar{i}}^k). \tag{35}$$

Step 4. The graph of $f_i(x_i)$ is compared with the graph of $d_i(x_i|x_{\bar{i}}^{m_i})$. If we can regard that both curves are coincident within the allowable error, we decide that $X_i(SDI_{m_i})X_{\bar{i}}$. If not $m_i+1 \to m_i$ and then go back to Step 1.

These steps can be easily realized by using a graphic terminal of a large-scale computer interactively.

4.3 Scaling Coefficients

Scaling coefficients λ_1 and λ_2 appeared in Theorems 2 and 3 can be estimated as follows:

Step 1. We ask the decision maker to choose alternative (x_1*, x_2^o) or (x_1^o, x_2*) which he prefers. Suppose he prefers (x_1*, x_2^o) to (x_1^o, x_2*) and therefore $\lambda_1 \geqq \lambda_2$.

Step 2. For arbitrary $x_2^1 \varepsilon X_2$ $(x_2^o \neq x_2^1 \neq x_2*)$ we ask questions to the decision maker to find $x_1^1, x_1^2 \varepsilon X_1$ such that

$$(a) \quad (x_1^o, x_2*) \sim (x_1^1, x_2^o) \tag{36a}$$

$$(b) \quad (x_1^o, x_2*) \sim (x_1^2, x_2^1) \tag{36b}$$

This implies that we find two indifference points in $X_1 \times X_2$ space as shown in Fig. 3.

Step 3. Let

$$\alpha_1 = v_1(x_1^1), \quad \alpha_2 = v_1(x_1^2|x_2^1), \quad \alpha_3 = v_2(x_2^1), \quad \alpha_4 = v_2(x_2^1|x_1*) \tag{37}$$

then using eqs.(25a), (8a) ,(36) and (37) we obtain

$$\lambda_2 = v(x_1^o, x_2*) = v(x_1^1, x_2^o) = v_1(x_1^1)v(x_1*, x_2^o) = \alpha_1\lambda_1 \tag{38a}$$

$$\begin{aligned}\alpha_2 = v_1(x_1^2|x_2^1) &= [v(x_1^2, x_2^1) - v(x_1^o, x_2^1)]/[v(x_1*, x_2^1) - v(x_1^o, x_2^1)] \\ &= [\lambda_2 - \alpha_3\lambda_2]/[(1-\lambda_1)\alpha_4 + \lambda_1 - \alpha_3\lambda_2]\end{aligned} \tag{38b}$$

Solving eqn.(38) with respect to λ_1 and λ_2, we obtain

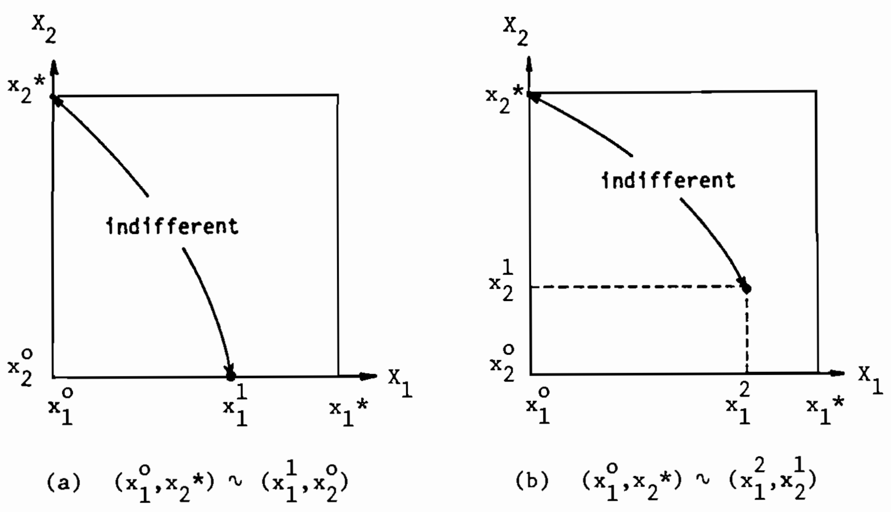


FIGURE 3 Assessment of indifference points for evaluating scaling coefficients.

$$\lambda_1 = \alpha_2\alpha_4/[\alpha_1(1-\alpha_3)-\alpha_2(1-\alpha_1\alpha_3\alpha_4)] \quad (39a)$$

$$\lambda_2 = \alpha_1\lambda_1 \quad (39b)$$

After obtaining the information for the order of structural difference independence, NCVFs and the scaling coefficients, we can construct a multiattribute measurable value function by using a decomposition form in Theorems 2 or 3. For two attribute cases we could draw indifference curves of the multiattribute measurable value function in two attribute space $X_1 \times X_2$.

5. CONCLUDING REMARKS

For obtaining multiattribute measurable value functions under certainty we have introduced a new concept of finite-order independence of structural difference as a natural extension of the concept of weak difference independence. The decomposition forms based on this concept includes Dyer and Sarin's additive and multiplicative decompositions as special cases. Therefore, depending upon the complexity of trade-offs among multiple attributes, this concept provides more flexible multiattribute measurable value functions as a riskless preference representation.

Although we didn't include measurable value functions for group decision making in this paper the concept of finite-order independence of structural difference would enable us to model a change of attitude of each decision maker towards the group value depending upon the consequence and the degree of satisfaction obtained by the other decision makers. Therefore, we could model various attitude of each decision maker in the group who are stubborn, selfish, sympathetic or ethical.

As a further research we need to clarify the relationship between utility functions under risk and measurable value functions under certainty (Sarin, 1982, Bell, 1982, and Krzysztofowicz, 1983). If we could discriminate a decision maker's strength-of-preference and the attitude towards risk, it might be possible to solve this problem.

REFERENCES

Bell, D. E. (1982). Regret in decision making under uncertainty, Opns. Res., 30(5):961-981.

Dyer, J. S., and Sarin, R. K. (1979). Measurable multiattribute value functions, Opns. Res., 27(4):810-822.

Fishburn, P. C. (1970). Utility Theory for Decision Making, John Wiley, New York.

Keeney, R.L., and Raiffa, H. (1976).Decisions with Multiple Objectives: Preferences and Value Tradeoffs, John Wiley, New York.

Krantz, D. H., Luce, R. D., Suppes, P., and Tversky, A. (1971). Foundations of Measurement, Academic Press, New York.

Krzysztofowicz, R. (1983). Strength of preference and risk attitude in utility measurement, Organizational Behavior and Human Performance, 31(1):88-113.

Sarin, R. K. (1982). Relative risk aversion, Management Sci., 28(8): 875-886.

Tamura, H. and Nakamura, Y. (1983). Decompositions of multiattribute utility functions based on convex dependence, 31(3):488-506.

PREFERENCE SEPARABILITY AND VALUE FUNCTIONS

P. L. Yu
School of Business, University of Kansas, Lawrence, Kansas 66045

1. Introduction.

In typical non-trivial multiple-criteria decision problems, the decision makers are faced with the ordering problems over a number of possible outcomes with a number of attributes. One of the human great inventions is the numerical system. If we could use the numerical systems for ordering so that the more preferred outcome will have a higher numerical value, then the decision making problem would become much easier. Unfortunately, this approach is not an easy one. A number of assumptions are needed for us to attach numerical ordering for the preference over the outcomes. (See [1-7], [10])

In the next section we shall sketch roughly under what conditions a set of revealed preference may be represented by partial order, weak order, value functions, monotonic value functions and additive value functions. As monotonic value functions and additive value functions are some of the most important forms in application, we shall explore the conditions under which a set of revealed preference may be represented by them. These conditions, which are hinged on whether or not the preference has the proper preference separability, will be discussed in section 3. In section 4 we shall focus on additive value functions which are most important in applications. We shall introduce a new concept, additive covering, to verify the preference separability for additive value function representations. Using "orthogonal designs," the new concept greatly reduces the labor for verifying preference separability for additive value functions. In second 5, we briefly sketch how additive or other kinds of value functions can be elicited through a number of methods. The concluding remark is then put in section 6.

2. Preference and Its Representation

2.1 Preliminary Definition

Let Y be the collection of all possible outcomes of a multiple criterion decision problems. The elements of Y will be denoted by $y = (y_1, \ldots, y_q)$ with y_j indicating the measurement of the jth attribute. Note that implicitly we assume that there are q attributes. For convenience, we shall use the superscript for the element index of Y, and the subscript for the component of y. In particular, y^i_j is the jth component of $y^i \in Y$.

For any y^1 and y^2 in Y, we write: (A) $y^1 \succ y^2$, if y^1 is preferred to y^2; (B) $y^1 \prec y^2$, if y^1 is less preferred than y^2; (C) $y^1 \sim y^2$, if y^1 is indifferent to y^2 or if the preference relation is unknown.

Definition 2.1 A Preference will be one or several of the following: (I) A preference based on $\succ$ (resp. $\prec$ or $\sim$) is a subset of YxY denoted by $\{\succ\}$ (resp. $\{\prec\}$ or $\{\sim\}$), so that whenever $(y^1, y^2) \in \{\succ\}$ (resp. $\{\prec\}$ or $\{\sim\}$), $y^1 \succ y^2$ (resp. $y^1 \prec y^2$, or $y^1 \sim y^2$); (II) $\{\succsim\} = \{\succ\} \cup \{\sim\}$, and $\{\precsim\} = \{\prec\} \cup \{\sim\}$.

Definition 2.2 $\{y^0 \prec\} = \{y \in Y \mid y^0 \prec y\}$: superior or better set w.r.t. y^0. $\{y^0 \succ\} = \{y \in Y \mid y^0 \succ y\}$: inferior or worse set w.r.t. y^0. $\{y^0 \sim\} = \{y \in Y \mid y^0 \sim y\}$: indifferent/indefinite set w.r.t. y^0.

Definition 2.3 (i) A preference $\{\succ\}$ is a partial order if it is transitive (i.e., if $y^1 \succ y^2$ and $y^2 \succ y^3$, then $y^1 \succ y^3$). (ii) A preference $\{\succ\}$ is a weak order if it is transitive, and $\{\succsim\} = \{\succ\} \cup \{\sim\}$ is also transitive.

Definition 2.4 $v: Y \to R^1$ is called a value function for $\{\succ\}$ on Y if for every $y^1, y^2 \in Y$, we have $y^1 \succ y^2$ iff $v(y^1) > v(y^2)$.

Definition 2.5 A value function v(y) is additive iff there are $v_i(y_i)$: $Y_i \to R^1$, $i = 1, \ldots, q$, such that:

$$v(y) = \sum_{i=1}^{q} v_i(y_i)$$

2.2 Hierarchy of Preference Representations

Depending on the features and/or assumptions we made on preference, a set of revealed preference may be represented by domination structures, partial order, weak order, value functions, monotonic value functions and additive value functions. In Figure 1, we give a hierarchy of preference representation in six classes.

In the initial revealed set of preference (Box 0) the preference may be represented as in Definition 2.1 and 2.2. The preference, without further assumptions, may be represented by domination structures. (Ch. 7 of [10]). In this class, nondominated solutions may be located as tentative "good" solutions.

When the preference has the structure or assumption that $\{\succ\}$ is transitive, then the preference becomes a partial order (Box 1). This class includes Pareto preference or preference represented by constant domination cones (See Chapters 3 and 7 of [10].) Here efficient solutions or nondominated solutions again can be located for tentatively "good" solutions.

Suppose $\{\succsim\}$ is also transitive. Then the preference $\{\succ\}$ can be represented by weak order (Box 2). In this class of preference, we could talk about indifference curves. Lexicographic order is one of the ordering in this class (See [4, 10]).

Suppose that $\{y \succ\}$ and $\{y \prec\}$ are open for each $y \in Y$. Then the preference can be represented by continuous value functions (Box 3). For details see Theorem 2.1 of this article and [1, 2, 4, 10].

Suppose there is preference separability for some components of the index set of the attributes. Then the preference may be represented by monotonic value functions (Box 4). The details of this condition are given in Theorem 3.1 - 3.3 For the details see [1, 2, 4, 10].

Suppose that the preference enjoys preference separability for each subset of the attributes index set. Then the preference can be represented by additive value functions (Box 5). See [1, 2, 4, 10] for details.

Note that additive value functions are most easily understood and most important in applications. However, the additive value functions involve the strongest assumptions. The verification of preference separability for each subset of the attribute index set is very laborious. In Section 4, we shall describe a streamline method which could release the laborious work of the verification work to a new minimum.

Theorem 2.1 (See Debreu [1]) Let (Y, T) be a topological space. Then there exists a continuous value function v for $\{\succ\}$ on Y in the topology T if; (i) $\{\succ\}$ on Y is a weak order; (ii) (Y, T) is connected and separable; (iii) $\{y\succ\}, \{y\prec\} \in T$, for every $y \in Y$.

(0) Preference; {Domination structures }

$\downarrow$ $\{\succ\}$ transitive

(1) Partial Order; Pareto Preference; constant cone domination structures

$\downarrow$ $\{\succsim\}$ transitive

(2) Weak Order; Lexicographic order

$\downarrow$ $\{y\prec\}$ and $\{y\succ\}$ are open for each $y \in Y$ (Theorem 2.1)

(3) Continuous Value functions

$\downarrow$ preference separability (Theorem 3.1 - 3.3)

(4) Monotonic Value functions

$\downarrow$ preference separability for each sub-collection of attribute index set (Theorem 3.1 - 3.3)

(5) Additive value functions

Figure 1

3. Preference separability and value function forms

To facilitate our discussion, let us introduce the following notation

(i) $Q = \{1, 2, \ldots, q\}$, the index set of the attributes;

(ii) given $\{I_1, \ldots, I_m\}$ being a partition of Q, define $z_k = y_{I_k}$, the vector with $\{y_i \mid i \in I_k\}$ as its components, $k = 1, \ldots, m$;

(iii) $Y_{I_k} = \prod_{i \in I_k} Y_i$;

(iv) $y = (y_{I_1}, \ldots, y_{I_m}) = (z_1, \ldots, z_m) = z$;

(v) $\bar{I}_j = Q \setminus I_j$.

Definition 3.1 Given that $I \subset Q$, $I \neq Q$, $u \in Y_I$ and $w \in Y_{\bar{I}}$, we say that u (or I) is preference separable, of $\succ$ - separable, iff $(u^0, w^0) \succ (u^1, w^0)$ for any $u^0, u^1 \in Y_I$ and some $w^0 \in Y_{\bar{I}}$ implies that $(u^0, w) \succ (u^1, w)$ for all $w \in Y_{\bar{I}}$.

Definition 3.2 If $\{I_1, \ldots, I_m\}$ and $z = (z_1, \ldots, z_m)$ are a partition of Q and y respectively, and if $v(z) = F(v_1(z_1), \ldots, v_m(z_m))$, then v is said to be strictly increasing in v_i ($i \in \{1, \ldots, m\}$), iff v is strictly increasing in v_i with v_k ($k = 1, \ldots, m$; $k \neq i$) fixed.

Theorem 3.1 (i) If v(y) as defined in Definition 3.2 is strictly increasing in v_i, $i \in \{1, \ldots, m\}$, then z_i and I_i are $\succ$-separable. (ii) if v(y) is additive then $\{\succ\}$ enjoys $\succ$ - separability for any subset of Q. (See [10].)

Under suitable conditions, the converse of the above holds.

Definition 3.3 Let $I \subset Q$, $I \neq Q$. I is said to be essential if there exists some $y_{\bar{I}} \in Y_{\bar{I}}$ such that not all elements of Y_I are indifferent at $y_{\bar{I}}$. I is strictly essential if for each $y_{\bar{I}} \in Y_{\bar{I}}$ not all elements of Y_I are indifferent at $y_{\bar{I}}$. If I is not essential it is called inessential.

Assumption 3.1 (i) Each topological space $(Y_i, \mathcal{T}_i)$, $i=1,\ldots, q$, is topologically separable and connected. So, $(Y, \mathcal{T})$, with $Y = \prod_{i=1}^{q} Y_i$, $\mathcal{T} = \prod_{i=1}^{q} T_i$, is topologically separable and connected as well.

(ii) $\{\succ\}$ on Y is a weak order, and for each $y \in Y$, $\{y \prec\}$ and $\{y \succ\} \in \mathcal{T}$.

Note that Assumption 3.1 implies that $\{\succ\}$ can be represented by continuous value function. (Thereom 2.1).

Therorem 3.2 (Debreu) Assume that assumption 3.1 holds.

(i) $v(y)$ can be written as $v(y)=F\,(v_1(y_1),\ldots,v_q(y_q))$, where F is continuous and strictly increasing in v_i ($i=1,\ldots, q$) which are all continuous, iff each $\{i\}$, $i=1, \ldots, q$, is $\succ$-separable.

(ii) if there are at least three components of q that are essential, then we can write $v(y) = \sum_{i=1}^{q} v_i(y_i)$, where each v_i is continuous, iff each possible subset $I \subset Q$ is $\succ$- separable.

Theorem 3.3 (Gorman) Let $\mathcal{I} = \{I_0, I_1,\ldots, I_m\}$ and $(z_0, z_1,\ldots,z_m)$ be a partition of Q and y respectively. Assume that assumption 3.1 holds.

Then: (i) $v(y)=F(z_0,v_1(z_1),\ldots,v_m(z_m))$, (3.1)

where $f(z_0, \ )$ is continuous and strictly increasing in v_i($i=1,\ldots, m$), iff each I_i, $i=1,\ldots,m$ is $\succ$-separable.

(ii) Assume $m > 3$, and each $\{i\}$, $i \in q$, is strictly essential. Then we can write:

$$v(y) = \sum_{i=0}^{m} v_i(z_i) \tag{3.2}$$

iff $\bigcup_{k \in S} I_k$, $S \subset M = \{0, 1,\ldots m\}$ (i.e. the union of any subsets of $\mathcal{I}$) is $\succ$-separable.

4. Additive value functions

In this section we shall sketch the concept and applications of additive covering for verifying preference separability for additive value functions. For the details, see Yu and Takeda [11].

Definition 4.1 (i) Two subsets I_1 and I_2 of Q are said to overlap iff none of $I_1 \cap I_2$, $I_1 \backslash I_2$, $I_2 \backslash I_1$ are empty. (ii) Let $\mathcal{I}$ be a collection of subsets of Q. Then: (1) $\mathcal{I}$ is said to be connected if for any A, B of $\mathcal{I}$ there is a sequence $\{I_1, I_2, \ldots, I_s\}$ of $\mathcal{I}$ such that I_{k-1} overlaps with $I_k (k=2, \ldots, s)$, and $I_1 = A$ and $I_s = B$; (2) $\mathcal{I}$ is $\succ$ - separable if each element of $\mathcal{I}$ is $\succ$ - separable.

Definition 4.2 A collection of nonempty subsets of Q, $\mathcal{I} = \{I_1, \ldots, I_r\}$, $r \geq 2$, is an additive covering of Q if (i) $\mathcal{I}$ is connected; (ii) Q is contained by the union of the elements of $\mathcal{I}$ and (iii) each element of Q is contained by no more than two elements of $\mathcal{I}$.

Example 4.1 Let $Q = \{1, 2, 3, 4, 5\}$. Then $\mathcal{I}_1 = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 5\}\}$ and $\mathcal{I}_2 = \{\{1, 2, 3\}, \{3, 4\}, \{1, 5\}\}$ are two additive coverings of Q. But $\mathcal{I}_3 = \{\{1, 2, 3\}, \{2, 3, 5\}, \{3, 4\}\}$ is not an additive covering of Q because 3 is contained by three elements of $\mathcal{I}_3$.

Given an additive covering $\mathcal{I} = \{I_1, \ldots, I_r\}$, Define:

$$A_{ij} = I_i \cap I_j \tag{4.1}$$

$$\hat{I}_k = I_k \setminus (\cup \{I_i \in \mathcal{I} \mid i \neq k\}) \tag{4.2}$$

where $i, j, k = 1, \ldots, r$.

Note that $\hat{I}_k$ is the collection of elements in I_k which are not contained by other I_i, $i \neq k$. All elements of $\{A_{ij} | i > j\}$ and $\{\hat{I}_k \mid k = 1, \ldots, r\}$ are mutually disjoint and each element of Q must be in some I_i and so must be in some A_{ij} or $\hat{I}_k$. The totality of all nonempty A_{ij} and $\hat{I}_k$ therefore forms a partition of Q.

Define $\mathcal{D}(\mathcal{I}) = \{J_t | t = 1, \ldots, m\}$ (4.3)

where $J_t \neq \emptyset$ is either an element of $\{A_{ij}\}$ or an element of $\{\hat{I}_k\}$.

Lemma 4.1 For each additive covering $\mathcal{I}$ of Q, there is a unique partition $\mathcal{D}(\mathcal{I})$ of Q which is derived by (4.1)-(4.3). $\mathcal{D}(\mathcal{I})$ contains $m \geq 3$ elements.

Theorem 4.1 Suppose that the preference $\{\succ\}$ on Y satisfies Assumption 3.1 and enjoys the following properties: (i) each $\{i\}$ of Q is strictly essential; and (ii) there is an additive covering $\mathcal{I} = \{I_1, \ldots, I_r\}$, $r \geq 2$, of Q such that $\mathcal{I}$ is $\succ$ -separable. Then the preference can be additively represented by

$$v(y) = v(z_1, \ldots, z_m) = \sum_{t=1}^{m} v_t(z_t) \tag{4.4}$$

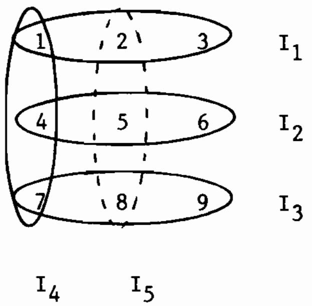
where $(z_1, \ldots, z_m)$ is the partition of y corresponding to $\mathcal{D}(\mathcal{I})$ defined in (4.3).

Example 4.2 Let $\mathcal{I}_1$ and $\mathcal{I}_2$ be as in Example 4.1. Let $\mathcal{D}_i$, $i = 1, 2$, be the corresponding partition of $\mathcal{I}_i$ defined by (4.3). Then for both i=1, 2,

$\mathcal{D}_1 = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}\}$. If the assumptions of the theorem hold, then the value function for $\{\succ\}$ can be written as $v(y) = \sum_{i=1}^{5} v(y_i)$.

Note that the special additive covering of the type of $\mathcal{I}_1$, was discussed in [6, 10] ; while the type of $\mathcal{I}_2$ is new. With respect to $\mathcal{I}_2$, <u>three</u> subsets of Q need to be verified for $\succ$ - separability but, with respect to $\mathcal{I}_1$, <u>four</u> subsets of Q need to be verified for $\succ$ - separability.

<u>Example 4.3</u> Let $Q = \{1, 2, \ldots, 9\}$ and $\mathcal{I}_1 = \{\{1, 2, 3\}, \{4, 5, 6\}, \{7, 8, 9\}, \{1, 4, 7\}\}$. (see Figure 2.)



<u>Figure 2</u>

Then $\mathcal{D}_1 = \{\{1\}, \{4\}, \{7\}, \{2, 3\}, \{5, 6\}, \{8, 9\}\}$ is the partition of Q. If the assumptions hold, we can write

$v(y) = v_1(y_1) + v_4(y_4) + v_7(y_7) + v_{23}(y_2, y_3) + v_{56}(y_5, y_6) + v_{89}(y_8, y_9)$.

Now if we add $I_5 = \{2, 5, 8\}$ to $\mathcal{I}_1$, then the corresponding partition is $\mathcal{D}_2 = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{7\}, \{8\}, \{9\}\}$; and $v(y) = \sum_{i=1}^{9} v_i(y_i)$ when the assumptions of the theorem hold.

The following orthogonal square designs show some efficient ways to verify $\succ$- separability for additive value functions. ([$\mathcal{I}$] is the number of element in $\mathcal{I}$).

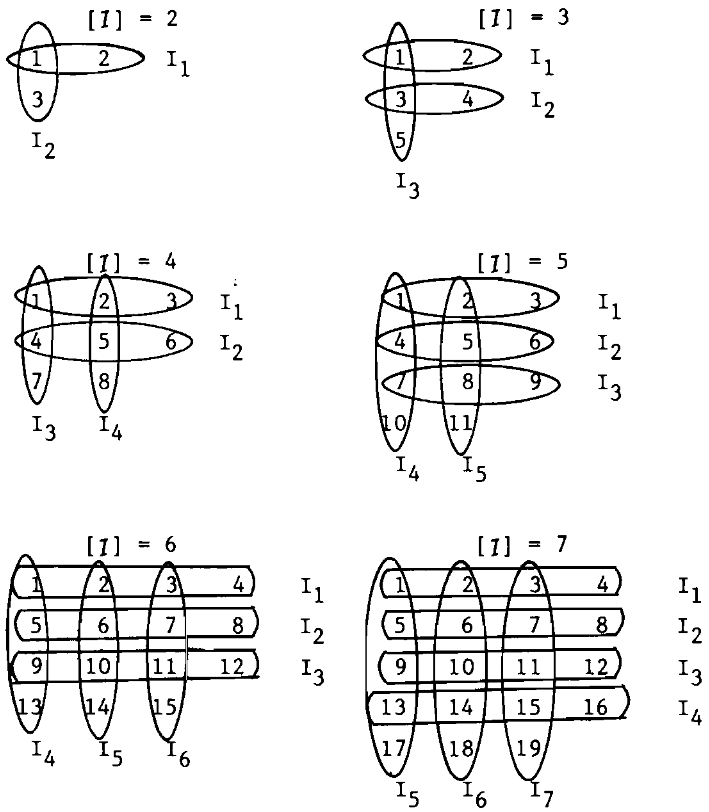


Figure 3 Orthogonal Square Designs

5. Elicitation techniques for constructing value functions

There are a number of methods to approximate the preference by value functions. These methods are usually called elicitation techniques for constructing value functions. Being limited by space we shall just sketch four class of the techniques. Hopefully by the title, the reader could relate it to the methods. For the details, the reader may refer to Chapter 6 of [10] and references quoted therein.

Class 1: Direct Applications of Calculus: Trade-off ratios; Tangent planes; Gradients; Line integrals.

Class 2: Additive value functions: (1) Indifference method; (2) Mid- value method.

Class 3: Minimization of inconsistencies in revealed preference: (1) Statistical methods-regression; (2) Mathematical programming models: minimizing inconsistence of { }; eigen weight methods; holistic assessment.

Class 4: Distance functions and compromise solutions.

6. Conclusion

We have sketched the hierarchy of preference representations, the preference separability and its relations to monotonic value functions and additive value functions, and the concept of additive covering for verifying preference separability for additive value function representation. The new method introduced can greatly reduce the work of verifying preference separability. As additive value functions are most easily understood and most easily analyzed, they have been used in many applied problems. Precautions are needed when we apply the additive value functions. We must verify whether the suitable preference separability are valid; otherwise, we may oversimplify the problem and obtain the inferior solutions. On the other hand, as with any real-life problem in analysis, we must not hesitate to make valid assumptions. Otherwise, the analysis would become extremely complex. The balance between making assumptions for the ease of analysis and the complexity of the reality needs to be maintained.

Many research problems are open. For instance, what would be the best procedure as to verify preference separability in real-life settings? Can one derive an effective interactive method which can verify the preference separability in the most efficient way? How to extend our results of preference separability to "membership functions" of fuzzy sets? This problem needs to be answered as to classify the membership functions. The above questions and their related ones are certainly waiting for the readers to explore.

References

1. Debreu, G., "Representation of a Preference Ordering by a Numerical Function," in Decision Processes, Thrall, R. M., Coombs, C. H. and Davis, R. L., Eds., Wiley, New York, 1954.

2. Debreu, G., "Topological Methods in Cardinal Utility," in Mathematical Methods in Social Science, Arrow K. J., Karlin, S. and Suppes, P., Eds., Stanford Univ. Press, Stanford, California, 1960.

3. Dyer, J. S. and Sarin, R. K., "Measurable Multiattribute Value Function," Operations Research, Vol. 27, pp. 810-822, 1979.

4. Fishburn, P. C., "Utility theory for Decision Making," John Wiley and Sons, New York, 1970.

5. Gorman, W. M., "The Structure of Utility Functions," Review of Economic Studies, 35, 367-390, 1968.

6. Keeney, R. L. and Raiffa, H., "Decisions with Multiple Objectives: Preferences and Value Tradeoffs," John Wiley and Sons, New York, 1976.

7. Krantz, D. H., Luce, R. D., Suppes, P. and Tversky, A., Foundations of Measurement, Vol. 1, Academic Press, New York, 1971.

8. Leitmann, G., An Introduction to Optimal Control, McGraw-Hill, New York, 1966.

9. Saaty, T. L., The Analytic Hierarchy Process, McGraw-Hill, New York, 1980.

10. Yu, P. L., Multiple Criteria Decision Making: Concepts, Techniques and Extensions, Plenum Publishing Co., New York, 1985.

11. Yu, P. L. and E. Takeda "A Verification Theorem of Preference Separability for Addition Value Functions." Working paper, School of Business, University of Kansas, Lawrence, Kansas 1985.

TWO-PERSON FINITE GAMES WITH MESSAGES ON THE NATURE CHOICES

B. S. Metev* and G. G. Slavov**

*Institute of Industrial Cybernetics & Robotics, BAS, 1113 Sofia, Bulgaria
**Bulgarian Industrial Association, 134 Rakovski Str.,1000 Sofia, Bulgaria

1. INTRODUCTION

A finite 2-persons game with the participation of the Nature is considered. There is a fixed sequence of moves of the partners, they have different information and communicate between themselves. A method for finding a strategy - the best in some sense for the first player - is proposed.

2. DESCRIPTION OF THE GAME

We shall consider a 2-persons game with the participation of the Nature. Let P1 denote the first player, and P2 - the second player.

P1 chooses $x \in X$, $|X| = m$, P2 chooses $y \in Y$, $|Y| = n$, the Nature chooses $a \in A$, $|A| = p$. The payoff function for P1 is $w_1 = f_1(x,y,a)$ and this for P2 is $w_2 = f_2(x,y,a)$. We will not take into account the interests of the Nature in this paper. P1 knows w_1, w_2 and the sets X, Y, A. P2 knows w_2 and the sets X,Y,A.

The Nature chooses an arbitrary $a_c \in A$. P2 comes to know the choosen a_c before to make his own choice or communication. On the other hand, P1 does not know a_c before his choice, but he knows that a_c is known to P2. Depending on circumstances each of these players wants to maximize his own payoff (or to get maximin). This behaviour of P2 is known to P1. P1 makes the first move.

There are different possibilities for P1 under the described conditions. In particular, instead of choosing $x \in X$ directly, P1 can put his own choice in dependence on the choice or the message of P2, and can inform P2 about this dependence in advance. The following case is considered in the paper: P1 proposes to P2 a strategy $\tilde{x}(y,a)$ with the meaning: if P2 reports a_r and chooses y_c, then P1 chooses $x_c = \tilde{x}(y_c,a_r)$. Subsequently, P1 follows the proposed strategy, P2 reports always exactly the choosed y_c, but the inequality $a_r \neq a_c$ is possible.

The sequence of the moves in the game is as follows. P1 determines a strategy $\tilde{x}(y,a)$. The Nature chooses a_c. P2 comes to know a_c. P1 announces the strategy $\tilde{x}(y,a)$ to P2. P2 chooses y_c and communicates y_c and a_r to P1. P1 chooses $x_c = \tilde{x}(y_c,a_r)$. P2 gets $f_2(x_c,y_c,a_c)$ and P1 gets $f_1(x_c,y_c,a_c)$.

3. SOME NOTIONS. PROBLEM FORMULATION.

The set of all possible functions $\tilde{x}(y,a)$ is denoted by $\tilde{X}$. $|\tilde{X}| = m^{np}$. The strategy $\tilde{x}(y,a)$ divides the set $X \times Y$ of all pairs (x,y) into two parts: realizable and unrealizable ones. The pair (x_t, y_t) is realizable with the strategy $\tilde{x}(y,a)$ if there exists such $a_t \in A$, that $x_t = \tilde{x}(y_t, a_t)$. If there does not exist such $a_t \in A$, the pair (x_t, y_t) is an unrealizable one with the strategy $\tilde{x}(y,a)$. The pair (x_c, y_c) is realized, if P2 has chosen y_c, and P1 has chosen x_c (in accordance with the strategy $\tilde{x}(y,a)$) at some $a_c \in A$.

When P2 has to make a move, he knows a_c and $\tilde{x}(y,a)$. P2 reports such a_r and chooses such y_c, which both force the pair (x_c, y_c), $x_c = \tilde{x}(y_c, a_r)$ to be realized, and this pair maximizes the payoff to P2 at $a = a_c$ on the set of pairs, realizable with $\tilde{x}(y,a)$.

Keeping the strategy $\tilde{x}(y,a)$ fixed, P1 knows the set of realizable pairs (x,y). Because P1 knows w_2, too, he knows for each $a \in A$ the set of pairs, maximizing the payoff for P2. The minimum of payoffs for P1 on this set (at the same a) and then on the whole A is his guaranteed result with the strategy $\tilde{x}(y,a)$. The problem is: to find a strategy, which maximizes this guaranteed result. We call this strategy the best one for P1.

In order to obtain an expression (a formula) for the maximal guaranteed result R (maximin) for P1, we can use the following reasoning. Let us assume that P1 has proposed to P2 a strategy $\tilde{x}(y,a) \in \tilde{X}$. P2, knowing a_c, reports such a_r and chooses such $y_c \in Y$, that

$$f_2(\tilde{x}(y_c,a_r),y_c,a_c) = \max_{\substack{t \in A \\ v \in Y}} f_2(\tilde{x}(v,t),v,a_c). \tag{1}$$

Denote $u = \tilde{x}(y,a)$. P1 must know the sets

$$B(u,a_c) = \left\{ (y_c,a_r) \middle/ f_2(\tilde{x}(y_c,a_r),y_c,a_c) = \max_{\substack{t \in A \\ v \in Y}} f_2(\tilde{x}(v,t),v,a_c) \right\}$$

for each $a_c \in A$. Therefore, for a fixed $a_c \in A$, P1 gets at least

$$\min_{(y_c,a_r) \in B(u,a_c)} f_1(\tilde{x}(y_c,a_r),y_c,a_c)$$

Minimizing for all $a_c \in A$ and maximizing on the whole $\tilde{X}$ we obtain the following expression for R:

$$R = \max_{u \in \tilde{X}} \min_{a_c \in A} \min_{(y_c,a_r) \in B(u,a_c)} f_1(\tilde{x}(y_c,a_r),y_c,a_c) \tag{2}$$

4. FINDING THE BEST STRATEGY FOR P1

A. Let $D_o = X \times Y$. Let $B_o(a)$ be the set of all such pairs $(x,y) \in X \times Y$, which maximize f_2 on the whole $X \times Y$ for a fixed $a \in A$. Let $C_o(a) \subseteq B_o(a)$ contain exactly these elements of $B_o(a)$, which minimize $f_1(x,y,a)$ on the $B_o(a)$ for the same a. We construct $\tilde{x}_o(y,a)$ in the following way. For each

$a_t \in A$ we fix one pair $(x_t,y_t) \in C_\circ(a_t)$. For the same a_t but for $y \neq y_t$ we choose an arbitrary $x \in X$.

For $a_t \in A$ P1 gets at least $f_1(\tilde{x}_\circ(y_t,a_t),y_t,a_t) = f_1(x_t,y_t,a_t)$, because if $a_c = a_t$ the realization of a pair $(x,y) \in B_\circ(a_t) \setminus C_\circ(a_t)$ would increase the payoff for P1, and the realization of a pair $(x,y) \in (X \times Y) \setminus B_\circ(a_t)$ would decrease the payoff for P2.

In general, it is possible to choose different pairs $(x_t,y_t) \in C_\circ(a_t)$ for a fixed $a_t \in A$, when constructing the strategy $\tilde{x}_\circ(y,a)$. But the payoff for P1 is constant for this a_t, therefore, we can denote

$$d_\circ = \min_{a_t \in A} f_1(x_t,y_t,a_t)$$

for the strategy $\tilde{x}_\circ(y,a)$.

Let $M_\circ \subseteq X \times Y$ be the set of all pairs (x,y), realizable with the strategy $\tilde{x}_\circ(y,a)$. Let $M_\circ^- \subseteq M_\circ$ contain exactly these elements of $M_\circ$, for which $w_1 = d_\circ$ at some $a \in A$.

If a strategy $\tilde{x}(y,a)$ allows to realize some pair $(x,y) \in M_\circ^-$, this strategy cannot guarantee to P1 more than $d_\circ$, because for each pair $(x,y) \in M_\circ^-$ there exists some $a \in A$, at which P2 get his own obtainable maximum with this pair (obtainable with this strategy) and P1 gets $d_\circ$.

It is possible to divide $\tilde{X}$ into two parts: $T_\circ'$ and $T_\circ''$, $T_\circ' \cap T_\circ'' = \emptyset$, $T_\circ' \cup T_\circ'' = \tilde{X}$. It is possible to realize at least one pair $(x,y) \in M_\circ^-$ with each strategy $\tilde{x}(y,a) \in T_\circ'$, i.e. there exists such a_r and y_t, that $(\tilde{x}(y_t,a_r),y_t) \in M_\circ^-$. There is no strategy $\tilde{x}(y,a) \in T_\circ''$, allowing to realize a pair $(x,y) \in M_\circ^-$. In other words, each strategy $x(y,a) \in T_\circ'$ guarantees $d_\circ$ to P1. If there exists a strategy, guaranteeing more to P1, such strategy belongs to $T_\circ''$.

B. Let $D_1 = (X \times Y) \setminus M_\circ^-$. We must check if there exists at least one admissible x for each $y \in Y$, beacuse a rejection of pairs $(x,y) \in M_\circ^-$ can lead to arising of such y_v, for which all pairs (x,y_v), $\forall x \in X$, are rejected. If this happens, the searching is over. The maximal guaranteed result (the maximin) for P1 is equal to $d_\circ$. This result can be obtained by the strategy $\tilde{x}_\circ(y,a)$.

If there exists at least one admissible $x \in X$ for each $y \in Y$ (the corresponding pair (x,y) is not yet rejected), the searching goes on. Now, $B_1(a)$ is the set of such pairs $(x,y) \in D_1$, which maximize f_2 on the whole D_1 for a fixed $a \in A$. $C_1(a) \subseteq B_1(a)$ contains exactly these elements of $B_1(a)$, which minimize f_1 on the whole $B_1(a)$ for the same a. The strategy $\tilde{x}_1(y,a)$ is determined like $\tilde{x}_\circ(y,a)$: for each $a_t \in A$ we choose one pair $(x_t,y_t) \in C_1(a_t)$; x is choosen under condition $(x,y) \in D_1$ for the same a_t and for $y \neq y_t$. $M_1 \subseteq D_1$ is the set of all pairs (x,y), realizable with strategy $\tilde{x}_1(y,a)$; d_1 is the guaranteed result for P_1 with this strategy; and $M_1^- \subseteq M_1$ contains exactly these elements for M_1, for which $w_1 = d_1$ for some $a \in A$.

Now, we can divide the set T''_o into two parts: T'_1 and T''_2 (this is a partition of T''_o). Each strategy $\tilde{x}(y,a) \in T'_1$, (in particular $\tilde{x}_1(y,a)$ guarantees d_1 to P1, because such a strategy allows to realize some pair $(x,y) \in M^-_1$. If there exists a strategy which guarantees more to P1, it belongs to the set T''_1.

It is clear, that the pairs rejection can continue considering the set $D_2 = D_1 \setminus M^-_1$, then $D_3 = D_2 \setminus M^-_2$ and so on.

C. Any time when we reject pairs (x,y), we must carry out the computations and verifications, described in parts A and B. The current set T''_i of strategies is divided into two parts: T'_{i+1} and T''_{i+1}. (For the first time this was the set $\tilde{X}$). The guaranteed result for P1 with any $\tilde{x}(y,a) \in T'_{i+1}$ is determined and the analysis continues with the set T''_{i+1}.

The described procedure leads to a sequence of strategies

$\tilde{x}_o(y,a)$, $\tilde{x}_1(y,a)$ ... $\tilde{x}_i(y,a)$ ...

and a corresponding sequence of guaranteed results $d_o, d_1, \ldots d_i \ldots$.
Both sequences are finite. The maximal guaranteed result R for P1 is: $R = \max_i d_i$. The corresponding strategy cguarantees R to P1. As we know $|\tilde{X}| = m^{np}$, but the number of investigated strategies is no greater than m.n.

5. EXAMPLE

Let us consider an example with $|X| = 4$, $|Y| = 3$, $|A| = 5$. Table 1 contains the payoff functions of P1 and P2.

TABLE 1. Payoff functions

		P2					P1				
		a=1	a=2	a=3	a=4	a=5	a=1	a=2	a=3	a=4	a=5
y=1	x=1	13	9	1	12	7	1	6	4	11	1
	x=2	9	1	11	9	9	3	7	1	2	1
	x=3	4	1	7	1	5	9	3	1	8	4
	x=4	1	10	13	5	3	11	1	7	10	8
y=2	x=1	11	6	10	1	8	2	5	1	1	10
	x=2	3	1	3	3	1	1	4	3	1	5
	x=3	1	7	4	4	10	4	1	1	3	9
	x=4	6	3	1	8	11	1	9	6	1	3
y=3	X=1	2	1	9	2	13	7	1	9	6	6
	x=2	1	11	8	1	6	8	1	10	7	11
	x=3	5	2	1	10	2	1	8	2	1	1
	x=4	10	8	5	11	1	5	2	1	4	1

The left half of the table shows the payoff function of P2, and the right one - this function of P1. The first four rows correspond to x = 1,2,3,4 and to y = 1. The second four rows correspond to the same x's and to y = 2, and so on. The first column in the left half and the first column in the right half corresponds to a = 1, the second columns correspond to a = 2, and so on. It can be seen that for each pair (i.e. in each row) there exists some $a \in A$, for which P1 gets minimal payoff, equal to 1. Constructing the described game leads to the following.

P1 cannot permit to P2 the following payoffs when $a_c = 1,2,3$:

$$a_c = \quad 1 \qquad 2 \qquad 3$$
$$f_2 = \quad 13 \qquad 11;10 \qquad 11;10$$

because P1 gets 1 with the corresponding pairs (x,y) at the same a_c. The rejection of corresponding rows leaves only x = 3 admissible for y = 1. In this moment the following strategy (for example) guarantees 2 to P1:

TABLE 2

y \ x \ a	1	2	3	4	5
1	3	3	3	3	3
2	2	3	4	2	3
3	4	4	1	4	1

The players receive (get) the payoffs encircled in the table. If P2 wants to get maximum, he must always choose $y_c = 3$.

It is not very difficult to see that further rejection of pairs does not increase the guaranteed result for P1, i.e. R = 2.

6. CONCLUSION

An algorithm for fast finding of best strategy for the first player in a finite 2-persons game with fixed sequence of moves, with different information and with participation of the Nature is presented. It is possible that the main idea of this algorithm would be useful in the analysis of some more general games.

7. REFERENCES

Burkov, V.N., Enaleev, A.K. (1985). Optimality of the open control principle. Necessary and sufficient conditions for information reliability in active systems. Automat. and Remote Control, No. 3. (in Russian).

Germeier, Yu.V. (1976). Games with non-opposite interests. "Science" ("Nauka"), Moscow, (in Russian).

Kukuschkin, N.S., Morozov, V.V. (1984). Theory of nonantagonistic games. University of Moscow, Moscow, (in Russian).

Mettev, B.S. (1977). Game models of expert investigation (expertise). J. of Comput. Math. and Math. Physics, v. 17, No. 4, (in Russian).

A METHODOLOGICAL APPROACH TO COMPARING PARAMETRIC CHARACTERIZATIONS OF EFFICIENT SOLUTIONS

Andrzej P. Wierzbicki
Institute of Automatic Control, Technical University of Warsaw
Nowowiejska 15/19, 00-665 Warsaw, Poland

INTRODUCTION.

The vector optimization problem considered here is to minimize a continuous vector-valued function $f: S \rightarrow R^m$ on a constraint set $C \subset S$. Let $F = f(C)$ be a compact set (though much weaker assumptions are sufficient for the existence of optimal solutions - see Benson, 1978). While keeping in mind that the set F is usually defined implicitely and that an attainable decision outcome y F means that $y = f(x)$ for some admissible decision $x \in C$, we can restrict the discussion to the outcome or objective space only. We assume that all objectives are minimized and use the notation $D = -R^m_+$ and $\tilde{D} = -R^m_+ \setminus \{0\}$ while int D denotes the interior of $-R^m_+$. Thus, $y' \in y'' + D$ denotes here that $y'_i \leqq y''_i$ for all $i=1,.. m$, while $y' \in y'' + \tilde{D}$, $\tilde{D} = D \setminus \{0\}$, denotes $y'_i \leqq y''_i$ for all $i=1,..$ m and $y'_j < y''_j$ for some $j=1,.. m$, and $y' \in y'' + \text{int } D$ denotes $y'_i < y''_i$ for all $i=1,..$ m where $y + D$ is the cone D shifted by y. The problem of vector minimization of $y = f(x)$ over C can be equivalently stated as the problem of finding D-optimal elements of F. The set of all such elements, defined by:

$$\bar{F} = \{\bar{y} \in F: F \cap (\bar{y} + \tilde{D}) = \emptyset\} \tag{1}$$

is called the efficient set (D-optimal set, Pareto set) in objective or outcome space.

Several other concepts of efficiency are essential for the discussion of characterizations. The weakly efficient elements belong to the set

$$\bar{F}^w = \{\bar{y} \in F: F \cap (\bar{y} + \text{int } D) = \emptyset\} \tag{2}$$

Although important for theoretical considerations, weakly efficient elements are not useful in practical decision support, since there might be too many of them. Another concept is that of properly efficient elements; there are many almost equivalent definitions of such elements, see Sawaragi et al.(1985). We adopt here the definition of Henig (1982) that characterizes properly efficient elements as D'-optimal elements for any cone D' that contains D in its interior

$$\bar{F}^p = \bigcup_{D' \in \mathcal{D}} \{\bar{y} \in F: F \cap (\bar{y} + \tilde{D}') = \emptyset\} \ ; \qquad \mathcal{D} = \{D': D \subset \text{int } D'\} \tag{3}$$

Properly efficient elements have bounded marginal substitution rates that indicate how much one of the objectives must be deteriorated in order to improve another one by a unit. In applications, it is more useful to restrict further the concept of properly efficient elements and consider only such that have marginal substitution rates bounded by some a priori number. This corresponds to the concept of properly efficient elements with bound ε or D_ε-optimal elements that belong to the set

$$\bar{F}^{\varepsilon} = \{\bar{y} \in F:\ F \cap (y + \tilde{D}_\varepsilon) = \emptyset\} \tag{4}$$

where

$$D_\varepsilon = \{y \in R^m:\ \text{dist}(y, D) \leqq \varepsilon \| y \|\} \tag{5}$$

and $\tilde{D}_\varepsilon = D_\varepsilon \setminus \{0\}$, while $\varepsilon > 0$ is some given number - see Wierzbicki (1977); marginal substitution rates are then bounded by ε and $1/\varepsilon$.

PARAMETRIC CHARACTERIZATIONS AND REPRESENTATIONS, THEIR COMPLETENESS AND OTHER PROPERTIES.

In multiobjective optimization, most characterizations - sufficient and necessary conditions - of efficiency are related to the use of some substitute scalarizing function that typically depends not only on the objective values but also on some additional parameters. There are two classes of such parameters that are important for applications in decision support systems: weighting coefficients and objective function levels which can be interpreted as reference, aspiration or reservation levels. Generally, we consider a set A of such parameters α, $A \subset R^m$. Let a substitute scalarizing function be denoted by $s: F \times A \rightarrow R^1$; important examples are (bi)linear functions, norms, so called achievement functions. Such a function should desireably have two basic properties.

(S) The sufficiency property: for each $\alpha \in A_s$

$$\operatorname*{Arg\ min}_{y \in F \cap Y_s(\alpha)} s(y, \alpha) \subset \bar{F} \tag{6}$$

where A_s is a subset of A for which the condition (6) holds and $Y_s(\alpha)$ represents possible additional constraint set. An analogous property could characterize weakly efficient points or properly efficient points. If (S) holds, then a point-to-set mapping of A_s into $\bar{F}$ can be defined:

$$\Psi(\alpha) = \operatorname*{Arg\ min}_{y \in F \cap Y_s(\alpha)} s(y, \alpha) \tag{7}$$

Such a mapping is typically used as a basis of interaction between a decision maker and a decision support system. In such applications, however, we need a single point in the set $\Psi(\alpha)$, that is, a selection $\psi(\alpha) \in \Psi(\alpha)$. The decision maker, called also the user, specifies some $\alpha \in A_s$ and the system responds with an efficient outcome $\bar{y} = \psi(\alpha) \in \bar{F}$; hence, parameters α will be called controlling parameters, while the mapping $\Psi(\alpha)$, or its selection $\psi(\alpha)$, will be called a parametric representation of $\bar{F}$.

However, the scalarizing function $s(y, \alpha)$ that implies a parametric representation should also desireably have the following property:

(N) The necessity property: for each $\bar{y} \in \bar{F}$, there exists $\bar{\alpha} \in A_n$ such that:

$$\bar{y} \in \operatorname*{Arg\ min}_{y \in F \cap Y_n(\bar{\alpha})} s(y, \bar{\alpha}) \tag{8}$$

This property can be also modified for weakly efficient or properly efficient points. Observe that, if $Y_n(\bar{\alpha}) = Y_s(\bar{\alpha})$ and $\bar{\alpha} \in A_s$, then we check both necessary and sufficient conditions when (N) and (S) both hold. The pair of conditions (S) and (N) will be called here a parametric characterization of solutions to multiobjective optimization problems. We shall say that (S) and (N) completely characterize parametrically the efficient set F if $A_s = A_n$, $Y_s(\alpha) = Y_n(\alpha)$ for all $\alpha \in A_s$.

An important aspect of parametric characterizations is their controllability. If a characterization is complete, then the related parametric representation has a specific 'onto' property:

$$\bigcup_{\alpha \in A_s} \Psi(\alpha) = \bar{F} \tag{9}$$

which, in fact, can be taken as a precise definition of the completeness of characterizations. For incomplete characterizations, the equality sign in (9) must be substituted by an inclusion; if (9) holds with a limit or a closure added on the left-hand side, we shall call such characterization almost complete.

Complete or almost complete characterizations provide for a kind of global controllability of the parametric representation by a user: he can reach (almost) all $\bar{y} \in \bar{F}$ by suitably changing α. However, a user of a decision support system needs also local controllability of a parametric representation in the sense of being able to easily and continuously influence his selection of $\bar{y} \in \bar{F}$. This means that the computable selection $\psi(\alpha) \in \Psi(\alpha)$, $\psi: A_s \rightarrow \bar{F}$, should be Lipschitz-continuous:

$$\| \psi(\alpha') - \psi(\alpha'') \| \leq \beta \| \alpha' - \alpha'' \| \quad \text{for all } \alpha', \alpha'' \in A_s \tag{10}$$

which, in turn, necessitates a Lipschitz-continuity of the mapping Ψ. Unfortunately, there are until now very few results on Lipschitz-continuity of parametric characterizations. We give later an example of such result for a simple case; in other cases, intuitive or negative statements can be still made, based on logical evaluation or simple counterexamples.

The sets $\bar{F}^w$ of weakly efficient solutions and $\bar{F}^p$ of properly efficient solutions have several characterizations that are complete, hence globally controllable. Characterizations of the efficient set $\bar{F}$ are either almost complete or they have other drawbacks. The sets A_s and A_n might depend on the set F and thus on computational accuracy; the intersections of F and $Y_s(\alpha)$ or $Y_n(\alpha)$ might become empty by computational inaccuracies; the mathematical operations required in characterizations might be unreasonable from a computational point of view. Thus, we shall say that a characterization of the type (S), (N) is robustly computable if it satisfies the following conditions:

(i) The conditions (S), (N) do not contain additional requirements of m-time repetition of minimization nor requirements of uniqueness of minima.

(ii) The intersection of F with $Y_s(\alpha)$ or $Y_n(\alpha)$ should not become empty when the set F is slightly perturbed. For F of arbitrary, a priori unknown shape this implies that for each $\bar{y} \in \bar{F}$ and the corresponding $\bar{\alpha}$ in (N) or in (S), there must be a neighborhood $U(\bar{y})$ such that $U(\bar{y}) \subset Y_n(\bar{\alpha})$, $U(\bar{y}) \subset Y_s(\bar{\alpha})$. Thus, (S), (N) cannot contain additional constraints that might be active at any $\bar{y} \in \bar{F}$; all such constraints should be included in the form of the function $s(y, \alpha)$, say, by penalty techniques.

Unfortunately, completeness and robust computability of characterizations of efficiency do not coincide, which will be shown later in an impossibility theorem.

A special issue is that of constructive computability of the necessary conditions (N). Some of them specify, in their proof, the value of parameters $\bar{\alpha} \in A_n$ for which these conditions should be checked; such necessary conditions will be called direct. Other asssure us only of the existence of such $\bar{\alpha}$ while searching for this $\bar{\alpha}$ might be computationally cumbersome; such necessary conditions will be called indirect.

Beside controllability and robust or direct computability, there are several other aspects of constructiveness of parametric characterizations and representations of efficient solutions to multiobjective optimization problems that are important in applications in decision support systems. One of such aspects is independence on a priori information. Many characterizations use information about so-called ideal or utopia point. Abstractly, this point is defined as the strict lower bound to the efficient set or as the unique (strong) D-maximal point of the set $\{y \in R^m: F \subset y - D\}$. If $D = -R^m_+$, then the utopia point is the vector composed of results of scalar minimization of each objective function separately. A characterization should not depend on the precise knowledge of the utopia point, because it would not then be robustly computable. As long as only approximate information about the utopia point is required in a characterization, it does not constitute an excessive dependence on a priori information, because the utopia point (or its approximation in cases of large dimensionality corresponding to multiobjective trajectory optimization) can be computed once for entire F.

While the use of approximate bounds to the set F is quite constructive, the requirements of further a priori knowledge of F are not. For example, if a priori knowledge of entire $\bar{F}$ is used in (N), it makes the necessary condition rather useless, since we cannot then apply (N) to check whether y belongs to a priori unknown $\bar{F}$. (N) shall be called tautological in such a case.

Experience in applications of parametric representations in multiobjective optimization and interactive decision support has led most authors to agree more or less explicitely on several further attributes of their constructiveness:

Simplicity. A parametric representation should be concepually simple and easy to grasp mentally.

Generality. A parametric representation should be, if possible, applicable not only to linear and convex problems , but also to nonconvex, discrete and dynamic problems of multiobjective trajectory optimization.

Interpretability of parameters. The parameters in the sets A_s, A_n should have an easy and reasonable interpretation for the user (who needs such an interpretation when changing these parameters in order to control the parametric representation), not for theorists only.

Computability. Beside the requirements of robust computability and directness of necessary conditions, parametric representations should be computable by means of algorithms that do not require excessive computer time and can be relied upon to produce results without the need of adjustment by the user.

ALTERNATIVE CHARACTERIZATIONS AND PARAMETRIC REPRESENTATIONS OF EFFICIENT SOLUTIONS.

There are many characterizations that imply various parametric representations. We shall subdivide them into three classes: (A) those based on weighting coefficients used in (bi)linear functions and various norms; (B) those based on aspiration or reservation levels used in various norms and achievement functions; (C) other possible characterizations. We shall discuss here only the classes (A) and (B); for examples of other possible characterizations see, e.g., Zionts and Wallenius (1976).

(A) Characterizations by weighting coefficients. These characterizations are obtained if α is a vector composed of weighting coefficients α_i used in (bi)linear functions or various norms that scalarize the components y_i of the objective vector. All characterizations in this class have one fundamental disadvantage in common: experience in applications of decision support systems shows that weighting coefficients are not easy to be understood well and interpreted by an average user.

(A1) (Bi)linear functions used as substitute scalarizing functions have the following form:

$$s(y,\alpha) = \sum_{i=1}^{m} \alpha_i y_i \tag{11}$$

with $\alpha = (\alpha_1, \dots \alpha_i, \dots \alpha_m)$; the sets A_s, A_n are defined by:

$$A_s = \left\{ \alpha \in \operatorname{int} R^m_+ : \sum_{i=1}^{m} \alpha_i = 1 \right\}; \qquad A_n = \left\{ \alpha \in R^m_+ : \sum_{i=1}^{m} \alpha_i = 1 \right\} \tag{12}$$

Theorem 1. Let $s(y,\alpha)$, A_s, A_n be defined as above. If $\alpha \in A_s$, then each $\bar{y}$ that minimizes $s(y,\alpha)$ over $y \in F$ is properly efficient, hence efficient. If $\bar{y}$ is efficient or weakly efficient and F is convex, then there exists $\bar{\alpha} \in A_n$ such that $\bar{y}$ minimizes $s(y,\bar{\alpha})$ over $y \in F$. If $\alpha \in A_n$, then each $\bar{y}$ that minimizes $s(y,\alpha)$ over $y \in F$ is weakly efficient. If $\bar{y}$ is properly efficient and F is convex, then there exists $\bar{\alpha} \in A_s$ such that $\bar{y}$ minimizes $s(y,\alpha)$ over $y \in F$.

For the proofs of various parts of this well-known theorem see, for example, Jahn (1985) or Sawaragi et al. (1985); originally, this characterization dates back to Koopmans (1951), Kuhn and Tucker (1951) and Geoffrion (1968). For convex cases, this characterization is complete for weak and proper efficiency and almost complete for efficiency. Moreover, these characterizations are robustly computable but indirect for necessary conditions. They are also independent of a priori information, conceptually simple, rather general (with the restriction of necessary conditions to the convex case) and easily computable for sufficient conditions. The main drawback of these characterizations, beside bad interpretability of weighting coefficients, is the fact that the related parametric representations are not Lipschitz-continuous for such basic cases as when F is a convex polyhedral set.

Similar properties to the above characterizations have those based on the l_1 norm:

$$s(y,\alpha) = \sum_{i=1}^{m} \alpha_i \left| y_i - \tilde{y}_i \right|; \qquad \alpha \in A_s \tag{13}$$

with A_s defined as in (12) and a lower bound point $\tilde{y}$ restricted by the strict lower bound, that is, utopia point $\hat{y}$:

$$\tilde{y} \in \hat{y} + \operatorname{int} D \tag{14}$$

Actually, $\tilde{y} \leqq \hat{y}$ would suffice, but the strong inequality in (14) is assumed to obtain computational robustness. These characterizations will not be discussed separately.

(A2) The <u>weighted</u> l_p <u>norm</u> is also often used as a substitute scalarizing function:

$$s(y,\alpha) = \left[\sum_{i=1}^{m} \alpha_i |y_i - \tilde{y}_i|^p\right]^{1/p}; \qquad \alpha \in A_s \tag{15}$$

with A_s defined as in (12) and $\tilde{y}$ restricted as in (14); the parameter $p \in (1;\infty)$ can be also treated as (m+1)-th component of the parameter vector α.

<u>Theorem 2</u>. Let $s(y,\alpha)$, α, p, $\tilde{y}$, be selected as above. Then each y that minimizes $s(y,\alpha)$ over $y \in F$ is properly efficient. If $\bar{y}$ is efficient, then for each $\varepsilon > 0$ there exist such $\bar{\alpha} \in A_s$, such $p \in (1;\infty)$ and such $\bar{y}'$ with $\|\bar{y}' - \bar{y}\| < \varepsilon$ that $\bar{y}'$ minimizes $s(y,\bar{\alpha})$ over $y \in F$.

This form of this theorem is due to Gearhart (1983); see also Zeleny (1973), Yu and Leitman (1974), Wierzbicki (1977), Salukvadze (1979). This characterization is almost complete for proper efficiency and efficiency also in non-convex cases; in this sense, it is stronger than this by (bi)linear functions.

This characterization is robustly computable, but the necessary condition is indirect. The Lipschitz-continuity of the related parametric representation has not been studied, but we might suspect local controllability. This characterization depends on a priori information, but not excessively and is not tautological. It is not quite simple conceptually, but rather general. The interpretability of the parameter pair (α,p) for an average user is bad; moreover, this representation might be not easily computable if p is very large, since it leads to badly conditioned nonlinear programming problems.

(A3) The <u>weighted</u> l_∞ <u>(Chebyshev)</u> <u>norm</u> is a very useful substitute scalarizing function:

$$s(y,\alpha) = \max_{1 \leq i \leq m} \alpha_i |y_i - \tilde{y}_i|; \qquad \alpha \in A_s \tag{16}$$

where A_s is defined as in (12) and $\tilde{y}$ restricted as in (14).

<u>Theorem 3</u>. Let $s(y,\alpha)$, $\tilde{y}$, A_s be defined as above. Then each $\bar{y}$ that minimizes $s(y,\alpha)$ over $y \in F$ is weakly efficient. If the minimum is unique, then such $\bar{y}$ that minimizes $s(y,\alpha)$ over $y \in F$ is efficient. If $\bar{y}$ is weakly efficient, then there exists such $\bar{\alpha} \in A_s$ that $\bar{y}$ minimizes $s(y,\bar{\alpha})$ over $y \in F$. If $\bar{y}$ is efficient, then there exists such $\bar{\alpha} \in A_s$ that $\bar{y}$ uniquely minimizes $s(y,\bar{\alpha})$ over $y \in F$.

This theorem is due to Dinkelbach (1971) and Bowman (1976). This characterization is complete for weak efficiency and also for efficiency even in a nonconvex case, but at the cost of the requirement of uniqueness and thus loosing robust computability of efficiency conditions. Beside this basic drawback, this characterization depends on a priori information but not excessively and is not tautological, is rather simple conceptually, general, and rather easily computable for weak efficiency.

The basic drawback of all weighting coefficient methods - their bad interpretability - can be overcome in this case by making these coefficients dependent on aspiration or reference levels w_i for objective functions. Under the restriction that $w_i > \tilde{y}_i$, we can take:

$$\alpha_i = [1/(w_i - \tilde{y}_i)] / \sum_{j=1}^{m} 1/(w_j - \tilde{y}_j) \tag{17}$$

which has the interpretation that the closer is an aspiration or reference level w_i to the lower bound level $\tilde{y}_i$, the more important is the objective. When checking necessary conditions in Theorem 3, the application of (22) with $w_i = \bar{y}_i$ makes these conditions direct. This modification has been used by Steuer and Choo (1983), Nakayama (1985), Korhonen and Laakso (1985); however, if aspiration or reference levels are used as the controlling parameters, then the method belongs to another class since the norm (16) changes its form of dependence on controlling parameters and should be interpreted as an achievement function. In this sense, we shall show later that the corresponding parametric representation is Lipschitz-continuous and thus locally controllable.

(A4) A <u>composite norm</u>, in particular - a combination of weighted l_1 and l_∞ norms is one of the strongests substitute scalarizing functions:

$$s(y,\alpha) = \max_{1 \le i \le m} \alpha_i |y_i - \tilde{y}_i| + \alpha_{m+1} \sum_{i=1}^{m} \alpha_i |y_i - \tilde{y}_i| ,$$

$$\alpha \in A_s, \qquad \alpha_{m+1} \in (0;\ 1] \tag{18}$$

where A_s is defined as in (12) and $\tilde{y}$ restricted as in (14).

<u>Theorem 4</u>. Let $s(y,\alpha)$, $\tilde{y}$, α_{m+1} and A_s be defined as above. Then each $\bar{y}$ that maximizes $s(y,\alpha)$ over $y \in F$ is properly efficient; if $\bar{y}$ is properly efficient, then there exists a (sufficiently small) α_{m+1} and $\bar{\alpha} \in A_s$ such that $\bar{y}$ maximizes $s(y,\bar{\alpha})$ over $y \in F$.

This theorem is due to Dinkelbach and Iserman (1973). It completely characterizes proper efficiency without convexity assumptions; since (18) converges to (16) with $\alpha_{m+1} \to 0$, it implies also an almost complete characterization of efficiency. This characterization is robustly computable and its necessary condition becomes direct if we apply (22) with $w_i = \bar{y}_i$ and α_{m+1} smaller than an a priori bound ε for marginal substitution rates.

This characterization depends on a priori information but not excessively and is not tautological. It is not quite simple conceptually but rather general and easily computable. Thus, it might be one of the best characterizations - provided, however, that we use the transformation (17) of weighting coefficients in order to assure easy interpretability and local controllability. This has been used by Lewandowski at al. (1985), although not as a norm but as an achievement function.

(B) <u>Characterizations by objective function levels</u>. These characterizations assume that α is a vector composed of objective function levels, denoted here by w, that are interpreted either as reservations (values that must be achieved), aspirations (values that should be desireably achieved) or reference values (which can be, in fact, interpreted as aspirations). While much better interpretable for an average user than the characterizations by weighting coefficients, most of the characterizations by objective function levels have several disadvantages that can be overcome first by introducing the concept of <u>order-consistent achievement functions</u> - a class that includes functions such as (16) and (18) under the transformation (17) but is much more general.

(B1) <u>Directional search</u>. If a direction $w \in R^m_+$ and the utopia point $\hat{y}$ are given, we can construct a substitute scalarizing function for the directional search:

$$s(y, w, \bar{t}) = \| y - \hat{y} - \bar{t}w \| ; \quad w \in R^m_+ \tag{19}$$

with an arbitrary norm in R^m and with $\bar{t}$ selected as the smallest value of $t \in [0; \infty)$ for which the minimum of $s(y, w, t)$ over $y \in F$ is equal zero. This is actually an additional minimization requirement; moreover, $\hat{y}$ should be known exactly in the corresponding sufficient condition, hence the following incomplete characterization is certainly not robustly computable:

<u>Theorem 5</u>. Let $y \in F$ be efficient and let a lower bound point $\tilde{y} \lneq \hat{y}$ be given. If $w = \bar{y} - \tilde{y}$, then $\bar{t} = 1$ is the lowest value of t such that $y + tw \in F$ and the minimum of the function $s(y, w, t)$ over $y \in F$ is equal zero. If $m = 2$ and F is convex, then, for each $w \in R^m_+$, the smallest value of $t \gneq 0$ such that $\hat{y} + tw \in F$ results in an efficient $\bar{y} \in \bar{F}$. If $m > 2$, counterexamples show that an analogous sufficient condition cannot be proven even under convexity assumptions.

The above theorem is well known, see, for example, Yu and Leitman, (1974). This characterization cannot be used to generate a priori unknown efficient solutions in response to user requirements; it is only a very good tool for checking the efficiency of a given $\bar{y}$.

(B2) <u>Reservation levels or constraints on objective functions</u>. Here several simple substitute functions and constraints are used:

$$s_k(y, w) = y_k; \quad y \in F \cap Y_k(w) \tag{20}$$

where:

$$Y_k(w) = \{ y \in R^m : y_i \leq w_i,\ i = 1, \dots m,\ i \neq k \};$$
$$w = (w_1, \dots w_i, \dots w_m) \in \hat{y} - D \tag{21}$$

<u>Theorem 6</u>. Let $s_k(y, w)$, $Y_k(w)$ be defined as above. If, for some $k = 1, \dots m$, $\bar{y}$ minimizes $s_k(y, w)$ over $y \in F \cap Y_k(w)$ with some $w \in \hat{y} - D$, then $\bar{y}$ is weakly efficient; if $\bar{y}$ is weakly efficient, then there exists $k = 1, \dots m$ such that $\bar{y}$ minimizes $s_k(y, w)$ over $y \in F \cap Y_k(w)$ with $w = \bar{y}$. If, for all $k = 1, \dots m$, $\bar{y}$ minimizes $s_k(y, w)$ over $y \in F \cap Y_k(w)$ with some $w \in \hat{y} - D$, then $\bar{y}$ is efficient; if $\bar{y}$ is efficient, then $\bar{y}$ minimizes $s_k(y, w)$ over $y \in F \cap Y_k(w)$ with $w = \bar{y}$ for all $k = 1, \dots m$. Let F be convex. Then $\bar{y}$ is properly efficient if and only if the problems of minimizing $s_k(y, w)$ over $y \in F \cap Y_k(w)$ are stable, that is, the perturbation functions:

$$\bar{s}_k(w) = \min_{y \in F \cap Y_k(w)} s_k(y, w) \tag{22}$$

are Lipschitz-continuous at $w = \bar{y}$ for all $k = 1, \dots m$.

The proof of this theorem, due to earlier results in Haimes et al.(1975), Changkong and Haimes (1978) can be also found in Sawaragi et al

(1985). This characterization is complete and rather general (valid without any convexity assumptions but not easy to generalize for the case of trajectory optimization). However, it is not robustly computable and only the weak efficiency part of this characterization has found broader applications. Moreover, the proper efficiency part of this characterization is not direct. On the other hand, we can expect local controllability.

The characterization depends on a priori information but not excessively and is not tautological; it conceptually simple and the parameters are easily interpretable as reservation levels for objective values. For the weak efficiency part of this characterization, it is also easy to compute.

Another, early variant of characterization by using reservation levels is related to one of two possible interpretations of goal programming: this of trying to improve given attainable upper bounds or reservations for objective values. Originally suggested by Charnes and Cooper (1961), further developed by Fandel (1972), Ecker and Kuada (1975), it has been studied extensively in various modifications - see Gal (1982). Its prototypical formulation is:

$$s(y,\ w) = \sum_{i=1}^{m} \alpha_i (w_i - y_i); \qquad y \in F \cap Y(w);$$

$$Y(w) = \left\{ y \in R^m:\ y_i \leq w_i,\ i = 1, \ldots m \right\}; \qquad w \in F \tag{23}$$

with some fixed $\alpha \in A_s$ defined as in (12). This gives a complete characterization of efficient solutions:

Theorem 7. Let $s(y,\ w), \alpha,\ Y(w)$ be defined as above with $w \in F$. If $\bar{y}$ maximizes $s(y,\ w)$ over $y \in F \cap Y(w)$, then $\bar{y}$ is efficient; if $\bar{y}$ is efficient and we set $\bar{w} = \bar{y}$, then $\bar{y}$ maximizes $s(y,\ \bar{w})$ over $y \in F \cap Y(w)$.

This theorem is well known, see Charnes and Cooper, (1975) and Gal (1982), but this characterization has a basic drawback: it is not robustly computable. In fact, we use here $F \cap Y(\bar{w}) = \{\bar{y}\}$ in necessary conditions and this singleton set becomes empty by any, however slight, perturbation of F. Except for this essential drawback, this characterization does not depend on a priori information, is simple conceptually, very general (no convexity assumptions are needed and a generalization to multiobjective trajectory optimization is easy), well interpretable and easily computable if we do not come with w too close to F.

The drawbacks of this otherwise excellent class of characterizations could be overcome when substituting constraints by penalty functions - but this leads to the concept of an achievement function. Before adressing this concept, yet another class of characterizations must be considered.

(B3) Aspiration levels with various norms. This class consists of two subclasses. The first subclass, called compromise programming or displaced ideal, see Zeleny (1973), (1982), corresponds to the case where aspiration levels for objective function values are below utopia point and thus far from being attainable, $w \leq \hat{y}$. This is actually the case of classes A2, A3, A4 with the lower bound point $\tilde{y}$ treated as an additional parameter and interpreted as aspiration level point; this case will not be considered here any further. The second subclass corresponds to the second, widely used interpretation of goal programming: this of trying to come close to given aspiration levels or goals which are typically not far from being attainable. In fact, consider formula (15) with another interpretation:

$$s(y,\ w) = \Big[\sum_{i=1}^{m} \alpha_i |w_i - y_i|^p\Big]^{1/p} \tag{24}$$

where $\alpha \in A_s$ is treated not as the controlling parameter but as a constant and w is the controlling parameter instead. The limit case when $p=\infty$ is a form similar to (16). If $p= 1$, we obtain a form similar to (23), however, there is a basic difference: the function above should be minimized and not maximized as it was the case with (23). Theorem 7 implies that one must maximize a norm or a measure of improvement from attainable reservation levels in order to get to the efficient set; from unattainable aspiration levels, however, one must minimize the distance to the attainable and efficient set. Thus, there are two precisely opposite interpretations of goal programming techniques. To distinguish between them, we must have additional means of checking that their boundary - the efficient set - has been crossed.

Theorem 8. Let $s(y, w)$ be defined as above with any $p \in (1; \infty)$, and either F be convex or $w \in \bar{F} + D$. If $\bar{y}$ minimizes $s(y, w)$ over $\bar{y} \in F$ and $\bar{y}_i > w_i$ for all $i= 1,.. m$, then $\bar{y}$ is properly efficient. If $\bar{y}$ is properly efficient, then there exists such $\bar{w}$ with $\bar{w}_i < \bar{y}_i$ for all $i = 1,.. m$ that $\bar{y}$ minimizes $s(y, \bar{w})$ over $y \in F$.

The proof of this complete characterization of proper efficiency is given in Wierzbicki (1986); similar, though not exactly the same results are given in Jahn (1984). The above characterization is tautological for noncovex sets F: if we require $w \in \bar{F} + D$, then we must know $\bar{F}$ a priori and would have much simpler means of checking whether $\bar{y} \in \bar{F}$. For convex compact F, we can use this characterization constructively only in its sufficient part; the necessary conditions are indirect.

On the other hand, goal programming is simple conceptually, easily interpretable and relatively easily computable; therefore, it has been widely used, see Dyer (1972), Charnes and Cooper (1975), Ignizio (1983). The difficulty in reaching efficient solutions if the goal w is attainable can be ignored if goal programming is treated as a tool of supporting strictly satisficing decisions; however, we assume here that a decision support system should not only inform the user that a given goal is attainable, but also propose a corresponding efficient solution. In the terminology of goal programming, the components $|w_i - y_i|$ of the distance function are often called achievement functions (or under-achievement and over-achievement functions, if the sign of $w_i - y_i$ is taken into account). The drawbacks of goal programming suggest that a strengthening of this concept would be useful.

CONCEPTS AND PROPERTIES OF ORDER-CONSISTENT ACHIEVEMENT FUNCTIONS.

When trying to specify a class of characterizations based on objective function levels that would have good properties in applications for decision support, it is essential to choose first appropriate concepts that correspond to the nature of the vector optimization problem. We adress here two such concepts: this of monotonicity, essential for sufficiency parts of characterizations, and that of separation of sets, essential for the necessity parts of characterizations.

The role of monotonicity in vector optimization is explained by the following basic theorem:

Theorem 9. Let a function $r: F \rightarrow R^1$ be strongly monotone, that is, let $y' < y''$ (equivalent to $y' \in y'' + \tilde{D}$) imply $r(y') < r(y'')$. Then each minimal point of this function is efficient. Let this function be strictly - sometimes called weakly - monotone, that is, let $y' \ll y''$ (equivalent to $y' \in y'' + \text{int } D$) imply $r(y') < r(y'')$. Then each minimal point of this function is weakly efficient. Let this function be ε-strongly monotone, that is, let $y' \in y'' + \tilde{D}_\varepsilon$ imply $r(y') < r(y'')$, where D_ε, $\tilde{D}_\varepsilon$ are defined as in (5). Then each minimal point of this function is properly efficient with bound ε.

Various parts of this theorem are well-known - see Yu and Leitman (1974), Wierzbicki (1977), Jahn (1984), Sawaragi et al. (1985), see also Wierzbicki (1986) for the proof of the proper efficiency with bound ε part. Observe that a function constructed with the help of a norm, $r(y) = \| y - \tilde{y} \|$, is strictly monotone for all $y \geqslant \tilde{y}$ if the Chebyshev norm is used and strongly monotone for all $y \geqslant \tilde{y}$ if any other norm is used; a composite norm of the form (18) where $r(y) = s(y, \alpha)$ with some $\alpha \in A_s$ is ε-strongly monotone for all $y > \tilde{y}$ if ε is sufficiently small when compared to α_{m+1}.

The second concept, that of separation of sets, is actually used implicitely or explicitely whenever necessary conditions of scalar or vector optimality are derived. We say that a function $r: R^m \rightarrow R^1$ strongly separates two disjoint sets Y_1 and Y_2 in R^m, if there is such $\beta \in R^1$ that $r(y) < \beta$ for all $y \in Y_1$ and $r(y) \geqslant \beta$ for all $y \in Y_2$. Since the definition of efficiency (1) requires that the sets F and $\bar{y} + \tilde{D}$ are disjoint (or F and $\bar{y}$+ int D for weak efficiency, or F and $\bar{y} + \tilde{D}_\varepsilon$ for proper efficiency with bound), they can be separated by a function. If F is convex, these sets can be separated by a linear function of the form (11); this separation of sets is precisely the primal concept beyond the dual concept of weighting coefficients. If F is not convex, the sets F and $\bar{y} + \tilde{D}$ could be stil separated at an efficient point $\bar{y}$, but we need for this a nonlinear function with level sets $\{y \in R^m: r(y) \leqslant \beta\}$ which would closely approximate the cone $y + D$. There might be many such functions; we shall define first their desireable properties and then give several examples of them.

(B4) Order-representing achievement functions are defined generally as such continuous functions $s: F \times W \rightarrow R^1$ that $s(y, w)$ is strictly monotone (see Theorem 9) as a function of $y \in F$ for any $w \in W$ and, moreover, posseses the following property of order representation:

$$\{y \in R^m: s(y, w) < 0\} = w + \text{int } D, \text{ for all } w \in W; \tag{25}$$

which implies, together with the continuity of $s(y, w)$, that:

$$s(y, w) = 0 \text{ for all } w = y \in F \tag{26}$$

Here we assume $W = R^m$ or any reasonably large subset of R^m containing F or, at least, $\bar{F}^w$; the controlling parameter w is interpreted as aspiration level point that might be attainable or not. A simple example of such function is:

$$s(y, w) = \max_{1 \leqslant i \leqslant m} \alpha_i (y_i - w_i) \tag{27}$$

with $W = R^m$ and some fixed $\alpha \in A_s$. Other examples of order-representing functions will be given later. At any weakly efficient point $\bar{y}$, an order-representing function strictly separates the sets $\bar{y}$+ int D and F.

However, an order-representing function cannot be strongly monotone, since it could not be continuous in such a case.

(B5) <u>Order-approximating achievement functions</u> are defined generally as such continuous functions $s: F \times W \rightarrow R^1$ that $s(y, w)$ is strongly monotone (see Theorem 9) as a function of $y \in F$ for any $w \in W$ and, moreover, posseses the following property of order aproximation:

$$w + D_{\bar{\varepsilon}} \subset \{y \in R^m: s(y, w) \leqslant 0\} \subset w + D_{\varepsilon}, \text{ for all } w \in W; \tag{28}$$

with some small $\varepsilon > \bar{\varepsilon} \geqslant 0$, for some reasonably large set W containig F or, at least, $\bar{F}$; the requirement (28) implies also (26). A simple example of order-approximating function is:

$$s(y, w) = \max_{1 \leqslant i \leqslant m} \alpha_i (y_i - w_i) + \alpha_{m+1} \sum_{i=1}^{m} \alpha_i (y_i - w_i) \tag{29}$$

with $W = R^m$ and some $\alpha_{m+1} > 0$ that is sufficiently small as compared to ε and large as compared to $\bar{\varepsilon}$; this function is not only strongly monotone, but also $\bar{\varepsilon}$-strongly monotone. Other examples of order-approximating functions will be given later. At any point $\bar{y}$ that is properly efficient with bound ε, an order-approximating function strictly separates the sets $y + \tilde{D}_{\bar{\varepsilon}}$ and F.

Order-representing and order-approximating functions are jointly called <u>order-consistent achievement functions</u>. When the concepts of monotonicity and separation of sets are used, the following theorem that characterizes efficient solutions by minima of order-consistent functions might appear simple to the point of triviality; but this is precisely the power of arguments based on separation of sets that they simplify complex problems.

<u>Theorem 10</u>. Let $s(y, w)$ be an order-representing function. Then, for any $w \in W$, each point that minimizes $s(y, w)$ over $y \in F$ is weakly efficient; if $\bar{y}$ is weakly efficient (or efficient), then the minimum of $s(y, w)$ with $w = \bar{y}$ over $y \in F$ is attained at $\bar{y}$ and is equal zero. Let $s(y, w)$ be an order-aproximating function with some $\varepsilon, \bar{\varepsilon}$ as in (28). Then, for any $w \in W$, each point that minimizes $s(y, w)$ over $y \in F$ is efficient; if $\bar{y}$ is properly efficient with bound ε (D_{ε}-optimal), then the minimum of $s(y, \bar{w})$ with $\bar{w} = \bar{y}$ over $y \in F$ is attained at $\bar{y}$ and is equal zero. Let, in addition, $s(y, w)$ be $\bar{\varepsilon}$-strongly monotone in y; then each point that minimizes $s(y, w)$ over $y \in F$ is properly efficient with bound $\bar{\varepsilon}$.

For proofs of various parts of this theorem, also for infinite-dimensional normed spaces, see Wierzbicki (1977), (1980), (1982), (1986). Classes (B4, B5), without any convexity assumptions nor restrictions on controlling parameters w, completely characterize weakly efficient elements and almost completely characterize properly efficient and efficient elements (if we take the closure of sets of maximal points of an order-approximating achievement function as $\varepsilon \rightarrow 0$). By adding the requirement of uniqueness of minima in Theorem 10, we could make this characterization complete also for efficient solutions, but we forego this generalization because it would mean the loss of robust computability. The requirement that $w = \bar{y}$ in necessary conditions is not tautological, if we want to use these conditions to check the efficiency of a given element: it is direct and robustly computable, since we do not assume any a priori knowledge of F, nor do we limit the minimization to a single point.

These characterizations are not quite simple conceptually, but the controlling parameters w and the values of the achievement function $s(y, w)$ are very well interpretable: while w is interpreted as aspiration levels, the sign of the minimum of achievement function indicates whether these aspirations are attainable or not, and the value zero of this minimum indicates that aspirations are attainable and efficient. These characterizations are also very general, valid not only for nonconvex and discrete or integer cases, but also easy to extend for problems of multiobjective trajectory minimization - see Wierzbicki (1980). Computationally, their applications are either simple - if F is a convex polyhedral set, then the problem of minimizing (27) or (29) can be rewritten as a linear programming problem - or more complicated for nonlinear or nonconvex problems. In such cases, we must either represent (27), (29) by additional constraints, or apply nondifferentiable optimization techniques, since the definitions of order-consistent achievement functions imply their nondifferentiability at $y = w$.

These characterizations are also, most probably, locally controllable; before establishing Lipschitz-continuity of a parametric representation corresponding to the simple achievement function (27) we must, however, indicate the use of order-consistent functions for checking the uniqueness of minima. The concept of separation of sets used in Theorem 10 implies the following corollary:

<u>Corollary</u>. If $\bar{y}$ is a minimal point of an $\tilde{\varepsilon}$-strongly monotone order-approximating function $s'(y, w)$ over $y \in F$ with any $w \in W$, then $\bar{y}$ is also the unique minimal point of an order-representing function $s''(y, \bar{w})$ with $\bar{w} = \bar{y}$ over $y \in F$.

This corollary is an immediate consequence of the separation of the sets $y + \tilde{D}$ and F by the cone $y + D_{\tilde{\varepsilon}}$. On one hand, this confirms only an easy theoretical conclusion that an order-representing function has unique minima at all properly efficient points. On the other hand, however, the corollary gives a constructive computational way of checking the uniqueness of minima of an order-representing function.

If $\bar{y}$ is, for example, a minimal point of function (27), we can take function (29) with some small α_{m+1} and $w = \bar{y}$ and minimize the latter function; if we obtain the same result of this second minimization, we are sure that the minimum of the former function is unique. This applies, however, only to order-consistent functions in multiobjective minimization, and is by no means a general way of checking the uniqueness of minima of other functions, for which task we do not have constructive computational methods.

The above corollary explains also why we can use rather strong assumptions in the following theorem.

<u>Theorem 11</u>. Let the order-representing function $s(y, w)$ be defined as in (27) and consider the set W of such $w \in R^m$ that the minima of this function are properly efficient elements of F, that is, are unique. Then the parametric representation:

$$\bar{y} = \Psi(w) = \arg\min_{y \in F} s(y, w) \tag{30}$$

is Lipschitz-continuous with the Lipschitz constant 4, that is, $\|\Psi(w') - \Psi(w'')\| \leqslant 4\,\|w' - w''\|$ for all $w', w'' \in W$ and for the Chebyshev norm which implies also Lipschitz-continuity in any other norm in R^m.

Finally, next theorem explains the impossibility of complete and robustly computable characterization of efficient elements $\bar{y} \in \bar{F}$.

Theorem 12. Let $s: F \times A \rightarrow R^1$ be a continuous substitute scalarizing function for vector minimization problems over an arbitrary set $F \subset R^m$.

(a) Suppose that for each efficient $\bar{y} \in \bar{F}$ there exists an $\bar{\alpha} \in A_n \subset A$ such that $\bar{y}$ is a minimal point of $s(y, \bar{\alpha})$ over $y \in F \cap Y(\bar{\alpha})$, where $Y(\bar{\alpha})$ is an additional constraint set, and that each minimal point of $s(y, \alpha)$ over $y \in F \cap Y(\alpha)$ is weakly efficient for any $\alpha \in A_s \subset A$; let $A_s \cap A_n \neq \emptyset$. If, for each $\bar{y} \in \bar{F}$ and the corresponding $\bar{\alpha} \in A_n$, the set $Y(\bar{\alpha})$ contains a neighborhood $U(\bar{y})$ of $\bar{y}$, then the function $s(y, \alpha^n)$ has the following property of local order-representation:

$$\{y \in U(\bar{y}): s(y, \alpha) < s(\bar{y}, \alpha)\} = (\bar{y} + \text{int } D) \cap U(\bar{y}) \text{ for all } \alpha \in A_n \cap A_s \tag{31}$$

(b) If a continuous function $s(y, \alpha)$ has the property (31) then, for sets F of arbitrary form, there exist minimal points $\bar{y}'$ of this function over $y \in F \cap U(\bar{y})$ that are weakly efficient but not efficient; hence, a complete characterization of efficiency by minimal points of such a function is impossible, if we do not apply additional conditions of uniqueness or repetitive minimization.

For proofs of Theorems 11, 12 see Wierzbicki (1986).

FURTHER EXAMPLES OF ORDER-CONSISTENT ACHIEVEMENT FUNCTIONS.

The definitions of order-consistent functions do not require that all level sets of the function s(y, w) should represent or approximate order; only the zero level set should have this property. Hence, there are many examples of order-consistent functions.

A general form of an order-representing function can be written as follows:

$$s(y, w) = \begin{cases} -v(w - y), & \text{if } y - w \in D \\ \text{dist } (y - w, D), & \text{if } y - w \notin D \end{cases} \tag{32}$$

where $v: R^m \rightarrow R^1$ is a strongly monotone value (or utility) function with the property that $v(w - y) = 0$ for all $y - w \in D \setminus \text{int } D$, and any norm in R^m can be taken to define the distance. If we take a multiplicative form of v - for example, the Nash (1950) compromise function - and use the norm l_p with $p \geq 2$, then the function s(y, w) is differentiable except for $y - w \in D \setminus \text{int } D$:

$$s(y, w) = \begin{cases} -\prod_{i=1}^{m} (w_i - y_i), & \text{if } y - w \in D \\ \left[\sum_{i=1}^{m} (y_i - w_i)_+^p\right]^{1/p}, & \text{if } y - w \notin D \end{cases} \tag{33}$$

where $(y_i - w_i)_+ = \max(0, y_i - w_i)$. Another form of order-representing function is piece-wise linear and can be interpreted as an exact internal penalty function for the characterization (23) of efficient solutions:

$$s(y, w) = \max \left(\rho \max_{1 \leq i \leq m} (y_i - w_i), (1/m)\sum_{i=1}^{m}(y_i - w_i)\right) \tag{34}$$

where the function is determined by the sum only for such $y - w \in D$ that

$$(1/m) \sum_{i=1}^{m} (y_i - w_i) > \rho \max_{1 \le i \le m} (y_i - w_i).$$

The above function is useful when applied to linear vector optimization problems, where F is a convex polyhedral set. In such cases, we rewrite the problem of minimizing (34) by using additional variables $z_i = y_i - w_i$, $i = 1,.. m$, $z_{m+1} = s(y, w)$, to the following form:

$$s(y, w) = z_{m+1}, \quad z_{m+1} \geq (1/m)\sum_{i=1}^{m} z_i, \quad z_{m+1} \geq \rho z_i, \; i=1,.. m \tag{35}$$

This function has been used in the DIDAS system of decision support - see Kallio et al. (1980), Lewandowski et al (1982), Grauer et al (1984). Similar transformations are possible for all convex or convex-like - see Jahn (1984) - piece-wise linear functions s(y, w), such as (27), (29) or their further modifications given below.

The prototype order-representing function (27) has also several modifications in cases when additional information about F is available. Suppose we know not only a lower bound $\tilde{y} < \hat{y}$, but also an upper bound $\tilde{\tilde{y}}$ such that F or $\bar{F}^w$ is contained in $\tilde{\tilde{y}} + D$. Thus, for each objective variable y_i, a reasonable but not necessarily tight lower bound $\tilde{y}_i$ and upper bound $\tilde{\tilde{y}}_i$ are known. Suppose a user of a decision support system wants to control his selection of efficient solutions by two parameters: his aspiration levels w' and his reservation levels w", where $\tilde{y}_i < w'_i < w''_i < \tilde{\tilde{y}}_i$ for all $i = 1,... m$. When assuming a satisficing behavior, we can use the fuzzy set theory and membership functions to describe the satisfaction of the user - see, e.g., Sakawa (1983). The membership functions describing the satisfaction of achieving individual objectives can be postulated, for example, in the form:

$$\mu_i = \begin{cases} 0, & \text{if } w''_i < y_i \leq \tilde{\tilde{y}}_i \\ (w''_i - y_i)/(w''_i - w'_i), & \text{if } w'_i \leq y_i \leq w''_i \\ 1, & \text{if } \tilde{y}_i \leq y_i < w'_i \end{cases} \tag{36}$$

and the aggregate mambership function can be taken as the minimum of the component membership functions. However, when maximizing such satisfaction function, we must be certain that the reservation levels are attainable and the aspiration levels are not attainable, which is equivalent to a priori knowledge of $\bar{F}$; otherwise, the maximization of this function would not necessarily result in efficient solutions and only strictly satisficing solutions would be obtained.

This drawback can be overcome if we forego the strict adherence to fuzzy set theory and satisficing behavior. Instead, we postulate a <u>quasisatisficing behavior</u> - see Wierzbicki (1985): a decision maker, aware of his objectives and their scale $(\tilde{y}; \tilde{\tilde{y}})$, is quasisatisficing if he optimizes when his resevation or aspiration levels are not yet attained, but he can further optimize or forego the optimization for additional good reasons if his aspiration levels are attained. In any case, the decision to forego optimization should be reserved for the user and a decision support system should always propose to him efficient solutions, consistently related to his aspiration and reservation levels but obtained through optimization; such efficient solutions will be called <u>quasisatisficing solutions</u>.

In order to characterize and compute quasisatisficing solutions, we have to extend monotonically the functions μ_i also for $w''_i < y_i \leqslant \tilde{\tilde{y}}_i$ and $\tilde{y}_i \leqslant y_i < w'_i$, while changing their signs because we consider minimization problems here:

$$\mu_i(y,w',w'') = \begin{cases} \gamma((y_i - \tilde{\tilde{y}}_i)/(\tilde{\tilde{y}}_i - w''_i) + 1), & \text{if } w''_i < y_i \leqslant \tilde{\tilde{y}}_i \\ (y_i - w''_i)/(w''_i - w'_i), & \text{if } w'_i \leqslant y_i \leqslant w''_i \\ \beta(y_i - w'_i)/(w'_i - \tilde{y}_i) - 1, & \text{if } \tilde{y}_i \leqslant y_i < w'_i \end{cases} \tag{37}$$

where β, γ are given positive parameters; we have $\mu_i = \gamma > 0$ if $y_i = \tilde{\tilde{y}}_i$, $\mu_i = 0$ if $y_i = w''_i$, $\mu_i = -1$ if $y_i = w'_i$, $\mu_i = -(1+\beta)$ if $y_i = \tilde{y}_i$, whereas $\beta = 0$ corresponds to foregoing minimization if $y_i \leqslant w'_i$. The function:

$$s(y, w, w'') = \max_{1 \leqslant i \leqslant m} \mu_i(y, w, w'') \tag{38}$$

is an order-representing achievement function, since it is strictly monotone and has the property (25) if we interprete w'' as the main controlling parameter and w' as an auxiliary parameter. A special case of this function, obtained when $w' \rightarrow y$ (we cannot let $w' \rightarrow w''$, because the function would become discontinuous) and without using the knowledge of $\tilde{\tilde{y}}$:

$$s(y, w) = \max_{1 \leqslant i \leqslant m} (y_i - w_i)/(w_i - \tilde{y}_i) = \max_{1 \leqslant i \leqslant m} (y_i - \tilde{y}_i)/(w_i - \tilde{y}_i) - 1 \tag{39}$$

can be interpreted as a generalization of the Raiffa-Kalai-Smorodinsky compromise solution - see Luce and Raiffa (1957), Kalai and Smorodinsky (1975) - which is obtained by directional search in the direction $\tilde{y}_i - w_i$; minimization of (39) instead gives more robust results and is applicable also to nonconvex problems. Observe that (39) can be either interpreted as an order-representing function or as the Chebyshev norm (16) with weighting coefficients (17). The latter interpretation was used by Steuer and Choo (1983), Nakayama (1985), Sawaragi et al. (1985); however, the proofs of characterization properties of this function become much simpler if based on its order-representing properties and Theorem 10.

Order-approximating achievement functions can be obtained from order-representing functions by adding linear terms. For example, function (38) can be made order-approximating by modifying its form to:

$$s(y,w',w'') = \Big[\max_{1 \leqslant i \leqslant m} \mu_i(y,w',w'') + (\varrho/m)\sum_{i=1}^{m} \mu_i(y,w',w'')\Big]/(1 + \varrho) \tag{40}$$

where $\varrho \in (0; m)$. A similar modification of (39) leads to a form which can be interpreted as a transformation of the composite norm (18):

$$s(y, w) = \max_{1 \leqslant i \leqslant m}(y_i - w_i)/(w_i - \tilde{y}_i) + (\varrho/m)\sum_{i=1}^{m}(y_i - w_i)/(w_i - \tilde{y}_i) \tag{41}$$

This form has been used in Lewandowski et al. (1985) for evaluating discrete alternatives.

Extensions of the concept of order-consistent achievement functions to multiobjective trajectory optimization lead also to many forms of achievement functions - see Wierzbicki (1980). Another extension is to weaken the requirement of order-approximation (28) in such a way that it would admit functions that are differentiable even at y = w (at the cost of the strength and clarity of neccessary conditions but preserving sufficiency). An example of such functions is a simple transformation of the l_p norm:

$$s(y, w) = \left[1/m\sum_{i=1}^{m}\left|(y_i - \tilde{y}_i)/(w_i - \tilde{y}_i)\right|^p\right]^{1/p} - 1 \qquad (42)$$

which has the zero-level set that approximates w+D rather closely but differentiably for sufficiently large p. Various penalty functions lead to other forms of smooth order-approximating achievement functions , see Wierzbicki (1975), (1978).

REFERENCES

Benson, H.P. (1978) Existence of efficient solutions for vector-maximum problems. JOTA 26: 569-580.

Bowman, V.J. Jr. (1976) On the relationship of the Chebyshev norm and efficient frontier of multiple-criteria objectives. In: Thiriez, H., Zionts, S. (eds.) Multiple criteria decision making. Springer, Berlin Heidelberg New York (Lecture Notes in Economic and Mathematical Systems 130).

Changkong, V., Haimes, Y.T. (1978) The interactive surrogate worth trade-off (ISTW) for multiobjective decision making. In: Zionts, S. (ed.) Multiple Criteria Problem Solving. Springer, Berlin Heidelberg New York (Lecture Notes in Economic and Mathematical Systems 155).

Charnes, A., Cooper, W. (1961) Management models and industrial applications of linear programming. Wiley, New York.

Charnes, A., Cooper, W. (1975) Goal programming and multiple objective optimization. J. Oper. Res. Soc. 1: 39-54.

Dinkelbach, W. (1971) Uber einen Losungsansatz zum Vectormaximumproblem. In: Beckman, M. (ed) Unternehmungsforschung Heute. Springer, Berlin Heidelberg New York (Lecture Notes in Operational Research and Mathematical Systems 50: 1-30).

Dinkelbach, W., Iserman, H. (1973) On decision making under multiple criteria and under incomplete information. In: Cochrane, J.L., Zeleny, M. (eds.) Multiple criteria decision making. University of South Carolina Press, Columbia, South Carolina.

Dyer, J.S. (1972) Interactive goal programming. Management Science 19: 62-70.

Ecker, J.G., Kouada, I.A. (1975) Finding efficient points for linear multiple objective programs. Mathematical Programming 8: 375-377.

Fandel, G. (1972) Optimale Enscheidung bei mehrfacher Zielsetzung. Springer, Berlin Heidelberg New York (Lecture Notes in Economic and Mathematical Systems 76).

Gal, T. (1982) On efficient sets in vector maximum problems - a brief survey. In: Hansen, P. (ed.) Essays and surveys on multiple criteria decision making. Proceedings, Mons 1982. Springer, Berlin Heidelberg New York (Lecture Notes in Economic and Mathematical Systems 209).

Gearhart, W.B. (1983) Characterization of properly efficient solutions by generalized scalarization methods. JOTA 41: 618-630.

Geoffrion, A.M. (1968) Proper efficiency and the theory of vector optimization. J. Math. Anal. Appl. 22: 618-630.

Grauer, M., Lewandowski, A., Wierzbicki, A.P. (1984) DIDAS: Theory, implementation and experiences. In: Grauer, M., Wierzbicki, A.P. (eds.) Interactive decision analysis. Springer, Berlin Heidelberg New York Tokyo (Lecture Notes in Economic and Mathematical Systems 229).

Haimes, Y.Y., Hall, W.A., Freedman, H.B. (1975) Multiobjective optimization in water resources systems, the surrogate trade-off method. Elsevier Scientific, New York.

Henig, M.I. (1982) Proper efficiency with respect to cones. JOTA 36: 387-407.

Ignizio, J.P. (1983) Generalized goal programming. Comp. Oper. Res. 10: 277-291.

Jahn, J. (1984) Scalarization in vector optimization. Mathematical Programming 29: 203-218.

Jahn, J. (1985) Some characterizations of the optimal solutions of a vector optimization problem. OR Spectrum 7: 7-17.

Kalai, E., Smorodinsky, M. (1975) Other solutions to Nash's bargaining problem. Econometrica 43: 513-518.

Kallio, M., Lewandowski, A., Orchard-Hays, W. (1980) An implementation of the reference point approach for multiobjective optimization. WP-80-35, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Korhonen, P., Laakso, J. (1985) A visual interactive method for solving the multiple criteria problem. European Journal of Operational Research, to appear.

Koopmans, T.C. (1951) Analysis of production as an efficient combination of activities. In: Koopmans, T.C. (ed.) Activity analysis of production and allocation. Yale University Press, New Haven.

Kuhn, H.W., Tucker, A.W. (1951) Nonlinear programming. In: Neyman, J. (ed.) Proceedings of the 2-nd Berkeley Symposium on Mathematical Statistics and Probability.

Lewandowski, A., Grauer, M. (1982) The reference point optimization approach - methods of efficient implementation. WP-82-019, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Lewandowski, A., Toth, F., Wierzbicki, A. (1985) A prototype selection committee decision support system - implementation, tutorial example and user's manual. Mimeograph. International Institute for Applied Systems Analysis, Laxenburg, Austria.

Luce, R.D., Raiffa, H. (1957) Games and decisions. Wiley, New York.

Nakayama, H. (1985) On the components in interactive multiobjective programming methods. In: Grauer, M., Thompson, M., Wierzbicki, A.P. (eds.) Plural rationality and interactive decision processes. Springer, Berlin Heidelberg New York Tokyo (Lecture Notes in Economic and Mathematical Systems 248).

Nash, J.F. (1950) The bargaining problem. Econometrica 18: 155-162.

Sakawa, M. (1983) Interactive fuzzy decision making for muliobjective nonlinear programming problems. In: Grauer, M., Wierzbicki, A.P. (eds.) Interactive decision analysis. Springer, Berlin Heidelberg New York Tokyo (Lecture Notes in Economic and Mathematical Systems 229).

Salukvadze, M.E. (1979) Vector-valued optimization problems in control theory. Academic Press, New York.

Sawaragi, Y., Nakayama. H., Tanino, T. (1985) Theory of multiobjective optimization. Academic Press, New York.

Steuer, R.E., Choo, E.V. (1983) An interactive weighted Chebyshev procedure for multiple objective programming. Mathematical Programming 26: 326-344.

Wierzbicki, A.P. (1975) Penalty methods in solving optimization problems with vector performance criteria. Working Paper of the Institute of Automatic Control, Technical University of Warsaw (presented at the VI-th IFAC World Congress, Cambridge, Mass.).

Wierzbicki, A.P. (1977) Basic properties of scalarizing functionals for multiobjective optimization. Mathematische Operationsforschung und Statistik, s. Optimization, 8: 55-60.

Wierzbicki, A.P. (1978) On the use of penaly functions in multiobjective optimization. In: Oettli, W., Steffens, F. et al. (eds.) Proceedings of the III-rd Symposium on Operations Research, Universitat Mannheim. Athenaum.

Wierzbicki, A.P. (1980) Multiobjective trajectory optimization and model semiregularization. WP-80-181, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Wierzbicki, A.P. (1982) A mathematical basis for satisficing decision making. Mathematical Modelling 3: 391-405.

Wierzbicki, A.P. (1983) Negotiation nad mediation in conflicts I: The role of mathematical approaches and methods. In: Chestnut, H. et al. (eds.) Supplemental ways for improving international stability. Pergamon Press, Oxford.

Wierzbicki, A.P. (1983) Negotiation nad mediation in conflicts II: Plural rationality and interactive decision processes. In: Grauer, M., Thompson, M., Wierzbicki, A.P. (eds.) Plural rationality and interactive decision processes. Springer, Berlin Heidelberg New York Tokyo (Lecture Notes in Economic and Mathematical Systems 248).

Wierzbicki, A.P. (1986) On the completeness and constructiveness of parametric characterizations to vector optimization problems (extended version of this paper). OR Spektrum, to appear.

Yu, P.L., Leitmann, G. (1974) Compromise solutions, domination structures and Salukvadze's solution. JOTA 13: 362-378.

Zeleny, M. (1973) Compromise programming. In: Cochrane, J.L., Zeleny, M. (eds.) Multiple criteria decision making. University of South Carolina Press, Columbia, South Carolina.

Zeleny, M. (1982) Multiple criteria decision making. McGraw-Hill, New York.

Zionts, S., Wallenius, I. (1976) An interactive programming method for solving the multiple criteria problem. Management Science 22: 652-663.

ON SCALARIZING METHOD IN VECTOR OPTIMIZATION

Dinh The Luc

1. INTRODUCTION

Let K be a closed, convex and pointed cone with a nonempty interior int K in an n-dimensional Euclidean space R^n. The cone induces preference orders in the space as follows: $a,b \in R^n$,

$a \geqq b$ if $a - b \in K$,

$a \geq b$ if $a - b \in K \setminus \{0\}$,

$a > b$ if $a - b \in \text{int } K$.

Further, let X and X' be two nonempty subsets of R^m, let f be a vector valued function from X into R^n and let f' be a scalar function on X'. We have a vector optimization problem associated with (X,f):

$$\begin{array}{ll} \text{Max } f(x) & \\ \text{s.t. } x \in X, & \qquad (I) \end{array}$$

and a scalar optimization problem associated with (X',f'):

$$\begin{array}{ll} \max f'(x) & \\ \text{s.t. } x \in X'. & \qquad (II) \end{array}$$

Denote the sets of optimal solutions and weak optimal solutions of (I) by E(X,f) and WE(X',f'), respectively, i.e.

$E(X,f) = \{x \in X\colon \text{ no } y \in X \text{ satisfies } f(y) \geq f(x)\}$,

$WE(X,f) = \{x \in X\colon \text{ no } y \in X \text{ satisfies } f(y) > f(x)\}$.

The set of optimal solutions of (II) will be denoted by E(X',f').

A common method of solving (I) is to convert it to a scalar problem of type (II) (see Jahn (1984), Wierbicki (1985) and the cited references there). If the set of alternatives X' and the objective function f' are suitably chosen, then we may have:

$E(X',f') \subseteq E(X,f)$ or $E(X',f') \subseteq WE(X,f)$,

i.e. it is possible to find optimal solutions of (I) by solving scalar problem (II). This is very useful for the decision makers who deal with vector problems because for scalar problems the theory and computational algorithms are widely developed. The choice of X' and f' depends on the vector problem to solve, also on other requirements of the decision makers. The following requirements are considered basic:

(i) The invariance of the alternative set:

$X = X'$ (or at least $X' \subseteq X$),

(II) The reservation of the preference orders:

$x,y \in X$ and $f(x) \geqq f(y)$ imply $f'(x) \geqq f'(y)$,

x,y X and $f(x) \geq f(y)$ imply $f'(x) > f'(y)$

(for the weak case $f(x) > f(y)$ implies $f'(x) > f'(y)$).

Some other requirements such as the reservation of linearity, convexity (i.e. if (I) is linear or convex, then (II) is linear or convex, respectively) are also of interest. In this paper, we pay our attention to the scalarizing problems which satisfy (i) and (ii). In Section 2 we give some separation theorems by monotonic functions as they have a close relation to those functions which yield (ii). In Section 3, the results of Section 2 will be applied to the getting of scalarizations of vector problems.

2. SEPARATION BY MONOTONIC FUNCTIONS

Definition 1. A function g from a subset A of R^n into R is said to be monotonic on A (sometimes it is said to be monotonic with respect to K) if for $a,b \in A$,

$a \gtrneqq b$ implies $g(a) > g(b)$,

and it is said to be weakly monotonic if for a,b A,

$$a \geqq b \text{ implies } g(a) \geqq g(b), \tag{1}$$

$$a > b \text{ implies } g(a) > g(b). \tag{2}$$

Definition 2. Let A and B be two nonempty subsets of R^n and let g be a scalar function on R^n. We say that A and B are separated by g if $g(a) \geqq g(b)$ for every $a \in A$ and $b \in B$.

Theorem 1. Let $p \in R^n$. Then $(p + \text{int } K) \cap A = \emptyset$ if and only if there exists a continuous, weakly monotonic function on R^n so that $p + \text{int } K$ and A are separated by it.

Proof. Assume that g is a weakly monotonic function separating $p + \text{int } K$ and A and suppose to the contrary that $(p + \text{int } K) \cap A \neq \emptyset$, i.e. $p + b \in A$ for some $b \in \text{int } K$. By Definition 2, we have

$$g(p) \geqq g(a) \quad \text{for all } a \in A. \tag{3}$$

However, it follows from the weak monotonicity of g that

$$g(p + b) > g(p).$$

This contradicts (3) as $p + b \in A$.

Conversely, assume $(p + \text{int } K) \cap A = \emptyset$. Take a vector $e \in \text{int } K$ and consider a function g given by the relation

$$g(x) = \inf(t:\ x \geqq p + te\ ),\ x \in R^n. \tag{4}$$

We prove that g is well defined and continuous, weakly monotonic on R^n. Indeed, since $e \in \text{int } K$, for every $x \in R^n$, the intersection of $(x + [e])$ with $(p + K)$ is nonempty, where $[e]$ denotes the set $(te:\ t \geqq 0)$. This ensures the existence of t such that $x \geqq p + te$. Moreover, as K is pointed, when a number T being large enough, $x - te \notin p + K$ each $t \geqq T$. Consequently, the infimum of (4) exists and $g(.)$ is well defined. Its continuity is easily established by a direct verification. Now, for the weak monotonicity, let $x,y \in R^n$ with $x \geqq y$. We see that $y \geqq p + te$ implies $x \geqq p + te$. Hence $g(x) \geqq g(y)$. Further, let $x > y$, then there is a positive ε so that $x - \varepsilon e \geqq y$. By the definition of g we have

$$g(x - \varepsilon e) = g(x) - \varepsilon .$$

Thus, $g(x) \geqq g(y) + \varepsilon$ and (2) holds. The proof is complete.

Let $p \in R^n$, $t \geqq 0$. Denote

$B(p,t) = \{x \in R^n : \|x - p\| \leqq t\}$, and

let cone(A,p) be the cone generated by A - p and let clcone(A,p) be its closure.

Lemma 1 . Suppose that $\mathrm{clcone}(A,p) \cap K = \{0\}$. (5)
Then there exists a convex, closed and pointed cone C so that

(iii) $K\setminus\{0\} \subseteq \mathrm{int}\, C$,

(iv) $C \cap (A-p) \subseteq \{0\}$.

Proof. Observe first that $(\mathrm{int}\, K) \cap (\mathrm{int}\, K^*) \neq \emptyset$, where K^* is the nonnegative palar cone of K. Let b be a normed unit vector of that intersection, and let

$K(t) = \{ x \in K: \ x.b = t \}$, $t \geq 0$.

Consider the convex compact set $K(1)$. It follows from (5) that there is a positive ε , $0 < \varepsilon < 1/2$ such that

$\mathrm{clcone}(A,p) \cap (K(1)+B(0,\varepsilon)) = \emptyset$.

Let C be the cone generated by $K(1)+B(0,\varepsilon)$. It is obvious that C is the cone to be constructed.

Theorem 2. Let $p \in R^n$. Then (5) holfs if and only if there exist a continuous, monotonic function g on R^n and a closed, convex and pointed cone C containing $K\setminus\{0\}$ in its interior so that g separates (p+C) and A.

Proof. If (5) holds, then by Lemma 1 there exists a cone C satisfying (iii) and (iv). Take $e \in \mathrm{int}\, K \subseteq \mathrm{int}\, C$ and consider a function g given by the formula:

$g(x) = \inf \{t: \ x \in p+te+C \}$, $x \in R^n$.

It can be verified without any difficulties that g(.) is well defined, continuous and monotonic. (The monotonicity is derived from (iii) and the definition of g(.)). Moreover,

$g(p+c) \geq 0 \geq g(a)$ for each $c \in C$ and $a \in A$.

Thus, g separates p+C and A.
Conversely, if g separates the two sets, then

$\mathrm{clcone}(A,p) \cap (\mathrm{int}\, C) = \emptyset$.

Indeed, if that is not the case, then there is a vector $a \in A$ such that $a \in p+\mathrm{int}\, C$. Take a nonzero vector $b \in K$ with the norm small enough so that $a-b \in p+C$. Then by the monotonicity of g we have

$g(a) > g(a-b)$,

contradicting the fact that g separates p+C and A.
Further, since $K\setminus\{0\} \subseteq \mathrm{int}\, C$,

$(\mathrm{clcone}(A,p)) \cap K = \{0\}$,

and the proof is complete.

Lemma 2. Suppose that $p \in R^n$ and
$(\mathrm{clcone}(A\setminus B(p,t)) \cap K \subseteq \{0\}$ for all $t > 0$. (6)
Then there exists a closed set C with the following properties:

(v) $K \subseteq C$,

(vi) $C+(K\setminus\{0\}) \subseteq \mathrm{int}\, C$,

(vii) $C \cap (A-p) \subseteq \{0\}$.

Proof. As in the proof of Lemma 1, we consider $K(t)$, $t \geq 0$. By the condition of the lemma, for a fixed positive T, there is a positive ε such that

$(K(t)+B(0,\varepsilon)) \cap (A-p) = \emptyset$, each $t \geq T$

and for every t , $0 < t < T$ there is a positive $\varepsilon(t)$ such that

$(K(t)+B(0, \varepsilon(t))) \cap (A-p) = \emptyset$.

It is not difficult to construct a continuous function s(.) on R_+, the set of nonnegative numbers, so that

$s(t) \leq \min \{ t, \varepsilon(t), \varepsilon \}$, for each $t \geq 0$,

$s(t) < s(t')$ if $t < t'$.

Now, define

$$C = \bigcup_{t \geq 0} (K(t)+B(0, s(t)/2)).$$

The direct verification will show that C yields all the requirements of the lemma. The proof is complete.

Theorem 3. (6) holds if and only if there exists a continuous, monotonic function on R^n so that it separates (p+K) and A .

Proof. Assume that (6) holds. By Lemma 2, we get a closed set C satisfying (v),(vi) and (vii). Let $e \in \text{int } K$. Define a function g by the relation

$$g(x) = \inf\{t: x \in p+te+C\}, \quad x \in R^n.$$

Using a proof similar to that of Theorem 1, we can conclude that g is well defined, continuous and monotonic. Besides, it separates (p+K) and A.

Conversely, assume that g is a continuous, monotonic function separating (p+K) and A. Suppose to the contrary that (6) does not hold, i.e. there is a sequence $\{a_i\}$ from A with

$\|a_i-p\| \geq T$ for some positive T, and

$\lim d(a_i,p+K) = 0$,

where $d(.,.)$ stands for the distance from a_i to p+K.

The latter fact shows that there is a sequence $\{b_i\}$ from p+K such that

$$\lim \|a_i-b_i\| = 0. \tag{7}$$

Further, as g is monotonic, for a given T we have

$q = \inf\{g(x): x \in p+K, \|x-p\| \geq T/2\} > g(p)$.

We may assume $\|b_i-p\| \geq T/2$.

$$\text{Hence} \quad g(b_i) > g(p)+(q-g(p))/2. \tag{8}$$

For $\varepsilon = (q-g(p))/2$, it follows from (7) that

$a_i \in b_i - \varepsilon e+K$ if i is large enough.

Now, the latter relation gives us the relation:

$g(a_i) \geq g(b_i)- \varepsilon$.

Combining this with (8) we obtain

$g(a_i) \geq g(b_i)- \varepsilon > g(p)+ \varepsilon - \varepsilon = g(p)$.

In this way, g does not separate (p+K) and A. The proof is complete.

3. SCALARIZATIONS

Given a vector optimization problem (I). It can be seen that if problem (II) satisfies conditions (i) and (ii), then

$E(X',f') \subseteq E(X,f)$

(or $E(X',f') \subseteq WE(X,f)$ in the weak case).

The conditions (i) and (ii) are equivalent to the existence of a monotonic (weakly monotonic) function g on f(X) such that f' is the composition of f and g. In this section we shall study the existence of such functions by using the separation results developed in the previous section.

Proposition 1. $x \in WE(X,f)$ if and only if there exists a continuous, weakly monotonic function g on R^n such that $x \in E(X,g.f)$.

Proof. By definition, $x \in WE(X,f)$ if and only if

$(f(x)+\text{int } K) \cap f(X) = \emptyset$.

Now the proposition follows immediately from Theorem 1.

Proposition 2. $x \in E(X,f)$ and (6) holds for $A=f(X)$ and $p=f(x)$ if and only if there exists a continuous, monotonic function g on R^n such that $x \in E(X,g.f)$.

Proof. Note first that a monotonic function g separating $(f(x)+K)$ and $f(X)$ if and only if $x \in E(X,g.f)$. Now, apply Theorem 3 to get our proposition.

Definition 3. (see also Jahn(1984)) $x \in E(X,f)$ is said to be a proper optimal solution of (I) if

$(\text{clcone}(f(X),f(x)) \cap K = \{0\}$.

Denote the proper solution set of (I) by $PrE(X,f)$. Below we give some conditions for (6) to hold.

Proposition 3. (6) holds for $A=f(X)$ and $p=f(x)$ if one of the following conditions is satisfied:

(viii) $x \in PrE(X,f)$

(ix) $x \in E(X,f)$ and $f(X)$ is compact.

Proof. The first condition is obvious as (5) implies (6). For the second part, suppose to the contrary that (6) does not hold, i.e. there is a sequence $\{x_i\}$ from X with

$\lim (f(x_i)-f(x)) \in K \setminus \{0\}$.

As $f(X)$ is compact, we may assume $\lim f(x_i) = f(y)$ for some $y \in X$. Hence $f(y)-f(x) \in K \setminus \{0\}$, contradicting $x \in E(X,f)$. The proof is complete.

Proposition 4. $x \in PrE(X,f)$ if and only if there are a closed convex, pointed cone C containing $K \setminus \{0\}$ in its interior and a continuous function g which is monotonic with respect to C such that $x \in E(X,g.f)$.

Proof. This is immediate from Theorem 2.

CONCLUSION: Some results similar to that of Propositions 1 and 2 have been proved by Jahn (1984). However, the author uses seminorms in the role of g and therefore the function is merely weakly monotonic on some domain of the space (except for the convex case). The argument used in our paper allows to get some more interesting results. Namely, under appropriate assumptions about the alternative set and objective function, it is possible to construct a continuous, monotonic function g so that by solving the scalar problem associated with $(X,g.f)$ we can obtain all the optimal solutions of problem (I). This result and some others will be presented in a forthcoming paper.

REFERENCES

Jahn, J. (1984). Scalarization in vector optimization, Mathematical Programming 29, 203-218.

Wierzbicki,P.(1985).On the completeness and constructiveness of parametric characterizations to vector optimization problems, International Workshop on large-scale modelling and interactive decision analysis,Wartburg-GDR,reprint.

Multicriterial duality, examples and advances

A. Göpfert
Technical University Leuna-Merseburg
Dep. of Mathematics
4200 Merseburg, DDR
Germ. Democr. Repub.

1. Introduction

If we have constructed a mathematical model about a technical or economical topic, it is important for mathematicians as well as for decision makers, to analyse this mathematical model in order to get as much prior information as possible concerning the problem. Especially information about the existence of a solution, about estimations of the values of the solutions in the criteria space, and about how the solutions depend on perturbations, are desirable. This knowledge is the background for numerical procedures, particularly for interactive numerical procedures.

If the above mentioned mathematical modelling leads to an optimization problem, for instance to

$$f(x) \rightarrow \max,\ x \in \mathcal{L};\ f: \mathcal{L} \rightarrow Z = R^p,\ p \geq 1 \ \text{ integer}, \tag{P}$$

then in the case $p = 1$ (scalar optimization), the duality principle is often used to get information about stability in connection with perturbations and about estimations of the maximal value of (P) "from above". Words such as dual, polar, adjoint, complementary, conjugated and the co-terms, reflect different and useful applications of duality for scalar extremal problems (Göpfert 1982, 1986; Zeidler 1978).

If (P) is a real vector optimization problem, i. e. $p > 1$, it also seems possible to derive advantageous dual problems for the given primal multicriteria problem (P). I am convinced, that such duality theories must be developed, and I hope, that if we have a good duality theory, then we will be able to use it successfully for practical purposes. In picture 1 you see (in the case $Z = R^2$, ordered with the customary cone $K = R^2_+$) the set $f(\mathcal{L}) - K$ with the efficient set (in the Pareto sense) E_P. A useful dual program (D) to (P) would be

$$g(y) \rightarrow \min,\ y \in \mathcal{L}_D\ ;\ g : \mathcal{L}_D \rightarrow Z, \tag{D}$$

where no point of $g(\mathcal{L}_D)$ can be dominated by a point of $f(\mathcal{L})$, or more exactly, where

$$f(\mathcal{B}) - K \cap g(\mathcal{B}_D) + (K \setminus \{0\}) = \emptyset. \tag{1}$$

(1) is often called w e a k d u a l i t y. And (D) would be very good (or mathematically, s t r o n g d u a l i t y is valid), if there is "no gap" between E_P and the efficient set E_D of (D), look at picture 1. The both programs

$$(P_o)\ \begin{array}{l} Cx \to \min, \\ Ax \geqq b \\ x \geqq 0 \end{array} \qquad (D_o)\ \begin{array}{l} Ub \to \max \\ \exists v > 0 : v^T(UA-C) \leqq 0 \\ U \geqq 0 \end{array}$$

fulfil (1), but there is a gap (picture 2), as you see, choosing

$$C = \begin{pmatrix} -2 & 0 \\ 1 & -1 \end{pmatrix}, \quad A = (-1,-1), \quad b = -1.$$

If you would have constructed a "good" dual vector optimization problem, then (look at picture 1)
- firstly, if z_D is in $g(\mathcal{B}_D)$, y o u c a n ' t f i n d p o i n t s o f $f(\mathcal{B})$ or points of E_P in $g(\mathcal{B}_D) + (K \setminus \{0\})$,
- secondly, if you have some points z_D of $g(\mathcal{B}_D)$, you have in some sense an estimation of E_D f r o m a b o v e,
- finally, you can use points $z_D \in g(\mathcal{B}_D)$ for the reference point approach of Wierzbicki (1980). K. Lampe (1984) calculated, using the stochastic search procedure of Timmel (look up in K. Lampe 1984), the convex academic example

$$f(x) = \begin{pmatrix} -x_1 \\ x_1 + x_2^2 \end{pmatrix} \longrightarrow \min, \ x \in \mathcal{B},$$

$$\mathcal{B} = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : h_1(x) = x_1^2 - x_2 \leqq 0,\ h_2(x) = x_1 + 2x_2 - 3 \leqq 0 \right\},$$

and the Lagrange - dual problem

$$\begin{pmatrix} -x_1 + L_{11}(x_1^2 - x_2) + L_{12}(x_1 + 2x_2 - 3) \\ x_1 + x_2^2 + L_{21}(x_1^2 - x_2) + L_{22}(x_1 + 2x_2 - 3) \end{pmatrix} \xrightarrow[L,\ x \in \mathcal{B}_D]{} \max, \tag{2}$$

$$\mathcal{B}_D = \left\{ x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},\ L = \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix} : L_{ij} \geqq 0,\ \exists z_1^* > 0, \exists z_2^* > 0 \right.$$

$$\left.\begin{aligned} z_1^*(-1+2L_{11}x_1+L_{12}) + z_2^*(1+2L_{21}x_1+L_{22}) &= 0 \\ z_1^*(-L_{11} + 2L_{12}) + z_2^*(2x_2 - L_{21} - 2L_{22}) &= 0 \end{aligned}\right\}$$

with

(where the equations in $\mathcal{L}_D$ come from $z^{*T}(f_x + Lh_x) = 0^T$) a p p r o x i m a t e l y. She got the dotted curves in picture 3 and (because (2) is a "gap-free" dual problem) she considered them as e s t i m a t i o n s of E_P from above respectively of E_D from below.

Now I state some results concerning duality theory, which we got in Merseburg (Gerstewitz/Iwanow 1985). Some of them were sharpened by Nehse and Iwanow in Ilmenau (look e.g. in Gerstewitz/Iwanow 1985).

2. Duality for convex problems

Now about duality theory for convex problems in $Z = R^p$, ordered with a convex cone K, not necessary closed, but of course $0 \in K$ (and $K \cap (-K) = \{0\}$, $Z = K-K$). We used the duality results (of the convex scalar optimization) with perturbations. Imbedding the given primal convex problem

$$f(x) \rightarrow \min, \quad x \in \mathcal{L}, \qquad (P_1)$$

in a perturbed problem (perturbation u) with F(x,u) convex, F(x,0) = f(x), then scalarizing F with a vector $z^* \in R^p$, $z_i > 0$, $i=1,\dots,p$, z^*F convex, proper, lower semicontinuous, then conjugating $g=z^*F$ to $g^*_{z^*}(x^*,u^*)$, we can write the dual program

$$h \rightarrow \max, \quad h \in \mathcal{L}_{D_1}, \qquad (D_1)$$

$$\mathcal{L}_{D_1} = \{ h \in R^p : \exists\, z^* > 0,\ \exists\, u^* \in R^m \text{ with } z^{*T} h \leqq -g^*_{z^*}(0,-u^*) \}.$$

In the case p=1 the dual problem would be the customary one

$$-g^*_{z^*}(0,-u^*) \rightarrow \max.$$

Then (1) is fulfilled and s t r o n g d u a l i t y t h e o r e m s are valid (also in general spaces) such as:

Strong direct duality theorem: If $f(x_o)$ is properly efficient relativ to (P_1), that is, x_o solves the scalar problem

$$z^* f(x) \rightarrow \min, \quad x \in \mathcal{L}, \qquad (3)$$

with a certain $z^* \in R^p$, $z_i^* > 0$ $\forall$ i=1,...,p, then $f(x_o) \in E_{D_1}$, if the problem $z^{*T} F(x,0)$ is stable (in the sense of Rockafellar).

Inverse duality theorem: If $-g^*_{z^*}(0,u^*)$ is stable $\forall z^*$ mit $z_i^* > 0$, and if h^o is efficient relative to (D_1) with $(z^*)^o > 0$, then there exist $h^1 \in \mathcal{L}_{D_1}$, efficient in (D_1) and $\overset{1}{x} \in \mathcal{L}$, properly efficient in (P_1) with

$$(z^*)^o h^o = (z^*)^o h^1 \ , \ h^1 = f(\overset{1}{x}) \ .$$

Iwanow sharpened especially in the last theorem "stable" to "normal" in the sense of Ekeland-Temam. And now the following corollar to the last theorem is interesting (U. Lampe 1981): If $f(\mathcal{L})$ is compact, then to each efficient h^o in (D_1) with $(z^*)^o$ there is an efficient h^2 of (D_1) with a properly efficient $\overset{2}{x}$ of (P_1), such that

$$f(\overset{2}{x}) = h^2, \ (z^*)^o h^o = (z^*)^o h^2, \ h^o - h^2 \in \bar{K} \setminus (K \setminus \{0\}) \ .$$

The last inclusion reflects, that K is not necessary closed.

The dual problem D_L of Isermann for the linear problem (P_L)

$$Cx \to \min, \ Ax = b, \ x \geqq 0; \ C:R^n \to R^p, \ A:R^n \to R^m, \qquad (P_L)$$

is

$$Ub \to \max, \ \exists v > 0 : v^T(UA-C) \leqq 0 \qquad (D_L)$$

and is a special case of (D_1). For other examples look in Jahn (1986).

3. Nonconvex duality

Now some remarks concerning nonconvex duality. The above made demand "no gap between E_P and E_D" on strong duality theorems is, as the known duality theorems (see f. e. Gerstewitz/ Iwanow 1985, Jahn 1986, Nakayama 1984) show, somewhat sharp. In the duality of (P_1) and (D_1) only the "properly" efficient solutions of (P_1) work! An other definition of proper efficiency is the following (U. Lampe 1981): Under the same suppositions as in 1. we say, that $f(x) \in f(\mathcal{L})$ is p r o p e r l y e f f i c i e n t in the sense of Lampe, if there is an open convex cone $C \neq Z$, $C \supset K \setminus \{0\}$, and $f(x)$ is efficient even relative to $C \cup \{0\}$ (instead of K).

Denoting $x \in \mathcal{L}$ with f(x) properly efficient in this sense with M_e and such $x \in \mathcal{L}$, with f(x) properly efficient in the sense of Geoffrion (see (3)) with N, it is valid $N \subseteq M_e$. If P is concavelike, that is (for maximum-problems), if $f(\mathcal{X})-K$ is convex, then $N=M_e$. Especially $N=M_e$ for concave problems (P) (or convex (P_1)). The proof of $N=M_e$ is founded on a separation theorem: Two convex sets (one of them is C) are separated by a certain hyperplane, which finally determines the z^* in (3). With such z^* and such separation theorems we established the convex multicriterial duality as in 2. And now, Gerstewitz had the idea, to generalize the c o n e C to a certain convex s e t C and, having this in mind, to separate a nonconvex set (being in connection with efficient points of (P)) from C by a c o n c a v e functional s instead of a l i n e a r functional z^* (as in the convex optimization). So we can state (also in general spaces) with K as in 2. (Gerstewitz/Iwanow 1985):(K closed, C open)

Proper efficiency with a convex set $C(x \in M_{e'})$: If $C \subset Z$ is a convex set with $C \supset K\setminus\{0\}$, $C \neq Z$, $\bar{C} + (K\setminus\{0\}) \subseteq C$, then $x \in \mathcal{X}$ may belong to $M_{e'}$, if $f(x)$ is efficient relative to $C \cup \{0\}$ (instead of K).

Theorem: $x_o \in M_{e'}$ iff there exists a concave continuous strictly monotoneous functional[1] $s : Z \to R$ with

$$s(z_o) \geqq s(z) \;\; \forall z \in f(\mathcal{X}) \tag{4}$$

where $z_o = f(x_o)$.

(4) is nothing else than generalizing the (linear) z^* in the proper efficiency of Geoffrion to (the nonlinear, i.e. concave) s. And if

$$N' = \{x_o \in \mathcal{X} : \exists \text{ s as in (4) with: } x_o \text{ solves } s\, f(x) \to \max, x \in \mathcal{X}\},$$

then is valid $N' = M_{e'}$ (in pict. 4 we see an example for $x_o \in M_{e'} = N'$). The proof is done using a new separation theorem: If $A \subset Z$, $A \neq \emptyset$, $A \cap C = \emptyset$, where C is (e.g.) our convex set C, then $\exists$ s (as in (4)) with

$$s(A) \leqq 0, \; s(C) > 0.$$

With s and using ideas of Klötzler-Krotov (see Göpfert/Gerstewitz 1986), we can state a very general dual problem (D_g) to (P) and even can give strong duality theorems. With this duality theorems we thus obtained the same generality as in the scalar optimization. The reason for getting such satisfactory general problems is, that for the Lagrange functionals a l l proper $L : R^n \to \bar{R}$ are admissible, whereas in the customary (convex) case L is linear. And as in the scalar case it is desirable for special problems, to diminish the class of possible functionals L and s. For vector optimization problems, using an idea of Ester (e.g. in(Gerstewitz/Iwanow 1985)), Gerstewitz discovered, that the solution $\mathfrak{z}(t,x) : (R^1_+ \setminus\{0\}) \times R^k \to R$ of

---

1) $s(z_1) > s(z_2)$ for $z_1 - z_2 \in K\setminus\{0\}$

$$\prod_{k=1}^{p} (x_i - \mathfrak{z}(t,x)) = \frac{1}{t}, \quad x_i - \mathfrak{z}(t,x) > 0 \;\; \forall\, t > 0,\; x \in R^k,$$

is concave (continuous and strongly monotoneous), and stated together with Iwanow a dual problem and a strong duality theorem for finite dimensional problems (P) with s of the type as $\mathfrak{z}$, so that at least s is a member of a smaller class of functionals.

Finally we represent the forementioned dual problem (D_g) to the primal problem (P_g) and an appertaining strong duality theorem (which is also valid in more general spaces); K, K_y as in 2.

$$f(x) \to \max,\; x \in \mathcal{L} = \{v \in X = R^n : g(v) \underset{K_y}{\geqq} 0,\; K_y \subseteqq Y = R^m\}, \qquad (P_g)$$

$$h \to \min,\; h \in \mathcal{L}_{D_g} = \{k \in Z = R^p : \exists\, s \text{ as in } (4),\; \exists\, L : Y \to \bar{R},$$

$$L(u) \geqq 0 \;\; \forall\, u \underset{K_y}{\geqq} 0,\; s(k) \geqq \sup_{x \in X} (s(f(x)) + L(g(x)))\}. \qquad (D_g)$$

<u>Theorem:</u> $f(M_{e'}) \subseteqq E_{D_g}$, i.e., properly efficient points from (P_g) are efficient relative to D_g.

With additional assumptions there is also a converse duality theorem.

4. References

Gerstewitz, Chr. u. E. Iwanow (1985). Dualität für nichtkonvexe Vektoroptimierungsprobleme, Wiss. Zeitschr. TH Ilmenau 31(1985)2, 61-81.

Göpfert, A. (1982). Über einige Dualitätsbegriffe in der Mathematik, Wiss. Zeitschr. TH Ilmenau 28(1982)6, 53-68.

Göpfert, A., Bittner, L., Elster, K.-H., Nozicka, F., Piehler, J. u. R. Tichatschke (Ed.)(1986). Lexikon der Optimierung, Akademie-Verlag Berlin, appears in 1986/87.

Göpfert, A. u. Chr. Gerstewitz (1986). Über Skalarisierung und Dualisierung von Vektoroptimierungsproblemen, Zeitschr. f. Analysis u. Anwendgn., appears in 1986.

Jahn, J. (1986). Mathematical Vector Optimization in Partially Ordered Linear Spaces. Bd. 31 von: Methoden und Verfahren der mathematischen Physik, Ed. B.Brosowski u. E. Martensen, Verlag P. Lang, Frankfurt a.M., Bern, New-York 1986.

Lampe, K. (1985). Numerischer Vergleich des dualen und primalen Vektoroptimierungsproblems mit dem stochastischen Suchverfahren von Timmel, Wiss. Zeitschr. TH Ilmenau 31(1985)2, 93-100.

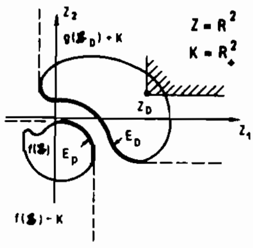
Lampe, U. (1981). Dualität und eigentliche Effizienz in der Vektoroptimierung, Sem.berichte Nr. 37 Humboldt-Universität Berlin 1981, 45-54.

Nakayama, H. (1984). Geometric consideration of duality in vektor optimization, J. Optim. Theory and Appl., 44(1984)4, 625-655.

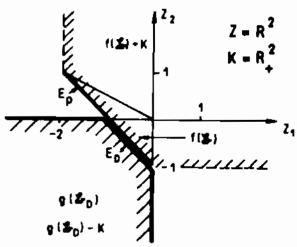
Wierzbicki, A. (1980). A mathematical basis for satisfieing decision making. Working Paper 80-90 IIASA (1980).

Zeidler, E. (1977). Vorlesungen über nichtlineare Funktionalanalysis III - Variationsmethoden und Optimierung - . Teubner-Texte zur Mathematik. BSB B.G. Teubner Verlagsgesellschaft Leipzig 1977.

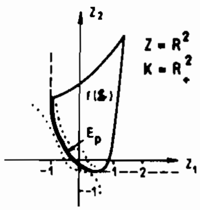
5. Pictures



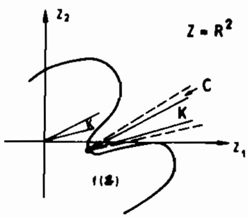
pict.1: weak duality



pict.2: gap



pict.3: enclosing of E_p



pict.4: nonconvexity

CONCEPTS OF EFFICIENCY AND FUZZY AGGREGATION RULES

Jochen Ester, Technische Hochschule Karl-Marx-Stadt
9010 Karl-Marx-Stadt, PSF 964, GDR

1. INTRODUCTION

Some concepts of efficiency can be covered by an unified approach with a substitute parametric optimization problem. This was already shown by Gearhart (1983) for properly efficient solutions in the sense of Henig (1982). In the paper presented here other definitions of solutions are included based on the classical Pareto concept (e.g. set of extremals, weakly efficient set).
Using the Theory of Fuzzy Sets we can give an interpretation of a special substitute scalarizing function as a generalized aggregation rule of fuzzy sets. This rule depends on parameters, which control the (logical) properties of the aggregation. Therefore, the aggregation rule can be adapted to the logical structure of the concept for efficiency. This concept must be given in the form of a hierarchical structure with additionally given weighting coefficients on each level.

2. THE MCDM-PROBLEM AND SEVERAL SOLUTIONS

Let us assume that we have a vectormaximum problem such as

$$\max \left\{ q(x) \;\middle|\; x \in X \subseteq R^n,\ q \in Z \subset R^m \right\} \tag{1}$$

with the denotations

$$q = (q_1 \ \dots \ q_m)^T \ , \quad x = (x_1 \ \dots \ x_n)^T \ .$$

A basic solution for (1) is the Pareto set, which is defined by

$$PM = \left\{ q^o \;\middle|\; \overline{\exists}\, q: \quad q \succ q^o,\ q, q^o \in Z \right\} , \tag{2}$$

where the binary preference relation "$\succ$" means (Elster, Nehse 1979)

$$q^1 \succ q^2 \Leftrightarrow q^1 \geqq q^2 \Leftrightarrow q^1 - q^2 \in R^m_+ \setminus \{0\} . \tag{3}$$

Z is a mapping of X and may be an arbitrary set. It can even be a discrete set. Now, we have the following demands

$$\max \left\{ q_i \mid q \in Z \right\} = q_i^{max} \leq M < +\infty \quad , \; i=1(1)m \quad , \tag{4}$$

$$\min \left\{ q_i \mid q \in PM \right\} = q_i^{min} \geq N > -\infty \quad , \; i=1(1)m \quad . \tag{5}$$

All further considerations are related to the space of the objectives and the set Z.
With

$$q_i^+ > q_i^{max} \; , \; q_i^- < q_i^{min} \; , \; i=1(1)m \tag{6}$$

and

$$q_i := \frac{q_i - q_i^-}{q_i^+ - q_i^-} \quad , \; i=1(1)m \tag{7}$$

we guarantee that the Pareto set of (1) lies in the unit cube after the transformation (7).

$$PM = \operatorname{Argmax} \left\{ q \mid q \in Z \right\} \subset \left\{ q \in R^m_+ \mid 0 \leq q_i \leq 1, \; i=1(1)m \right\} \tag{8}$$

Besides the Pareto set (or efficient set) we know a lot of other concepts for efficiency, for instance: weakly efficient set (Elster, Nehse 1979), properly efficient set in the sense of Schönfeld (Schönfeld 1970), in the sense of Salukvadse (Salukvadse 1975), in the sense of Nash, the set of extremals (Makarov et.al. 1982) or set of partial solutions etc.
The use of this great variety of concepts for efficiency could not be justified from a practical point of view. We can find a practical justification for the Pareto set only. But, unfortunately, the Pareto set has mostly a very high number of elements. Therefore, the Pareto set or one of its approximations can only be a solution of a MCDM-Problem in a first step of a decision making process. With the Theory of Fuzzy Sets we can give now a practical interpretation for some other solution sets.
This is possible due to the definitions of these sets by corresponding substitute parametric optimization problems.

2.1. Weakly efficient set

Using the following preference relation

$$q^1 \succ q^2 \Leftrightarrow q^1 > q^2 \Leftrightarrow q^1 - q^2 \in \operatorname{int}(R^m_+) \quad , \tag{9}$$

where $\operatorname{int}(R^m_+)$ means the interior of the positive orthant in the Euclidean space, we get the weakly efficient set of solutions PMS.

$$PMS = \left\{ q^s \,\middle|\, \overline{\exists} q: q \succ q^s, \quad q, q^s \in Z \right\} \tag{10}$$

It holds $PM \subseteq PMS$. The same set may be obtained by solving the parametric problem

$$PMS = \operatorname{Argmax} \left\{ \min_i (g_i q_i) \,\middle|\, q \in Z,\ g = (g_1 \ldots g_m)^T \in R_+^m \setminus \{0\} \right\}, \tag{11}$$

where g is a vector of parameters. This problem is known as the Germeier problem (Germeier 1971). Therefore, it is possible to name PMS the (properly) efficient set in the sense of Germeier.

2.2. Hyperbola efficient set

With another preference relation

$$q^1 \succ q^2 \Longleftrightarrow \left[\prod_{i=1}^{m} (q_i^1 - q_i^2)^{g_i} > 0 \right] \wedge \left[q_i^1 - q_i^2 > 0,\ i = 1(1)m \right] \tag{12}$$

we can determine the hyperbola efficient set PMN. We obtain the same solution for

$$\max \left\{ \prod_{i=1}^{m} (q_i)^{g_i} \,\middle|\, q \in Z,\ g = (g_1 \ldots g_m)^T \in R_+^m \setminus \{0\} \right\} . \tag{13}$$

This problem, well-known from Game Theory for cooperative games, was formulated by Nash. Therefore, it is possible to name PMN the (properly) efficient set in the sense of Nash. The following parametric problem belongs to the same class

$$\max \left\{ e \,\middle|\, \prod_{i=1}^{m} (g_i q_i - e) = a,\ q \in Z,\ g \in R_+^m \setminus \{0\},\ a \geq 0,\ g_i q_i - e \geq 0 \right\} \tag{14}$$

with the solution set PMH. In (Ester, Schwartz 1983) was shown

$$PMH(a > 0) \subseteq PM, \quad PMH(a = 0) = PMS,$$

$$\lim_{a \to +0} (PMH) = PM \qquad . \tag{15}$$

The separating hyperplanes in (13) and (14) belong to the same class (for m=2 we get hyperbolas).

2.3. Properly efficient set in the sense of Schönfeld

In this case the definition is given by the following substitute parametric optimization problem

$$\max \left\{ e = \sum_{i=1}^{m} g_i q_i \,\middle|\, q \in Z,\ g \in R_+^m \setminus \{0\} \right\}. \tag{16}$$

The solution set of (16) is called properly efficient set in the sense of Schönfeld

$$PMP = \mathrm{Argmax}(e) \; . \tag{17}$$

It holds PMP $\subseteq$ PMS. In the case of strongly convex Z we know

$$PM = PMS = PMP \; . \tag{18}$$

These results are well-known from the literature (Elster, Nehse 1979).

2.4. Set of extremals (Makarov et.al. 1982)

A generally rarely applied concept of solution is that of extremals or of partial solutions, defined by

$$PME = \bigcup_i \mathrm{Argmax}(q_i \mid q \in Z) \; , \tag{19}$$

which can be determined by (4). The elements of PME are needed for the computation of the so-called utopia point. It holds PME $\subseteq$ PMS. We can obtain the set PME by solving the problem

$$\max \left\{ \max_i (g_i q_i) \;\middle|\; q \in Z, \; g \in R^m_+ \setminus \{0\} \right\} \; . \tag{20}$$

2.5. Other concepts

A lot of different definitions for solutions of MCDM-problems has the same outcomes as described by (11),(13), (16) and (19). For instance, we can use distance- or achievement-functions with utopia points, nadir points, other reference points etc., but the solutions are the same as above. This can be explained by the fact that the separating hyperplanes in different substitute parametric optimization problems are similar.
Therefore, we do not discuss other definitions. In special cases it is quite easy to prove the shape of the corresponding hyperplanes.

3. GENERALIZED PARAMETRIC CHARACTERIZATION

We introduce the auxiliary objective

$$e(q,p,g) = \left[\sum_{i=1}^{m} g_i (q_i)^p \right]^{1/p} \tag{21}$$

with

$$g \in R^m_+ \setminus \{0\}, \; q \in Z \; , \; p \text{ real} \tag{22}$$

and obtain that the weakly efficient set, the hyperbola efficient set, the properly efficient set in the sense of Schönfeld and the set of extremals are special solutions of

$$\max\left\{e(q,p,g)\ \middle|\ q\in Z\right\}, \tag{23}$$

where $p = (-\infty)$, 0, 1, $(+\infty)$. In contrast to earlier investigations we have extensions in p from integer to real, from positive values to positive-negative values including zero, plus and minus infinite.
The character of the function (21) changes from convex to concave at the point (p=1), thus causing an essential deviation from the concept presented by Gearhart (1983). Note that (21) can not be used in the sense of a norm.
From the literature we know
<u>THEOREM 1:</u> The solution set E^p of

$$\max\left\{e(q,p,g)\ \middle|\ q\in Z,\ g\in \operatorname{int}(R^m_+),\ p \text{ real}\right\}$$

is a subset of the Pareto set

$$E^p = \bigcup_g E^p(g) = \operatorname{Argmax}(e) \subseteq PM\ . \tag{24}$$

In (Ester,Tröltzsch 1986) is shown the following
<u>THEOREM 2:</u> It holds

$$E^p \subseteq E^s \quad \text{for all} \quad -\infty \leq s \leq p \leq +\infty \tag{25}$$

$$\text{for } g\in R^m_+\setminus\{0\}\ .$$

This means, there exists a chain of inclusions by varying the parameter p. We get the highest number of solution elements for $p=(-\infty)$. If p is increasing now, some solution elements must be cancelled and we get the partial solution for $p=(+\infty)$.
It must be noted that the inclusion (25) is not strong.
In special cases the solution set is not strongly decreasing with increasing power p.
Without proof we give
<u>THEOREM 3:</u> Suppose the point-to-set mapping $(p\rightarrow E^p)$ is continuous for all $p<p'$. Then we find $p_o<p'$ with the following condition for a given arbitrary ε.

$$\forall\ q\in PM:\ \exists q'\in E^p:\ \ |q-q'|<\varepsilon\ ,\ p<p_o. \tag{26}$$

This means from a practical point of view that we can compute the Pareto set by using the auxiliary objective (21) if the power p lies under a certain limit.
Using this approximation theorem the Pareto set does fit in the chain of inclusions (25).
Without limitation of generality we can suppose

$$\sum_{i=1}^{m} g_i = 1\ . \tag{27}$$

Besides, that generalized parametric problem can be adapted to the conception of preferences of a decision maker. It has all properties of a preference function.

4. INTERPRETATION OF THE AUXILIARY PROBLEM

Using the Theory of Fuzzy Sets (Zadeh 1965), we shall investigate the connections between the solutions of (23) and the structure of the concepts for efficiency of a decision. Therefore, we need some membership functions of fuzzy sets.

- o The fuzzy set of all decisions, which are the best decisions with respect to the criterion q_i is

$$A_i = \left\{ q \mid z_i(q) \right\} \quad \forall\, q \in R_+^m \ , \ i=1(1)m \ . \qquad (28)$$

- o The fuzzy set of all decisions, which are the best decisions with respect to all criteria is an aggregation of A_i

$$A = \mathrm{aggr}(A_1,\dots,A_m) = \left\{ q \mid z(q) \right\} \quad \forall\, q \in R_+^m \ . \qquad (29)$$

z and z_i are the membership functions of the corresponding fuzzy sets, satisfying

$$0 \le z,\ z_i \le 1 \ .$$

Now we can choose

$$z_i(q) = q_i \ , \ i=1(1)m \qquad (30)$$

and for z we take the following aggregation rule

$$z = \left[\sum_{i=1}^{m} g_i (q_i)^p \right]^{1/p} = \left[\sum_{i=1}^{m} g_i (z_i)^p \right]^{1/p} = e \ . \qquad (31)$$

We find the best element (the best decision) as follows

$$\max \left\{ z \mid q=(z_1 \dots z_m)^T \in Z \cap R_+^m,\ g \in R_+^m \setminus \{0\},\ p \text{ real} \right\} \ . \qquad (32)$$

From fuzzy theory we know (Dubois, Prade 1980) a lot of aggregation rules.
For instance:

- o intersection of fuzzy sets ($A = A_1 \cap A_2 \cap \dots$)

$$z(A) = \min_i (z_i) \quad \text{or}$$

$$z(A) = \prod_{i=1}^{m} (z_i) \quad ,$$

- o union of fuzzy sets ($A = A_1 \cup A_2 \cup \dots$)

$$z(A) = \max_i(z_i) \qquad \text{or}$$

$$z(A) = \sum_{i=1}^{m}(z_i) \qquad .$$

These aggregation rules are special cases of (32) or (23). For $p=(-\infty)$ we obtain the "pure" intersection, and for $p=(+\infty)$ we arrive at "pure" union. The transition from intersection to union lies between $0 < p < 1$. The opportunity to vary p enables us to control the (logical) properties of the aggregation in a soft manner. However, we recognize a very surprising fact. The set of the best disjunctive decisions is a subset of the best conjunctive decisions being the weakly efficient set. At first sight we would expect the converse relation!

5. HIERARCHICAL STRUCTURE OF EFFICIENCY

In many cases we have such a situation that we can find disjunctive and conjunctive demands in the aggregation (29). For instance, the decision maker can state:
"The decision must be done with respect to the criteria q_1 and q_2 and ... and q_7 or q_1 and q_3 and ... and q_8 or ... etc."
This means

$$A = \operatorname{aggr}(A_1,\ldots,A_m) = \bigcup_{j\in J}\ \bigcap_{i\in I_j} A_i \quad , \qquad (33)$$

$$A'_j = \operatorname{aggr}_j(A_i,\ \forall i\in I_j)\ ,\quad \forall j\in J \quad . \qquad (34)$$

On the first level in the hierarchy we find the (basic) fuzzy sets A_i. With (34) we can compute the fuzzy sets on the second level A'_j. The fuzzy set A is the efficient set on the top level

$$A = \operatorname{aggr}(A_1,\ldots,A_m) = \operatorname{aggr}_o(A'_j,\ \forall j\in J)\ . \qquad (35)$$

The corresponding membership functions are

$$z_{A_i} = z_i = q_i\ ,\ i=1(1)m \quad , \qquad (36)$$

$$z_{A'_j} = z'_j = \left[\sum_{i\in I_j} g_i^j (z_i)^{p^j}\right]^{1/p^j} ,\ \forall j\in J \quad , \qquad (37)$$

$$z_A = z = \left[\sum_{j\in J} g_j^o (z'_j)^{p^o}\right]^{1/p^o} \quad . \qquad (38)$$

Each "subsystem" has relevante criteria, special "weighting factors" and its characteristic aggregation parameter p.

This idea renders possible to build up a model of efficiency for each problem. Therefore, it can be used in decision support systems and in automatic decision making. Obviously, the main problem is the determination of the parameters g and p.

REFERENCES

Dubois,D.;Prade,H. (1980). Fuzzy sets and systems - theory and applications. Academic Press, New York.

Elster,K.-H.;Nehse,R. (1979). Ergebnisse und Probleme der Vektoroptimierung. 24.IWK,TH Ilmenau, Reihe B1.

Ester,J.;Schwartz,B. (1983). Ein verallgemeinertes Effizienztheorem. MOS, ser.Optimization, 14(3).

Ester,J.;Tröltzsch,F. (1986). On generalized notions of efficiency in MCDM. Syst.Anal.Model.Simul., 3(2).

Gearhart,W.B. (1983). Characterization of properly efficient solutions by generalized scalarization methods. JOTA, 41: 491-502.

Germeier,J.B. (1971). Einführung in die Theorie der Operationsforschung (russ.). Nauka, Moskau.

Henig,M.I. (1982). Proper efficiency with respect to cones. JOTA, 36: 387-407.

Makarov,I.M.et.al. (1982). Auswahltheorie und Entscheidungsfindung (russ.). Nauka, Moskau.

Salukvadse,M.S. (1975). Aufgaben der Vektoroptimierung in der Theorie der Steuerung (russ.). Meznieraba, Tbilissi.

Schönfeld,P. (1970). Some duality theorems for the non-linear vectormaximum problem. Unternehmensforschung,14(1).

Zadeh,L.A. (1965). Fuzzy sets. Information and Control, 8: 338-353.

MULTIREGIONAL ECONOMIC POLICY-MODELLING

Marc Despontin
Centre for Statistics and Operations Research, Vrije Universiteit Brussel
Pleinlaan 2, B-1050 Brussels, Belgium

1. INTRODUCTION

In our experiments of policy-modelling with a multiregional econometric model for Belgium (Despontin 1981, 1982), both optimal control theory, using a classical symmetric quadratic penalty function, as multiple criteria linear programming methods were used to solve multiregional quantitative economic policy problems.

In this paper, a sketch of a possible approach is given trying to combine both methodologies for multiregional models in an interactive way.

2. MULTIREGIONAL DYNAMIC MODELS

Let us consider a multiregional dynamic linear(ized) model with r regions in reduced form:

$$Y_t = R_1 Y_{t-1} + R_2 Z_t + R_3 W_t$$

in which:

$$Y_t = \begin{pmatrix} Y_{1t} \\ Y_{2t} \\ \vdots \\ Y_{rt} \\ Y_{nt} \end{pmatrix} \quad ; \quad Z_t = \begin{pmatrix} Z_{1t} \\ Z_{2t} \\ \vdots \\ Z_{rt} \\ Z_{nt} \end{pmatrix}$$

where: Y_{jt} : vector of targets of economic policy for region j ($j=1,2,\dots,r$);

Y_{nt} : vector of supraregional (national) targets of economic policy;

Z_{jt} : vector of instruments of economic policy for region j ($j=1,2,\dots,r$);

Z_{nt} : vector of supraregional (national) instruments of economic policy;

W_t : vector of data of economic policy.

Considering a time horizont T, the general multiregional dynamic economic policy problem can be written as:

$$
\begin{cases}
\text{Max } f_{1it}(Y_{1t}, Y_{nt}, Z_{1t}, Z_{nt}) & i=1,2,\dots,c_1;\ t=1,2,\dots,T \\
\text{Max } f_{2it}(Y_{2t}, Y_{nt}, Z_{2t}, Z_{nt}) & i=1,2,\dots,c_2;\ t=1,2,\dots,T \\
\quad\vdots & \quad\vdots \\
\text{Max } f_{rit}(Y_{rt}, Y_{nt}, Z_{rt}, Z_{nt}) & i=1,2,\dots,c_r;\ t=1,2,\dots,T \\
\text{Max } f_{nit}(Y_{nt}, Z_{nt}) & i=1,2,\dots,c_n;\ t=1,2,\dots,T \\
Y_t = R_1 Y_{t-1} + R_2 Z_t + R_3 W_t & t=1,2,\dots,T \\
Y_o = Y_o^{\star} & \\
DY + EZ \leqslant B &
\end{cases}
\tag{1}
$$

in which: f_{jit} : ith criterion for region j for time period t;

c_j : number of criteria of region j;

$Y = (Y_1\ Y_2\ \dots\ Y_T)'$

$Z = (Z_1\ Z_2\ \dots\ Z_T)'$

D, E and B: matrices of appropriate order of coefficients of linear restrictions on the targets and instruments;

$Y_o^{\star}$: vector of current values of the target variables;

and in which, for the sake of simplicity, but without real restriction, we suppose that each region is only interested in the values of the instruments and targets of its own and of the supraregional level.

In a classical unicriterion optimizing approach, (1) is reduced to:

$$
\begin{cases}
\text{Max } F((f_{jit},\ j=1,2,\dots,r;\ i=1,2,\dots,c_j),(f_{nit},\ i=1,2,\dots,c_n),t=1,2,\dots,T) \\
Y_t = R_1 Y_{t-1} + R_2 Z_t + R_3 W_t \qquad t=1,2,\dots,T \\
Y_o = Y_o^{\star} \\
DY + EZ \leqslant B
\end{cases}
\tag{2}
$$

in which F is a super-criterion.

As clearly exposed by U. Reimers (1984, 1985), the traditional price-directive or resource-directive decomposition principles (e.g. G.B. Dantzig and P. Wolfe, 1961, J. Kornai and T. Liptak, 1965) are based on very restrictive hypotheses of which not only the economic relevance can be questioned in practice, but even are not applicable in a multiregional multiple objective approach.

In a classical unicriterion satisficing approach, by extension of H. Theil (1958), the problem is formulated as:

$$\begin{cases} \text{Min} \sum_{t=1}^{T} [(Y_t-Y_t^*)'G(Y_t-Y_t^*) + (Z_t-Z_t^*)'H(Z_t-Z_t^*)] \\ Y_t = R_1Y_{t-1} + R_2Z_t + R_3W_t \qquad t=1,2,\ldots,T \\ Y_o = Y_o^* \end{cases} \tag{3}$$

where F is a penalty function to be minimized and in which:

Y_t^* : desired values for the targets;

Z_t^* : desired values for the instruments;

G, H : matrices of penalty weights.

The satisficing approach appears to be very efficient for larger models, while a major difficulty consists in defining the penalty weights and desired values. From our experiences it became clear however that, in this simple quadratic model, often the penalty weights have to be chosen as to obtain realistic policies. As a consequence, they are not always longer a reflection of real preference structures, as they should be.

In the multicriteria optimizing approach (1), using the perturbation method explained in M. Despontin (1984), preference modelling is very flexible as long as the number of time periods, targets and instruments is not too large to allow an iterative interaction with the decision-maker(s).

Both approaches are not competing, but are complementary in a multicriteria framework.

In next model, we try to combine both in a multiregional setting.

3. AN INTERACTIVE FRAMEWORK

The interactive optimization proceeds in following steps:

(S1) The supraregional (national) objectives (Y_{nt}, Z_{nt}; t=1,2,...,T) and desired values (Y_{nt}^*, Z_{nt}^* ; t=1,2,...,T) are determined by the supraregional decision-maker (or decision-makers acting as one).

(S2) Each region j determines autonomously its own objectives (Y_{jt}, Z_{jt} ; t=1,2,...,T) and desired values (Y_{jt}^*, Z_{jt}^* ; t=1,2,...,T)

(S3) Given the desired values of the supraregional and regional targets and instruments, the vectors Y_t^* and Z_t^* (t=1,2,...,T) can be composed:

$$Y_t^* = \begin{pmatrix} Y_{1t}^* \\ Y_{2t}^* \\ \cdot \\ \cdot \\ \cdot \\ Y_{rt}^* \\ Y_{nt}^* \end{pmatrix} \quad ; \quad Z_t^* = \begin{pmatrix} Z_{1t}^* \\ Z_{2t}^* \\ \cdot \\ \cdot \\ \cdot \\ Z_{rt}^* \\ Z_{nt}^* \end{pmatrix}$$

(S4) Every region j determines its own matrices G_j and H_j of penalty weights for deviations on all targets and instruments.

(S5) For each region j following problem is solved:

$$\begin{cases} \text{Min} \sum_{t=1}^{T} \left[(Y_t - Y_t^*)' G_j (Y_t - Y_t^*) + (Z_t - Z_t^*)' H_j (Z_t - Z_t^*) \right] \\ Y_t = R_1 Y_{t-1} + R_2 Z_t + R_3 W_t \qquad t=1,2,\ldots,T \\ Y_o = Y_o^* \end{cases}$$

which gives an optimal control solution per region j we will denote by:

$$X_j^\circ = \begin{pmatrix} Y_1^\circ \\ Z_1^\circ \\ Y_2^\circ \\ Z_2^\circ \\ \cdot \\ \cdot \\ \cdot \\ Y_T^\circ \\ Z_T^\circ \end{pmatrix} \qquad j=1,2,\ldots,r$$

(S6) Let $Y_j(X_i^\circ)$ and $Z_j(X_i^\circ)$ respectively be the vectors of targets and instruments of the jth region and the supraregional level (over the whole time horizon), obtained by using the ith region optimal control solution X_i°.

Following (multidimensional) pay-off table can then be obtained:

objectives of region j / optimal control sol. of region j	1	2	...	r
1	$Y_1(X_1^\circ)$, $Z_1(X_1^\circ)$	$Y_2(X_1^\circ)$, $Z_2(X_1^\circ)$	...	$Y_r(X_1^\circ)$, $Z_r(X_1^\circ)$
2	$Y_1(X_2^\circ)$, $Z_1(X_2^\circ)$	$Y_2(X_2^\circ)$, $Z_2(X_2^\circ)$	...	$Y_r(X_2^\circ)$, $Z_r(X_2^\circ)$
.	.	.	...	.
.	.	.	...	.
.	.	.	...	.
r	$Y_1(X_r^\circ)$, $Z_1(X_r^\circ)$	$Y_2(X_r^\circ)$, $Z_2(X_r^\circ)$	...	$Y_r(X_r^\circ)$, $Z_r(X_r^\circ)$

Each region j compares the r optimal control solutions pairwise, from which, using Saaty's eigenvalue method, for each region j a vector of weights w_j is obtained (Saaty, 1977):

$$w_j = \begin{pmatrix} w_{j1} \\ w_{j2} \\ \cdot \\ \cdot \\ \cdot \\ w_{jr} \end{pmatrix} \qquad j=1,2,\ldots,r$$

(S7) By multiplying these by regional weights, $v_1, v_2, \ldots, v_r$, the final weights $q_1, q_2, \ldots, q_r$ are defined as:

$$\begin{pmatrix} q_1 \\ q_2 \\ \cdot \\ \cdot \\ \cdot \\ q_r \end{pmatrix} = \begin{pmatrix} w_1 & w_2 & \cdots & w_r \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ \cdot \\ v_r \end{pmatrix}$$

These regional weights are determined so as to optimize two criteria:

(i) Find a set of compromise weights $q_1, q_2, \ldots, q_r$ which are as near as possible to each individual set of regional weights separately, so as:

$$
\begin{aligned}
&\text{Min } d_{i1}^{+} + d_{i1}^{-} \\
&\text{Min } d_{i2}^{+} + d_{i2}^{-} \\
&\quad \vdots \qquad\qquad\qquad\qquad\qquad\qquad i=1,2,\ldots,r \\
&\text{Min } d_{ir}^{+} + d_{ir}^{-}
\end{aligned}
$$

$$
\begin{aligned}
q_1 &= w_{11}v_1 + w_{21}v_2 + \ldots + w_{r1}v_r \\
q_2 &= w_{12}v_1 + w_{22}v_2 + \ldots + w_{r2}v_r \\
&\vdots \\
q_r &= w_{1r}v_1 + w_{2r}v_2 + \ldots + w_{rr}v_r
\end{aligned}
$$

$$
\begin{aligned}
q_1 + d_{i1}^{+} - d_{i1}^{-} &= w_{i1} \\
q_2 + d_{i2}^{+} - d_{i2}^{-} &= w_{i2} \\
&\vdots \qquad\qquad\qquad\qquad i=1,2,\ldots,r \\
q_r + d_{ir}^{+} - d_{ir}^{-} &= w_{ir}
\end{aligned}
$$

$$
\begin{aligned}
&v_i \geqslant 0 && i=1,2,\ldots,r \\
&d_{ij}^{+},\ d_{ij}^{-} \geqslant 0 && j=1,2,\ldots,r;i=1,2,\ldots,r \\
&\sum_{i=1}^{r} v_i = 1 && (4)
\end{aligned}
$$

(ii) Maximize the priority value of the combination for each region:

$$
\begin{aligned}
&\text{Max } q_1w_{11} + q_2w_{12} + \ldots + q_rw_{1r} \\
&\text{Max } q_1w_{21} + q_2w_{22} + \ldots + q_rw_{2r} \\
&\vdots \\
&\text{Max } q_1w_{r1} + q_2w_{r2} + \ldots + q_rw_{rr} \qquad (5)
\end{aligned}
$$

(S8) Using the multicriteria linear programming perturbation method (M. Despontin, 1984), the multiple criteria problem (4), (5) is solved, giving final weights $q_1, q_2, \ldots, q_r$.

(S9) Penalty weights are now abtained through:

$$G = \sum_{i=1}^{r} q_i G_i \quad ; \quad H = \sum_{i=1}^{r} q_i H_i$$

by which:

$$\begin{cases} \text{Min} \displaystyle\sum_{t=1}^{T} [(Y_t - Y_t^*)' G (Y_t - Y_t^*) + (Z_t - Z_t^*) H (Z_t - Z_t^*)] \\ Y_t = R_1 Y_{t-1} + R_2 Z_t + R_3 W_t \qquad t=1,2,\ldots,T \\ Y_o = Y_o^* \end{cases}$$

can be solved, giving a compromise solution we will denote by X°_{r+1}.

(S10) The algorithm resumes with step (S6), but now comparing r+1 solutions, the original regional optimal solutions and the newly found compromise solution.

4. CONCLUSION

In this short text, we gave the principal components of an approach trying to combine both optimal control theory and multiple criteria decision making in order to solve large scale quantitative economic policy problems. It provides a framework to use the linear programming multiple criteria perturbation method, which should allow a flexible way to take account of the importance of the different regions and the impact on the overall national objectives.

Due to the very different nature of both methodologies, and taking into account the human involvement in multiple criteria decision making, it is clear that the effectiveness of the approach presented should be tested on real large scale models, which should be a next step.

5. REFERENCES

Dantzig, G.B. and Wolfe, P. (1961). The Decomposition Algorithm for Linear Programs, Econometrica, 29.

Despontin, M. (1981). Dynamic Optimization in a Multiregional Econometric Model for Belgium, in Brans, J.P. (ed.), Operational Research '81, North-Holland Publ. Co., Amsterdam.

Despontin, M. (1982). Regional Multiple Objective Quantitative Economic Policy: a Belgian Model, European Journal of Operational Research, vol. 10, 1.

Despontin, M. (1984). Interactive Economic Policy Formulation with Multiregional Econometric Models, in Despontin, M., Nijkamp, P. and Spronk, J. (eds.), Macro-Economic Planning with Conflicting Goals, Springer-Verlag, Berlin.

Kornai, J. and Liptak, T. (1965). Two Level Planning, Econometrica, 33.

Reimers, U. (1984). A Method for Solving the Decentralized Hierarchical Multiple Objective Decision Making Problem, Working paper nr. 154, Institut für Betriebswirtschaftslehre, Kiel.

Reimers, U. (1985). Koordination von Entscheidungen in hierarchischen Organisationen bei mehrfachen Zielsetzungen, Lang, Frankfurt am Main.

Saaty, T.L. (1977). A Scaling Method for Priorities in Hierarchical Structures, Journal of Math. Psych., June 1977.

Theil, H. (1958). Economic Forecasts and Policy, North-Holland Publ. Co., Amsterdam.

INTERACTIVE METHODS FOR MULTI-OBJECTIVE INTEGER LINEAR PROGRAMMING

J. TEGHEM Jr.
Faculté Polytechnique de Mons,
Rue de Houdain, 9, B - 7000 Mons (Belgium)
P.L. KUNSCH
BELGONUCLEAIRE S.A.,
Rue du Champ de Mars, 25, B - 1050 Brussels (Belgium)

1. INTRODUCTION AND GENERAL RESULTS

For the last 15 years, many Multi-Objective Linear Programming (MOLP) methods with continuous solutions have been developed. In many real world applications, however, discrete variables must be introduced representing, for instance, an investment choice, a production level, etc. The linear mathematical structure is then Integer Linear Programming (ILP), associated with MOLP giving a MOILP problem. Unfortunately, this type of problems has its own difficulties, as it cannot be solved by simply combining ILP and MOLP methods.

A previous paper [10] has been devoted to the characterizations of the set of efficient solutions to MOILP. In this paper, a survey of present interactive techniques for MOILP is attempted.

First, some important general results are pointed out, especially those related to the generalized Tchebycheff norm. In the following, a survey is made on several interactive methods to determine a best compromise among the set of efficient solutions. Sections 2 and 3 are devoted to all-integer and mixed integer linear programming problems, respectively. Some concluding remarks are presented in Section 4.

The problem (P) considered here is defined as follows :

$$\left[\begin{array}{l} \text{"max" } C\,X \\ X \in D = \left\{ X \in R^n \mid A\,X \leq d,\ X \geq 0,\ x_j \text{ integer } j \in J \right\} \end{array}\right. \qquad (P)$$

where C is a p x n matrix, and c_j ($j = 1,\dots,n$) and c^k ($k = 1,\dots,p$) are, respectively, the columns and the rows of C ;

A is a m x n matrix, and a_j ($j = 1,\dots,n$), the columns of A;

d is a m x 1 vector ;

J is a subset of $\left\{ 1,\dots,n \right\}$.

A solution X^* of D is called efficient, or non dominated, for problem (P), if there exists no other solution X in D, such that :

$$c^k X \geq c^k X^* \qquad k = 1,\dots,p$$

with at least one strict inequality holding.

Let us denote E(P), the set of all efficient solutions of problem (P). Different methods to determine and to characterize E(P) are reviewed in [10].

Let us just recall here that because of the presence of integer variables, the feasible solution space D is not convex, so that to generate E(P), it is no longer sufficient - contrary to MOLP - to solve the (P_α) programme :

$$\max \alpha' \cdot C X \qquad X \in D \qquad (P_\alpha)$$

for all values of α verifying

$$\alpha_k > 0 \qquad \forall k = 1,\dots,p$$

$$\sum_{k=1}^{p} \alpha_k = 1$$

The optimal solutions of (P_α) generate only a subset SE(P) of E(P), called the set of supported efficient solutions.

Nevertheless, Bowman [1] proved that E(P) is generated by the optimal solutions of the problem (P_λ), defined as :

$$\min_{X \in D} \| C X - \bar{Z} \|_\lambda$$

corresponding to all λ values such that $\lambda_k > 0$ and $\sum_{k=1}^{p} \lambda_k = 1$, and where $\bar{Z}$ is a goal point in the objective space with the properties

$$\bar{Z} \geq C X \qquad \forall X \in D$$

$$\| \in X - \bar{Z} \|_\lambda = \max_k \left\{ \lambda_k \mid c^k X - \bar{Z}_k \mid \right\}$$

is the generalized Tchebytcheff norm.

The existing interactive methods leading to the determination of a "best compromise" in E(P), according to the preferences of the decision-maker (D.M.), will be reviewed in the present paper.

For MOLP, there exists some interactive methods based on the following procedure :

- stepwise determination of a stability interval $[\underline{\alpha},\overline{\alpha}]$ of α values for the parametrized problem (P_α), consistent with the preferences of the D.M.;
- determination of the efficient solution corresponding to the optimal solution of (P_α), given a set of weights $\alpha \in [\underline{\alpha},\overline{\alpha}]$.

As pointed out by Zionts [15], this type of method cannot be directly extended to MOILP. The main reason is that for one single stability interval $[\underline{\alpha},\overline{\alpha}]$, (P_α) admits several integer solutions, as illustrated in the following example given in [15] :

$$p = 2 ,$$

$$Z_1 = x_1 \quad ; \quad Z_2 = x_2$$

$$D = \left\{ X \in R^2 \mid x_1 + \frac{1}{3} x_2 \leq \frac{25}{8} \; ; \; \frac{1}{3} x_1 + x_2 \leq \frac{25}{8} \; ; \; x_1 \geq 0 \geq x_2 \geq 0 \text{ integers} \right\}$$

Assuming that for the corresponding MOLP problem, the continuous best compromise is $x_1 = x_2 = 2.34$, corresponding to the stability interval $1/4 < \lambda_1 < 3/4$, there exists three completely different integer solutions for this same interval :

if $\frac{1}{4} < \lambda_1 < \frac{1}{3}$, then $x_1 = 0$, $x_2 = 3$

if $\frac{1}{3} < \lambda_1 < \frac{1}{3}$, then $x_1 = 2$, $x_2 = 2$

if $\frac{2}{3} < \lambda_1 < \frac{3}{4}$, then $x_1 = 3$, $x_2 = 0$

MOILP problems require thus specific interactive methods and the paper gives a survey of several of them, for all-integer problems and for mixed integer problems in Sections 2 and 3, respectively.

Some preliminary remarks must be made :

(i) only the methods related to general MOILP structure are reviewed. Several methods addressing specific combinatioral optimization problems (locational and scheduling problems, generalized networks, ...) have not been included in this survey ;

(ii) moreover, the present study is limited to published papers to be found in the open literature up to the end of 1985, but, of course, without any claim of exhaustivity. Technical reports (see Evans [2]) have been disregarded ;

(iii) some methods are considerably technical, therefore only the basic ideas of each approach are described.

2. MULTI-OJECTIVE ALL INTEGER LINEAR PROGRAMMING

This section corresponds to the case where $J \equiv \{1,\dots,n\}$ in problem (P).

2.1. Villareal-Karwan's method

This method [13] is based on an algorithm described in [12], by the same authors to characterize E(P). They use a dynamic programming approach : in the i-th stage, the $P_i(Y_i)$ problem is defined as :

$$\left[\begin{array}{l} \text{"max"} \sum_{j=1}^{i} c_j\, x_j \\ \sum_{j=1}^{i} a_j\, x_j \leq Y_i \\ x_j,\ \text{integers} \geq 0 \end{array}\right. \qquad (P_i(Y_i))$$

with $0 \leq Y_i \leq d$.

Let us call $E(P_i(Y_i))$, the set of efficient solutions of this subproblem.

The general principle of this method will consist in interactively selecting a set $\widetilde{S}_i(Y_i)$ of best subpolicies $(\widetilde{x}_1,\dots,\widetilde{x}_i)$ at each stage i and for each RHS vector Y_i, so that at the final stage, the D.M. will be able to select his best compromise within $S_n(d)$.

Villarreal and Karwan introduce the following conditions on the operators "preference" ($\succ$) and "indifference" ($\sim$) applicable to the set of policies :

- operators $\succ$ and $\sim$ are transitive and complete ;
- operator $\sim$ is symmetric and reflexive ;
- operator $\succ$ is asymmetric and irreflexive ;
- the following monotonicity property is satisfied :

$$(X^{(1)},X^{(3)}) \succ (X^{(2)},X^{(3)}) \qquad \text{iff} \qquad X^{(1)} \succ X^{(2)}$$

$$(X^{(1)},X^{(3)}) \sim (X^{(2)},X^{(3)}) \qquad \text{iff} \qquad X^{(1)} \sim X^{(2)}$$

Quite naturally, it is assumed that the D.M. prefers solution $X^{(1)}$ to $X^{(2)}$, whenever $X^{(1)}$ dominates $X^{(2)}$, i.e. if $C\,X^{(1)} \geq C\,X^{(2)}$ with at least one strict inequality.

It is then easy to prove that :

(1) a preferred subpolicy $(\tilde{x}_1,\ldots,\tilde{x}_i)$ of $\tilde{S}_i(Y_i)$ is a feasible solution of the problem $(P_i(Y_i))$ such that $(\tilde{x}_1,\ldots,\tilde{x}_{i-1})$ is a preferred subpolicy for stage (i-1) with the RHS vector $Y_{i-1} = Y_i - a_i\tilde{x}_i$, i.e. if it belongs to the set $S_{i-1}(Y_i - a_i\tilde{x}_i)$.

(2) $\tilde{S}_i(Y_i) \subseteq E(P_i(Y_i))$.

The basic procedure in [13] allows to reduce the effort by skipping some questions to the D.M. that would otherwise be necessary at each step i and for each RHS vector Y_i.

To this end, at each stage i, a set $\tilde{T}_i(Y_i)$ of potential preferred subpolicies is build.

$\tilde{T}_i(Y_i)$ is the set of optimal policies $(\tilde{x}_1,\ldots,\tilde{x}_i)$ of problem $(P_i(Y_i))$, such that $(\tilde{x}_1,\ldots,\tilde{x}_\ell)$, with $\ell < i$, is a preferred subpolicy for stage ℓ with the RHS vector $Y_\ell = Y_i - \sum_{j=+1}^{i} a_j \tilde{x}_j$, and

$$\tilde{S}_i(Y_i) \subseteq \tilde{T}_i(Y_i) \subseteq E(P_i(Y_i))$$

This way, the D.M.'s intervention is limited to the choice of subpolicies at some stage or for some RHS vector ; for instance, whenever too many potential preferred subpolicies are obtained,

i.e. if $|\tilde{T}_i(Y_i)| \geq \bar{T}$ given some level $\bar{T}$, or whenever too many stages are considered without any interactive phase,

i.e. if $|i - \ell| = \bar{L}$ given some level $\bar{L}$.

In the end, the best compromise is chosen in $\tilde{T}_n(d)$.

Remarks :

(i) As in [12], it is possible to decrease the storage requirements by an imbedded state approach using the concept of "resource efficient solution" ; a hybrid procedure can also be defined, incorporating bounding and fathoming criteria (see [10]).

(ii) The method can also be applied to some non linear objective functions (see [13]).

2.2. Gonzalez, Reeves and Franz algorithm

In this recent paper [3], the authors suggest to select a set $\tilde{S}$ of p efficient solutions, which is updated in each algorithm step according to the D.M.'s preferences. At the end of the procedure, $\tilde{S}$ will contain the most preferred solutions. The method is divided in two stages : in the first one, the supported efficient solutions are considered, while the second one deals with non-supported efficient solutions.

(a) Stage I - Determination of the best supported efficient solutions.

Step 0 :

In a preliminary step, the single objective ILP problems

$$\left[\begin{array}{l} \max c^k X \\ \\ \quad X \in D \end{array}\right.$$

are solved, successively for k=1,...,p.

At start, the set $\tilde{S}$ contains those optimal solutions.

Step j :

Let denote $\tilde{S} = \left\{ \tilde{X}^k,\ k=1,\dots,p \right\}$ and $\tilde{Z} = \left\{ \tilde{Z}^k = C\, X^k,\ k=1,\dots,p \right\}$ the set of the corresponding points in the objective space. For simplicity, index (j) is dropped in these notations. A linear direction of search G(X) is build : G(X) is the inverse mapping of the hyperplane defined by the p points $\tilde{Z}^k$ in the objective space into the decision space.

Note that this direction of search can be adjusted, if necessary (see [3], page 253).

A new supported efficient solution X* is determined by solving the single objective ILP problem :

$$\left[\begin{array}{l} \max G(X) \\ \\ \quad X \in D \end{array}\right. \qquad (P_a)$$

and $Z^* = C\, X^*$, is the corresponding p-vector.

Three different cases are possible :

(α) $Z^* \notin \tilde{Z}$ and the D.M. prefers solution X* to at least one solution X^k (k=1,...,p) : the least preferred solution $\tilde{X}^k$ is replaced in S by X*. A step (j+1) is initiated using the new set $\tilde{S}$.

(β) $Z^* \notin \tilde{Z}$ and X* is not preferred to any solution in $\tilde{S}$: $\tilde{S}$ is not modified and part b.2. is initiated.

(γ) $Z^* \in \tilde{Z}$: the set $\tilde{Z}$ defines a face of the efficient surface and part b.1. is initiated.

(b) Stage II - Introduction of the best non-supported solutions

b.1. Step j :

A non-supported efficient solution $\overline{X}$ is determined by solving the single ILP problem :

$$\left[\begin{array}{l} \max G(X) \\ X \in D \\ G(X) \leq \tilde{G} - \epsilon \end{array}\right. \qquad (P_b)$$

where, for $j > 1$, $\tilde{G}$ is the optimal value obtained for problem (P_b) at step j-1, or, for j=1, $\tilde{G}$ is the optimal value obtained for problem (P_a) at the last step of part (a) ; ϵ is chosen by the analyst small enough so that no relevant unsupported solution will be discarded.

The two following cases can occur :

(α) The D.M. prefers solution $\overline{X}$ to at least one solution $\tilde{X}^k$: this solution is replaced in $\tilde{S}$ by $\overline{X}$. Step (j+1) is initiated with the new set $\tilde{S}$ and the corresponding new direction G(X).

(β) Otherwise, the procedure stops and the D.M. selects his best compromise in $\tilde{S}$.

b.2. Further research should only aim in finding non-supported solutions around the most preferred solution of $\tilde{S}$. Problem (P_b) is solved once more with G(X) being the objective function that generated this most preferred solution.

2.3. Other methods

(a) White [14] presented recently a Lagrangian relaxation approach for the general multi-objective programming problem

$$\left[\begin{array}{l} \text{"max" } F(X) \\ X \in D = \left\{ X \mid H(X) \leq 0,\ X \text{ integers} \geq 0 \right\} \end{array}\right.$$

where F(X) and H(X) are vectors of real functions f_k, k=1,...,p and $h_i(X)$, i=1,...,m, respectively.

The author extends the concept of Lagrangian relaxation by introducing the following problem :

$$\text{"max" } L_\lambda(X) = F(X) - \lambda H(X) \cdot e$$

$$X \text{ integers} \geq 0$$

where $\lambda \in R^p$ and non-negative ;

e is the m sum vector of ones ;

this new problem provides vector bounds for F(X) to assist in the fathoming of nodes in an interactive Branch and Bound method.

(b) Most of the methods described in [10] and in the present paper use an implicit enumeration approach to handle the discrete variables. In single objective ILP problems, a cutting plane approach could be applied as well. There is a remark by Zionts [15,16] that such an approach does not appear to be promising for MOILP. Nevertheless, the interactive method developed by Musselman and Talavage [7] for MOLP uses a cutting plane procedure to progress to a best compromise, and these authors claim that their method can be adapted to MOILP.

(c) For the sake of completeness, let us mention that in an older paper, Jaikumar and Fisheries [5] described very shortly a "heuristic" interactive algorithm for MOILP problem with 0-1 variables. Their paper is only an abridgment.

3. MULTI-OBJECTIVE MIXED INTEGER LINEAR PROGRAMMING

3.1. Zionts method and some extensions

The Zionts-Wallenius method is a well known interactive procedure for MOLP (see Evans [2]). Let us recall that this method is based on the linear relaxation of problem (P_α) ; it is thus assumed that the decision-maker has a utility function in the back of his mind, which will be discovered during the procedure of questions and answers relative to his preferences. Following [15], let us briefly recall the main features :

(1) Choose an arbitrary p-vector $\alpha > 0$.

(2) First, solve the linear relaxation of problem (P_α), the solution of which is efficient. Identify the adjacent efficient extreme points in the space of the objective functions towards which the D.M. has no negative attitude. If there are none of them, the optimal solution has been found and the procedure stops. The marginal rates of change in the objectives from the given point to an adjacent one is a trade off offer, and the corresponding question is called an efficient question.

(3) Ask the D.M. if he likes or dislikes the trade off offer implied in each efficient question.

(4) Find a set of α consistent with all current and previous answers of the D.M. Go to step (2).

Even though problem (P_α) only generates supported efficient solutions, several studies attempt to extend the Zionts-Wallenius procedure to MOILP : the basic procedure is proposed by Zionts [15] and Villareal-Karwan-Zionts [11] ; some improvements have been recently given by Karwan-Zionts-Villareal-Ramesh [6].

(a) Basic procedure

The initial step consists in solving the linear relaxation of problem (P) using the Zionts-Wallenius method : a polyhedral set of possible α- values compatible with the D.M.'s preferences, is obtained, and a continuous best compromise is found.

If this solution is not integer, a branch and bound scheme is applied - based on classical rules and using down and up penalties, to select the branching variable. This is done until both the D.M.'s satisfaction and the achievement of an integer solution. In the branch and bound scheme, a node is fathomed if the following conditions hold :

(i) the D.M. prefers some other known integer solution ;
(ii) all the efficient trade off questions (if any) associated with the solution are viewed negatively or with indifference by the D.M.

During the procedure, the vector α is updated each time whenever it is not consistent with all prior answers of the D.M.

Some adaptations assist in reducing the storage requirements and the number of questions to be asked (see [11]).

(b) Additional improvements

The first application results were not very promising (see [6], page 264), so that three types of improvements have been developed in [6] :

(α) As the set of constraints on the weight α_k grows with the number of answers, a technique is used to eliminate redundant constraints.

(β) A better approach is given to determine a new feasible value of α in the polyhedral set consistent with the preferences of the D.M., since previously the new α value was quite often hardly different from the previous one. A technique maximizing the minimum slacks of the constraints imposed on the weights is proposed in order to find a "most consistent" or "middle most" set of weights.

(γ) The determination method of an initial incumbent solution is modified because, in many cases, the integer solution obtained from some heuristic did not compare well with the best solution obtained at a later stage. An improved method determines all rounded solutions obtained from the solution of the linear relaxation (each basic variable is rounded up and down), and compares them using the initial weights to obtain a first incumbent solution.

3.2. Steuer-Choo's method

Several approaches of more general MCDM problems can also be applied to MOILP ; among them, let us mention only the Steuer-Choo's method [9], which is valid for a general multi-objective programming problem of the form :

$$\left[\begin{array}{l} \text{"max" } Z = F(X) \\ \quad X \in D \end{array}\right.$$

where F(X) is a vector of functions $f_k(X)$ $(k=1,\ldots,p)$ with f_k being possibly non linear and D non convex.

This method is the result of several other techniques published previously (see references in [9]). It is based on the augmented generalized Tchebycheff norm defined as :

$$||| Z - \bar{Z} |||_{\lambda} = || Z - \bar{Z} ||_{\lambda} + \rho \cdot e'(\bar{Z} - Z)$$

where :

$\bar{Z}$ is a goal point in the objective space ;

- $|| Z - \bar{Z} ||_{\lambda}$ is the generalized Tchebycheff norm defined in Section 1 of this paper ;

- λ is such that $\lambda_k > 0$ and $\sum_{k=1}^{p} \lambda_k = 1$;

- e' is the sum vector of ones ;

- ρ is a sufficiently small positive scalar number (see section 3 of [9]).

Initialization

The chosen goal point $\bar{Z}$ is such that

$$\bar{z}_k = \max_{X \in D} f_k(X) + \epsilon_k$$

with $\epsilon_k \geq 0$; yet, ϵ_k has to be strictly positive if the following applies :

(i) either there is more than one efficient solution that minimizes the k-th objective, or
(ii) the only efficient solution that maximizes the k-th objective also maximizes one of the other objectives.

The first iteration uses a widely dispersed group of λ weighting vectors to sample the set of efficient solutions. The sample is obtained by solving the problem :

$$\min_{X \in D} \; ||| F(X) - \bar{Z} |||_{\lambda}$$

for each of the λ values in the set.

Then the D.M. is asked to identify the most preferred solution $X^{(1)}$ among the sample.

Iteration j

A more refined grid of weighting vectors λ is used to sample the set of efficient solutions in the objective space in the neighbourhood of the point $Z^{(j)} = F^{(j)}$. Again the sample is obtained by minimizing the augmented generalized Tchebytcheff norm and the most preferred solution $X^{(j+1)}$ is selected.

The procedure continues using increasingly finer sampling until the solution is deemed to be acceptable.

3.3. Adaptation of the STEM method

In our view, the STEM method of Benayoun et al. (see Evans [2]) remains one of the most efficient interactive method for MOLP. At each step, this method determines a new compromise solution which minimizes the maximal weighted distance to an ideal point in the objective space. It has therefore to be considered as a special case of the generalized Tchebycheff norm.

Taking into account Bowman's result presented in Section 1, an extension of STEM to MOILP seems to be possible without too many difficulties. Huckert et al. [4] and recently, Slowinski and Weglarz [8] have applied STEM to multi-objective scheduling problems. However, no general formulation of this concept has been given.

4. SOME CONCLUDING REMARKS

Interactive methods are especially important in MOILP problem solving because finding all efficient solutions does not seem to be a practicable nor an easy task (see [10]).

The questions to be asked to the D.M. are of outright importance. Their number should not be too large, and they should not be too cumbersome to answer.

With regard to these two aspects, we can criticize several existing procedures as follows :

- The method of Villareal-Karwan [13] requires much from the D.M. As pointed out by the authors [13, page 526], they assume that the D.M. is able to identify a preferred partial solution $(\tilde{x}_1,\dots,\tilde{x}_i)$ - sometimes already at a stage $i \ll n$ - and in our opinion, this should be the exception rather than the rule in real case studies.
- The Zionts-Wallenius method applicable to MOLP has the drawback that the analyst has to ask many questions to the D.M. In its extension to MOILP [15,11,6], this can take dramatic dimensions, because the number of questions increases very fast with the number of nodes in the branch and bound scheme. This drawback is analysed by the authors (see [11, 6]). Moreover, the questions are related to trade offs between the objectives and we feel that many times, it is not an easy task for the D.M. to evaluate them accurately. Let us add that the assumption of an implicit linear utility function is restrictive, particularly for MOILP.

On the contrary, some procedures are privileged :

- the method of Gonzalez et al. [3] is well adapted to interactive dialogues : the D.M. can easily tell, if he prefers a new solution X^* to at least one out of p solution $\tilde{X}^k$, $k=1,\dots,p$. Another favourable aspect of this method is that a unique objective ILP method at each step, has to be solved. Unfortunately, in our view, the distinction made between supported and non supported solutions seems to be somewhat artificial ;

- the procedure of Steuer-Choo [9], besides the fact that it is very general, has comparable advantages : the D.M. has to choose his most preferred solution from a set of efficient solutions. This should not be too difficult even if this set can sometimes be very large. Nevertheless, by its random generation of efficient solutions, this method is perhaps more suitable for general non linear problems.

- A STEM-like approach or methods based on the Tchebycheff norm could provide practitioners with useful instruments, keeping both the D.M.'s and the numerical effort at a reasonable level. We strongly feel the need for further work in that direction.

* * *

ACKNOWLEDGEMENT

We are indebted to Miss Lisbeth Rasmussen from the Danish Defense Research Establishment of Copenhagen, for having passed to us some useful references in 0-1 multi-objective linear programming.

* * *

REFERENCES

[1] V.J. BOWMAN Jr., "On the Relationship of the Tchebycheff Norm and the Efficient Frontier of Multiple-Criteria Objectives", in : H. Thiriez, S. Zionts, Eds. Multiple Criteria Decision Making (Springer Verlag, Berlin, 1976), pp. 76-85.

[2] G.W. EVANS, "Overview of Techniques for solving Multi-Objective Mathematical Programs", in : Mn.Sc. Vol 30, 11, (1984), pp. 1268-1282.

[3] J.J. GONZALEZ, G.R. REEVES, L.S. FRANZ, "An Interactive Procedure for solving Multiple Objective Integer Linear Programming Problems, in : Y. Haimes, V. Chankong, Eds., Decision Making with Multiple Objectives (Springer Verlag, 1985), pp. 250-260.

[4] K. HUCKERT, R. RHODE, O. ROGLIN, R. WEBER, "On the Interactive Solution to a Multi-Criteria Scheduling Problem", in : Z.Oper.Res., Vol 24 (1980), pp. 47-60.

[5] R. JAIKUMAR, B. FISHERIES, "A Heuristic 0-1 Algorithm with Multiple Objectives and Constraints", in : J.L. Cochrane, M. Zeleny, Eds., Multiple Criteria Decision Making (University of South Carolina Press, 1973), pp. 745-748.

[6] M.H. KARWAN, S. ZIONTS, B. VILLARREAL, R. RAMESH, "An Improved Interactive Multi-Criteria Integer Programming Algorithm", in : Y. Haimes, V. Chankong, Eds., Decision Making with Multiple Objectives (Springer Verlag, 1985), pp. 261-271.

[7] K.J. MUSSELMAN, J. TALAVAGE, "A Trade Off Cut Approach to Multiple Objective Optimization", in : Oper.Res. 28 (1980), pp. 1424-1435.

[8] R. SLOWINSKI, J. WEGLARZ, "An Interactive Algorithm for Multi-Objective Precedence and Resource Constrained Scheduling Problems", in : W. Urietholl, J. Visser, H.H. Boerma, Eds., Proceedings 8th International Congress (North Holland, 1985), pp. 866-873.

[9] R.E. STEUER, E.U. CHOO, "An Interactive Method Weighted Tchebycheff Procedure for Multiple Objective Programming", in : Math.Progr. 26 (1983), pp. 326-344.

[10] J. TEGHEM Jr., P.L. KUNSCH, "Characterization of Efficient Solutions for Multi-Objective Integer Linear Programming", in : Technical Report (Faculté Polytechnique de Mons, Belgium, 1985), submitted for publication.

[11] B. VILLARREAL, M. KARWAN, S. ZIONTS, "An Interactive Branch and Bound Procedure for Multicriterion Integer Linear Programming", in : Fandel, T. Gal, Eds., Multiple Criteria Decision Making, Theory and Application (Springer Verlag, 1980), pp. 448-467.

[12] B. VILLARREAL, M. KARWAN, "Multi-Criteria Integer Programming : a (Hybrid) Dynamic Programming Recursive Approach", in : Mathematical programming, 21 (1981), pp. 204-223.

[13] B. VILLARREAL, M. KARWAN, "An Interactive Dynamic Programming Approach to Multi-Criteria Discrete Programming", in : Journal of Mathematical Analysis and Applications, 81 (1981), pp. 524-544.

[14] D.J. WHITE, "A Multiple Objective Interactive Lagrangean Relaxation Approach", in : European Journal of Operational Research, 19 (1985), pp. 82-90.

[15] S. ZIONTS, "Integer Linear Programming with Multiple Objectives", in : Annals of Discrete Mathematics 1 (1977), pp. 551-562, North Holland.

[16] S. ZIONTS, "A Survey of Multiple Criteria Integer Programming Methods", in : Annals of Discrete Mathematics 5 (1979), pp. 389-398, North Holland.

* * *

INTERACTIVE DECISION MAKING FOR MULTIOBJECTIVE LINEAR PROGRAMMING PROBLEMS WITH FUZZY PARAMETERS †

Masatoshi Sakawa[1] and Hitoshi Yano[2]

1 Department of Systems Engineering, Faculty of Engineering, Kobe University, Kobe 657, Japan
2 Department of Information Science, College of Economics, Kagawa University, Kagawa 760, Japan

In this paper, we focus on multiobjective linear programming problems with fuzzy parameters and present a new interactive decision making method for obtaining the satisficing solution of the decision maker (DM) on the basis of the linear programming method. The fuzzy parameters in the objective functions and the constraints are characterized by fuzzy numbers. The concept of α-Pareto optimality is introduced in which the ordinary Pareto optimality is extended based on the α-level sets of the fuzzy numbers. In our interactive decision making method, in order to generate a candidate for the satisficing solution which is also α-Pareto optimal, if the DM specifies the degree α of the α-level sets and the reference objective values, the minimax problem is solved by making use of the linear programming method, and the DM is supplied with the corresponding α-Pareto optimal solution together with the trade-off rates among the values of the objective functions and the degree α. Then by considering the current values of the objective functions and α as well as the trade-off rates, the DM acts on this solution by updating his/her reference objective values and/or degree α. In this way the satisficing solution for the DM can be derived efficiently from among an α-Pareto optimal solution set. A numerical example illustrates various aspects of the results developed in this paper.

1. INTRODUCTION

In multiobjective decision making problems, multiple objectives are usually noncommensurable and cannot be combined into a single objective. Moreover, the objectives usually conflict with each other in that any improvement of one objective can be achieved only at the expense of another. Consequently, the aim is to find a compromise or satisficing solution of a decision maker (DM) which is also Pareto optimal based on his/her subjective value-judgement. However, when formulating the multiobjective programming problem which closely describes and represents the real decision situation, various factors of the real system should be reflected in the description of the objective functions and the constraints. Naturally these objective functions and the constraints involve many parameters whose possible values may be assigned by the experts. In the conventional approach, such parameters are fixed at some values in an experimental and/or subjective manner through the experts' understanding of the nature of the parameters.

---

† This research was partially supported by the Kajima Foundation's Research Grant.

In most practical situations, however, it is natural to consider that the possible values of these parameters are often only ambiguously known to the experts. In this case, it may be more appropriate to interpret the experts' understanding of the parameters as fuzzy numerical data which can be represented by means of fuzzy subsets of the real line known as fuzzy numbers (Dubois and Prade 1978,1980). The resulting multiobjective programming problem involving fuzzy parameters would be viewed as the more realistic version of the conventional one.

Recently, Tanaka and Asai (1981,1984) formulated the multiobjective linear programming problems with fuzzy parameters. Following the fuzzy decision or minimum operator proposed by Bellman and Zadeh (1970) together with triangular membership functions for fuzzy parameters, they considered two types of fuzzy multiobjective linear programming problems; one is to decide the nonfuzzy solution and the other is to decide the fuzzy solution.

More recently, Orlovski (1984) formulated general multiobjective nonlinear programming problems with fuzzy parameters. He presented two approaches to the formulated problems by making systematic use of the extension principle of Zadeh (1975) and demonstrated that there exist in some sense equivalent nonfuzzy formulations.

In this paper, we focus on the multiobjective linear programming problems with fuzzy parameters characterized by fuzzy numbers and introduce the concept of α-Pareto optimality by extending the ordinary Pareto optimality on the basis of the α-level sets of the fuzzy numbers. Then an interactive decision making method to derive the satisficing solution of the DM efficiently from among an α-Pareto optimal solution set is presented on the basis of the linear programming method as a generalization of the results obtained in Sakawa (1983a,1983b) and Sakawa et al. (1983,1984).

2. α-PARETO OPTIMALITY

Consider multiobjective linear programming (MOLP) problems of the following form:

$$\left.\begin{array}{l} \min \quad (c_1x, c_2x, \ldots, c_kx) \\ \text{subject to} \\ x \in X = \{ x \in E^n \mid a_jx \leq b_j, \ j=1,\ldots,m; \ x \geq 0 \} \end{array}\right\} \quad (1)$$

where x is an n-dimensional column vector of decision variables, $c_1, c_2, \ldots, c_k$ are n-dimensional cost factor row vectors, $a_1, a_2, \ldots, a_m$ are n-dimensional constraint row vectors and $b_1, b_2, \ldots, b_m$ are constants.

Fundamental to the MOLP is the Pareto optimal concept, also known as a noninferior solution. Qualitatively, a Pareto optimal solution of the MOLP is one where any improvement of one objective function can be achieved only at the expense of another.

Definition 1. (Pareto optimal solution)

$x^* \in X$ is said to be a Pareto optimal solution to the MOLP, if and only if there does not exist another $x \in X$ such that $c_ix \leq c_ix^*$, $i=1,\ldots,k$ with strict inequality holding for at least one i.

In practice, however, it would certainly be appropriate to consider that the possible values of the parameters in the description of the objective functions and the constraints usually involve the ambiguity of the experts' understanding of the real system. For this reason, in this paper, we consider the following fuzzy multiobjective linear programming (FMOLP) problem involving fuzzy parameters:

$$\left.\begin{aligned} &\min \quad (\tilde{c}_1 x , \tilde{c}_2 x , \ldots, \tilde{c}_k x) \\ &\text{subject to} \quad x \in X(\tilde{a},\tilde{b}) \triangleq \{ x \in E^n \mid \tilde{a}_j x \leqq \tilde{b}_j , j=1,\ldots,m ; x \geqq 0 \} \end{aligned}\right\} \quad (2)$$

Here $\tilde{c}_i = (\tilde{c}_{i1},\ldots,\tilde{c}_{in})$, and $\tilde{a}_j = (\tilde{a}_{j1},\ldots,\tilde{a}_{jn})$, $\tilde{b}_j$ represent respectively fuzzy parameters involved in the objective function $\tilde{c}_i x$ and the constraint $\tilde{a}_j x \leqq \tilde{b}_j$.

These fuzzy parameters are assumed to be characterized as the fuzzy numbers introduced by Dubois and Prade (1978,1980). It is appropriate to review here that a real fuzzy number $\tilde{p}$ is a convex continuous fuzzy subset of the real line whose membership function $\mu_{\tilde{p}}(p)$ is defined as:

(1) A continuous mapping from E^1 to the closed interval $[0,1]$,
(2) $\mu_{\tilde{p}}(p) = 0$ for all $p \in (-\infty,p_1]$,
(3) Strictly increasing on $[p_1,p_2]$,
(4) $\mu_{\tilde{p}}(p) = 1$ for all $p \in [p_2,p_3]$,
(5) Strictly decreasing on $[p_3,p_4]$,
(6) $\mu_{\tilde{p}}(p) = 0$ for all $p \in [p_4,+\infty)$.

We now assume that $\tilde{c}_{i1},\ldots,\tilde{c}_{in}$, $\tilde{a}_{j1},\ldots,\tilde{a}_{jn}$ and $\tilde{b}_j$ in the FMOLP are fuzzy numbers whose membership functions are $\mu_{\tilde{c}_{i1}}(c_{i1}),\ldots, \mu_{\tilde{c}_{in}}(c_{in})$, $\mu_{\tilde{a}_{j1}}(a_{j1}),\ldots, \mu_{\tilde{a}_{jn}}(a_{jn})$ and $\mu_{\tilde{b}_j}(b_j)$ respectively. For simplicity in the notation, define the following vectors:

$$c = (c_1,\ldots,c_k), \quad \tilde{c} = (\tilde{c}_1,\ldots,\tilde{c}_k) ,$$
$$a = (a_1,\ldots,a_m), \quad \tilde{a} = (\tilde{a}_1,\ldots,\tilde{a}_m) , \quad b = (b_1,\ldots,b_m), \quad \tilde{b} = (\tilde{b}_1,\ldots,\tilde{b}_m) .$$

Then we can introduce the following α-level set or α-cut (Dubois and Prade 1980) of the fuzzy numbers $\tilde{a}_{jr}$, $\tilde{b}_j$ and $\tilde{c}_{ir}$.

<u>Definition 2.</u> (α-level set)

The α-level set of the fuzzy numbers $\tilde{a}_{jr}, \tilde{b}_j$ and $\tilde{c}_{ir}$ is defined as the ordinary set $L_\alpha(\tilde{a},\tilde{b},\tilde{c})$ for which the degree of their membership functions exceeds the level α:

$$L_\alpha(\tilde{a},\tilde{b},\tilde{c}) = \{(a,b,c) \mid \mu_{\tilde{a}_{jr}}(a_{jr}) \geqq \alpha, \ \mu_{\tilde{b}_j}(b_j) \geqq \alpha, \ \mu_{\tilde{c}_{ir}}(c_{ir}) \geqq \alpha, \ i=1,\ldots,k, \ j=1,\ldots,m, \ r=1,\ldots,n \} \quad (3)$$

It is clear that the level sets have the following property:

$$\alpha_1 \leqq \alpha_2 \text{ if and only if } \quad L_{\alpha_1}(\tilde{a},\tilde{b},\tilde{c}) \supset L_{\alpha_2}(\tilde{a},\tilde{b},\tilde{c}) \quad (4)$$

For a certain degree α, the FMOLP (2) can be understood as the following nonfuzzy α-multiobjective linear programming (α-MOLP) problem.

$$\left.\begin{array}{ll} \min & (c_1x, c_2x ,\ldots, c_kx) \\ \text{subject to} & x \in X(a,b) \underline{\Delta} \{x \in E^n \mid a_jx \leqq b_j ,j=1,\ldots,m ; x \geqq 0 \} \\ & (a,b,c) \in L_\alpha(\tilde{a},\tilde{b},\tilde{c}) \end{array}\right\} \quad (5)$$

It should be emphasized here that in the α-MOLP the parameters (a,b,c) are treated as decision variables rather than constants.

On the basis of the α-level sets of the fuzzy numbers, we introduce the concept of α-Pareto optimal solutions to the α-MOLP.

Definition 3. (α-Pareto optimal solution)

$x^* \in X(a,b)$ is said to be an α-Pareto optimal solution to the α-MOLP (5), if and only if there does not exist another $x \in X(a,b)$, $(a,b,c) \in L_\alpha(\tilde{a},\tilde{b},\tilde{c})$ such that $c_ix \leqq c_i^*x^*$, $i=1,\ldots,k$, with strict inequality holding for at least one i, where the corresponding values of parameters (a^*,b^*,c^*) are called α-level optimal parameters.

It is significant to note here that from the property of the α-level set, the following relation holds, where $X^P(\alpha)$ denotes the set of α-Pareto optimal solutions for the fixed level α.

Proposition 1.

If $\alpha_1 \geqq \alpha_2$, then for any $x^1 \in X^P(\alpha_1)$, there exists $x^2 \in X^P(\alpha_2)$ such that $cx^1 \geqq cx^2$.

3. MINIMAX PROBLEMS

In order to generate a candidate for the satisficing solution which is also α-Pareto optimal, the DM is asked to specify the degree α of the α-level set and the reference levels of achievement of the objective functions, called reference levels. Observe that the idea of the reference levels or the reference point was first appeared in Wierzbicki (1979). For the DM's degree α and reference levels $\bar{z}_i$, $i=1,\ldots,k$, the corresponding α-Pareto optimal solution, which is in a sense close to his/her requirement (or better, if the reference levels are attainable) is obtained by solving the following minimax problem.

$$\min_{\substack{x \in X(a,b) \\ (a,b,c) \in L_\alpha(\tilde{a},\tilde{b},\tilde{c})}} \max_{1 \leqq i \leqq k} (c_ix - \bar{z}_i) \quad (6)$$

or equivalently

$$\min_{x,v,a,b,c} \quad v \quad (7)$$

$$\text{subject to} \quad c_ix - \bar{z}_i \leqq v, \quad i=1,\ldots,k \quad (8)$$

$$a_jx \leqq b_j , \ j=1,\ldots,m, \quad x \geqq 0 \quad (9)$$

$$\mu_{\tilde{a}_{jr}}(a_{jr}) \geqq \alpha, \quad \mu_{\tilde{b}_j}(b_j) \geqq \alpha, \quad \mu_{\tilde{c}_{ir}}(c_{ir}) \geqq \alpha,$$
$$i=1,\dots,k, \; j=1,\dots,m, \; r=1,\dots,n \tag{10}$$

In this formulation, however, constraints (8) and (9) are nonlinear because the parameters a,b and c are treated as decision variables. In order to deal with such nonlinearlities, we first introduce the following set-valued functions $S_i(.)$ and $T_j(.,.)$.

$$S_i(c_i) = \{ (x,v) \mid c_i x - \bar{z}_i \leqq v \} \tag{11}$$

$$T_j(a_j,b_j) = \{ x \mid a_j x \leqq b_j \} \tag{12}$$

Then it can be verified that the following relations hold for $S_i(.)$ and $T_j(.,.)$, when $x \geqq 0$.

Proposition 2

(1) If $c_i^1 \leqq c_i^2$, then $S_i(c_i^1) \supset S_i(c_i^2)$

(2) If $a_j^1 \leqq a_j^2$, then $T_j(a_j^1,b_j) \supset T_j(a_j^2,b_j)$

(3) If $b_j^1 \leqq b_j^2$, then $T_j(a_j,b_j^1) \subset T_j(a_j,b_j^2)$

Now from the properties of the α-level set for the vectors of fuzzy numbers $\tilde{c}_i,\tilde{a}_j$ and the fuzzy numbers $\tilde{b}_j$, it should be noted here that the feasible regions for c_i,a_j and b_j can be denoted respectively by the intervals $[c_{i\alpha}^L,\ c_{i\alpha}^R]$, $[a_{j\alpha}^L,\ a_{j\alpha}^R]$ and $[b_{j\alpha}^L,\ b_{j\alpha}^R]$.

Therefore we can obtain an optimal solution to (7)-(10) by solving the following linear programming problem.

$$\left.\begin{array}{lll} \min & v & \\ \text{subject to} & c_{i\alpha}^L x - \bar{z}_i \leqq v , & i=1,\dots,k \\ & a_{j\alpha}^L x \leqq b_{j\alpha}^R , & j=1,\dots,m , \quad x \geqq 0 \end{array}\right\} \tag{13}$$

The relationships between the optimal solutions to (13) and the α-Pareto optimal concept of the α-MOLP can be characterized by the following theorems.

Theorem 1.

If x* is a unique optimal solution to (13), then x* is an α-Pareto optimal solution to the α-MOLP.

(Proof)

Assume that x* is not an α-Pareto optimal solution to the α-MOLP, then there exist $x \in X(a,b)$ and $(a,b,c) \in L_\alpha(\tilde{a},\tilde{b},\tilde{c})$ such that $c_i x \leqq c_{i\alpha}^L x^*$, i=1,,,.,k, with strict inequality holding for at least one i. Then it holds that

$$\max_{1\leqq i\leqq k} (c_{i\alpha}^L x - \bar{z}_i) \leqq \max_{1\leqq i\leqq k} (c_i x - \bar{z}_i)$$
$$\leqq \max_{1\leqq i\leqq k} (c_{i\alpha}^L x^* - \bar{z}_i)$$

which contradicts the fact that x^* is a unique optimal solution to (13).

Theorem 2.

If x^* is an α-Pareto optimal solution and (a^*,b^*,c^*) is an α-level optimal parameter to the α-MOLP, then x^* is an optimal solution to (13) for some $\bar{z} = (\bar{z}_1,\ldots,\bar{z}_k)$.

(Proof)

Assume that x^* is not an optimal solution to (13) for any $\bar{z}$ satisfying

$$c_1^* x^* - \bar{z}_1 = \ldots = c_k^* x^* - \bar{z}_k = v^* .$$

Then there exist $x \in X(a,b)$ and $(a,b,c) \in L_\alpha(\tilde{a},\tilde{b},\tilde{c})$ such that

$$c_i^* x^* - \bar{z}_i > c_i x - \bar{z}_i \geqq c_{i\alpha}^L x - \bar{z}_i$$

This implies that $c_{i\alpha}^L x < c_i^* x^*$ which contradicts the fact that x^* is an α-Pareto optimal solution to the α-MOLP .

It should be noted here that for generating α-Pareto optimal solutions using Theorem 1 , uniqueness of solution must be verified. In general, however, it is not easy to check numerically whether an optimal solution to (13) is unique or not.

Consequently, in order to test the α-Pareto optimality of a current optimal solution x^*, we formulate and solve the following linear programming problem :

$$\left.\begin{array}{lll} & \max & \sum_{i=1}^{k} \varepsilon_i \\ \text{subject to} & c_{i\alpha}^L x + \varepsilon_i = c_{i\alpha}^L x^* , & \varepsilon_i \geqq 0 , \; i=1,\ldots,k \\ & a_{j\alpha}^L x \leqq b_{j\alpha}^R , \; j = 1,\ldots,m , & x \geqq 0 . \end{array}\right\} \quad (14)$$

Let $\bar{x}$ and $\bar{\varepsilon}$ be an optimal solution to (14). If all $\bar{\varepsilon}_i = 0$, then x^* is an α-Pareto optimal solution. If at least one $\bar{\varepsilon}_i > 0$, it can be easily shown that $\bar{x}$ is an α-Pareto optimal solution.

4. INTERACTIVE ALGORITHM

Now given the α-Pareto optimal solution for the degree α and the reference levels specified by the DM by solving the corresponding minimax problem, the DM must either be satisfied with the current α-Pareto optimal solution, or update his/her reference levels and/or the degree α .
In order to help the DM express his/her degree of preference, trade-off information between a standing objective function and each of the other objective functions as well as between the degree α and the objective functions is very useful. Such a trade-off information is easily obtainable since it is closely related to the simplex multipliers of the problem (13).

To derive the trade-off information, we define the Lagrangian function L for the problem (13) as follows ;

$$L = v + \sum_{i=1}^{k} \Pi_i (c_{i\alpha}^L x - \bar{z}_i) + \sum_{j=1}^{m} \lambda_j (a_{j\alpha}^L x - b_{j\alpha}^R) \quad (15)$$

If all $\Pi_i > 0$ for each i, then by extending the results in Haimes and

Chankong (1979), it can be proved that the following expression holds (Yano and Sakawa 1985).

$$\frac{\partial(c_1 x)}{\partial(c_i x)} = -\frac{\Pi_i}{\Pi_1}, \quad i=2,\ldots,k \tag{16}$$

Regarding a trade-off rate between $c_i x$ and α, if all $\Pi_i > 0$, $i=1,\ldots,k$ and $\lambda_j > 0$, $j=1,\ldots,m$, then the following relation holds based on the sensitivity theorem (e.g. Fiacco 1983).

$$\frac{\partial(c_i x)}{\partial\alpha} = \sum_{i=1}^{k} \Pi_i \frac{\partial(c_{i\alpha}^L)}{\partial\alpha} x + \sum_{j=1}^{m} \lambda_j \left\{ \frac{\partial(a_{j\alpha}^L)}{\partial\alpha} x - \frac{\partial(b_{j\alpha}^R)}{\partial\alpha} \right\}, \quad i=1,\ldots,k \tag{17}$$

It should be noted here that in order to obtain the trade-off rate information from (16) or (17), all the constraints of the problem (13) must be active for the current optimal solution.

Following the above discussions, we can now construct the interactive algorithm in order to derive the satisficing solution for the DM from among the α-Pareto optimal solution set. The steps marked with an asterisk involve interaction with the DM.

Step 0. Calculate the individual minimum and maximum of each objective function under given constraints for $\alpha = 0$ and $\alpha = 1$.

Step 1*. Ask the DM to select the initial value of α ($0 < \alpha < 1$) and the initial reference levels $\bar{z}_i$, $i=1,\ldots,k$.

Step 2. For the degree α and the reference levels specified by the DM, solve the minimax problem and perform the α-Pareto optimality test.

Step 3*. The DM is supplied with the corresponding α-Pareto optimal solution and the trade-off rates between the objective functions and the degree α. If the DM is satisfied with the current objective function values of the α-Pareto optimal solution, stop. Otherwise, the DM must update the reference levels and/or the degree α by considering the current values of the objective functions and α together with the trade-off rates between the objective functions and the degree α, and return to step 2. Here it should be stressed for the DM that (1) any improvement of one objective function can be achieved only at the expense of at least one of the other objective functions for some fixed degree α, and (2) the greater value of the degree α gives worse values of the objective functions for some fixed reference levels.

5. NUMERICAL EXAMPLE

To clarify the concept of α-Pareto optimality as well as the proposed method, consider the following three objective linear programming problem.

$$\left.\begin{aligned} &\min \quad (2x_1 + \tilde{c}_1 x_2 ,\ -3x_1 + \tilde{c}_2 x_2 ,\ \tilde{c}_3 x_1 - x_2) \\ &\text{subject to} \quad x \in X = \{(x_1,x_2) \mid 3x_1 + x_2 - 12 \leqq 0, \\ &\qquad\qquad x_1 + 2x_2 - 12 \leqq 0, \quad x_i \geqq 0,\ i=1,2\} \end{aligned}\right\} \tag{18}$$

where $\tilde{c}_1$, $\tilde{c}_2$, and $\tilde{c}_3$ are fuzzy numbers whose membership functions are given below :

$$\left.\begin{aligned}\mu_{\tilde{c}_1}(c_1) &= \max(1 - 0.5\,|c_1 - 4|\,,\,0)\,,\\ \mu_{\tilde{c}_2}(c_2) &= \max(1 - 2\,|c_2 + 0.75|\,,\,0)\,,\\ \mu_{\tilde{c}_3}(c_3) &= \max(1 - |c_3 - 2.5|\,,\,0)\,.\end{aligned}\right\} \quad (19)$$

Now, for illustrative purposes, we shall assume that the hypothetical DM selects the initial value of the degree α to be 0.5, and the initial reference levels ($\bar{z}_1$, $\bar{z}_2$, $\bar{z}_3$) to be (9, -13, -3) . Then the values of ($c^L_{1\alpha}$, $c^L_{2\alpha}$, $c^L_{3\alpha}$) become (3, -1, 2), and the corresponding α-Pareto optimal solution can be obtained by solving the following linear programming problem.

$$\left.\begin{aligned}&\min_{x \in X} \quad v\\ &\text{subject to} \quad 2x_1 + 3x_2 - 11 \leqq v\\ &\qquad\qquad\quad -3x_1 - x_2 + 11 \leqq v\\ &\qquad\qquad\quad 2x_1 - x_2 + 1 \leqq v\end{aligned}\right\} \quad (20)$$

Solving this problem, we obtain an optimal solution (x_1^*, x_2^*, v^*) = (2, 3, 4), optimal values (z_1^*, z_2^*, z_3^*) = (13, -9, 1) , and the simplex multipliers corresponding to the constraints for the objective function of (20) become ($\Pi_1^*, \Pi_2^*, \Pi_3^*$) = (1/4, 2/5, 7/20). From (16) the trade-off rates among the objective functions become as follows :

$$\left.\frac{\partial(c_1 x)}{\partial(c_2 x)}\right|_{\alpha=0.5} = -\frac{\Pi_2^*}{\Pi_1^*} = -\frac{8}{5}$$

$$\left.\frac{\partial(c_1 x)}{\partial(c_3 x)}\right|_{\alpha=0.5} = -\frac{\Pi_3^*}{\Pi_1^*} = -\frac{7}{5}$$

Concerning the trade-off rate between $c_i x$ and α, from (17) we have

$$\left.\frac{\partial(c_i x)}{\partial\alpha}\right|_{\alpha=0.5} = (\Pi_1^*, \Pi_2^*, \Pi_3^*)(2x_2^*, 0.5x_2^*, x_1^*)^T = \frac{14}{5}$$

Observe that the DM can obtain his/her satisficing solution from among an α-Pareto optimal solution set by updating his/her reference levels and/or the degree α on the basis of the current values of the objective functions and α together with the trade-off rates among the values of the objective functions and the degree α.

6. CONCLUSION

In this paper, we have proposed an interactive decision making method for the multiobjective linear programming problem with fuzzy parameters characterized by fuzzy numbers on the basis of the linear programming method. Through the use of the concept of the α-level sets of the fuzzy numbers, a new solution concept called the α-Pareto optimality has been introduced. In our interactive scheme, the satisficing solution of the DM can be derived from among an α-Pareto optimal solution set by updating the reference levels and/or the degree α based on the current values of the objective functions and α together with the trade-off rates between the objective functions and the degree α. An illustrative numerical example clarified the various aspects of both the solution concept of α-Pareto optimality and the proposed method. However, further applications must be carried out in cooperation with a person actually involved in decision making.

REFERENCES

Bellman, R.E. and Zadeh, L.A. (1970). Decision making in a fuzzy environment. Management Science, 17 (4):141-164.

Dubois, D. and Prade, H. (1978). Operations on fuzzy numbers. Int. J. Systems Sci.. 9 (6):613-626.

Dubois, D. and Prade, H. (1980). Fuzzy Sets and Systems : Theory and Applications. Academic Press.

Fiacco, A.V. (1983). Introduction to Sensitivity and Stability Analysis in Nonlinear Programming. Academic Press.

Haimes, Y.Y. and Chankong, V. (1979). Kuhn-Tucker multipliers as trade-offs in multiobjective decision-making analysis. Automatica, 15 (1):59-72.

Orlovski, S.A. (1984). Multiobjective Programming Problems with Fuzzy Parameters. Control and Cybernetics, 13 (3):175-183.

Sakawa, M. (1983a). Interactive computer programs for fuzzy linear programming with multiple objectives. Int. J. Man-Machine Studies, 18 (5):489-503.

Sakawa, M. (1983b). Interactive fuzzy decision making for multiobjective linear programming and its application. Proc. IFAC Symposium on Fuzzy Information, Knowledge Representation and Decision Analysis, Marseille, France :295-300.

Sakawa, M. and Yumine, T. (1983). Interactive fuzzy decision-making for multiobjective linear fractional programming problems. Large Scale Systems, 5 (2):105-114.

Sakawa, M., Yumine, T. and Yano, H. (1984). An Interactive Fuzzy Satisficing Method for Multiobjective Nonlinear Programming Problems. CP-84-18. International Institute for Applied Systems Analysis, Laxenburg, Austria.

Tanaka, H. and Asai, K. (1981). A formulation of linear programming problems by fuzzy function. Systems and Control, 25 (6):351-357 (in Japanese).

Tanaka, H. and Asai, K. (1984). Fuzzy linear programming problems with fuzzy numbers. Fuzzy Sets and Systems, 13 (1):1-10.

Wierzbicki, A.P. (1979). The Use of Reference Objectives in Multiobjective Optimization - Theoretical Implications and Practical Experiences. WP-79-66. International Institute for Applied Systems Analysis, Laxenburg, Austria.

Yano, H. and Sakawa, M. (1985). Trade-off rates in the weighted Tchebycheff norm method. Trans. S.I.C.E., 21 (3):248-255 (in Japanese).

Zadeh, L.A. (1975). The concept of a linguistic variables and its application to approximate reasoning-1. Information Science, 8:199-249.

Multiobjective Decisions for Growth Processes of Time-Series

*M. Peschel, W. Mende and F. Breitenecker**

Academy of Sciences of the GDR and *Technical University of Vienna

Abstract

Based on the Lotka-Volterra approach for systems analysis and the qualitative behavior of so-called exponential chains, a special binary code is introduced for these chains, the exponential code. From the correspondence between exponential chains and the code a transformation rule between chains and codes is deduced which can be used to construct a huge number of dynamical chains with similar qualitative properties to the exponential chains.

These chains lead to possible local models for time-series description. The aim is to approximate time-series by a sequence of local models from a dynamical chain concept with a high data-reduction efficiency. This leads to a multicriteria problem, which consists of decreasing the number of local models with small complexity (order) to represent a given time series with a mean square error which is as small as possible.

Introduction

In [1] we introduced the Lotka-Volterra approach for systems analysis. A given set of ordinary differential equations is transformed into a corresponding set of Lotka-Volterra equations

$$Fx_i = \Sigma G_{ij} x_j + \Sigma H_{ij} y_j \quad \text{with} \quad F = \mathrm{d}\ \ln/\mathrm{d}t$$

by applying the so-called structure design principle.

The dominating cluster elements in this constructive process are the so-called exponential chains

$$Fx_i = K_i x_i x_{i+1} \qquad i = 0, 1, \ldots$$

In this paper we refer to the qualitative properties of the exponential chains following [1] and then we derive a corresponding exponential code. After comparing this code with the usual binary code we discuss the possibility of constructing rather general binary codes each based on two rather arbitrary monotonic functions. For every such code we construct a certain dynamical chain in analogy to the exponential chain. Because of this important relationship between codes and dynamical chains most of the qualitative properties we found for the exponential chains also hold for these general dynamical chains. Then we describe the possible use of these dynamical chains as local models for the description of time-series in a spline-like manner but with interesting trends of parameter variation considering consecutively such local models.

If we want to find a relevant description of a given time-series by a sequence of such local models with a minimum mean square error, we have to solve a corresponding multicriteria problem.

The corresponding objective functions of the multicriteria problem are:

- Q_1: number of reference points, or of local models (minimum)
- Q_2: the order of the local models, length of local dynamical chains (minimum)
- Q_3: number of structural changes between the applied local models (minimum)
- Q_4: the overall mean square error should be small
- Q_5: the reduction factor R = number of sample points/number of all parameters of the sequence of local models (maximum).

1. Qualitative Properties of Exponential Chains

Exponential chains are given by

$$Fx_i = K_i x_i x_{i+1} \qquad i = 0,1,2,...$$

Of especially high interest are finite order exponential chains with $i = 0,1,2,...,N$ and normalized initial values $x_i(0) = 1$ at least for $1 = 1,2,...,N$ with $x_{N+1}(t) = 1$. (These special exponential chains we have called *exponential towers* [1].)

For data analysis purposes we want to expand a given time-series $x_0(t) = x(t)$ into such an exponential tower with a low representation order. This task is analogous to a Taylor series expansion which itself is also a chain expansion, namely of the form

$$\mathrm{d}x_i / \mathrm{d}t = K_i + x_i \quad i = 0,1,... \text{ with } x_i(0) = 0 \text{ at least for}$$
$$i = 1,2,...$$

and in the finite case with $x_{N+1}(t) = 0$.

For the analysis of the qualitative properties of exponential towers we can make a restriction using the signs of the coefficients, that is we can assume $K_i = +1$ or $K_i = -1$.

For the corresponding signal $x_{0,N}(t)$ at the bottom of the tower with normalized initial values $x_i(0) = 1$ for all $i = 0,1,...,N$ we introduce the notation

$$x_{0,N}(t) = e(s_0, s_1, ..., s_N) \text{ with } s_i = \operatorname{sign} K_i \quad .$$

We call these functions coordinate functions of an exponential tower of order $N + 1$.

For the first and second order coordinate functions we get by integration directly:

$$e(s_0) = \exp(s_0 t) \qquad e(s_0, s_1) = \exp(s_0 / s_1(\exp(s_1 t) - 1)) \quad .$$

Unfortunately, higher order coordinate functions cannot be expressed in a closed analytical form.

If we take into account that $\exp(KI)$ with $I = \int_0^t dt$ is for $K > 0$ a positive monotonic and for $K < 0$ a positive antitonic operator, we get the following monotonic inclusions:

$$e(+) > 1 > e(-)$$

$$e(+,+) > e(+) > e(+,-) > 1 > e(-,-) > e(-) > e(-,+)$$

$$e(+,+,+) > e(+,+) > e(+,+,-) > e(+) > e(+,-,-) > e(+,-) > e(+,-,+) > 1 >$$

$$e(-,-,+) > e(-,-) > e(-,-,-) > e(-) > e(-,+,-) > e(-,+) > e(-,+,+)$$

This process of splitting up the lower order coordinate functions into coordinate functions of the next higher order can be transparently demonstrated.

We designate the corresponding linearly ordered set of binary vectors of the length l with $or(W_l)$ and the corresponding reverse order with $or(W_l)$.

Obviously, the generation rule for the construction of $or(W_{l+1})$ from $or(W_l)$ or from $or(W_l)$ is given by

$$or(W_{l+1}) = \begin{cases} +\, or(W_l) \\ -\, or(W_l) \end{cases}$$

Using this linear order for binary vectors $W_l = (s_0,s_1,\dots,s_{l-1})$ we introduce now so-called monotony classes for binary vectors of arbitrary but finite length, namely

$$> W_l = \begin{cases} W_l + or\ (W) & \text{if } \Pi s_i = +1 \\ W_l - or\ (W) & \text{if } \Pi s_i = -1 \end{cases}$$

and

$$< W_l = \begin{cases} W_l - or\ (W) & \text{if } \Pi s_i = +1 \\ W_l + or\ (W) & \text{if } \Pi s_i = -1 \end{cases}$$

Here, W is any finite binary word.

Adjoining upper and lower monotony classes are not overlapping, they are always separated from each other by a certain binary word of length $l - 1$.

According to the monotonicity classes for binary vectors just introduced we define now the so-called monotonicity classes for trajectories with regard to the coordinate functions $e(s_0,s_1,\dots,s_{l-1})$ by considering one parametric family of curves of the following type:

$$> e(s_0,s_1,\dots,s_{l-1}) = \begin{cases} e(s_0,s_1,\dots,s_{l-1},+K_l) & \text{if } \Pi s_i = +1 \\ e(s_0,s_1,\dots,s_{l-1},-K_l) & \text{if } \Pi s_i = -1 \end{cases}$$

$$< e(s_0,s_1,\dots,s_{l-1}) = \begin{cases} e(s_0,s_1,\dots,s_{l-1},-K_l) & \text{if } \Pi s_i = +1 \\ e(s_0,s_1,\dots,s_{l-1},+K_l) & \text{if } \Pi s_i = -1 \end{cases}$$

$> e(s_0,s_1,\dots,s_{l-1})$ is the upper and $< e(s_0,s_1,\dots,s_{l-1})$ the lower monotonicity class with reference to $e(s_0,s_1,\ldots,s_{l-1})$.

In these expressions $e(s_0,s_1,\ldots,s_{l-1}, s_l K_l)$ is the solution $x_{0l}(t)$ of a chain with the parameter set $s_0,s_1,\ldots,s_{l-1},s_l K_l$ with $K_l > 0$ and normalized initial conditions.

Obviously these classes built up bands of curves which are completely covered by the corresponding curve families parametrized by the parameter $K_l > 0$.

It can easily be seen that all these monotonicity classes together give a decomposition of the whole positive quadrant into disjoint bands. We will now study which phenomena occur if we adjoin another base module either at the top or at the bottom of a given exponential chain. It is convenient to express the corresponding results in terms of the language of finite automata.

For this purpose we consider the above-introduced monotonicity classes of trajectories as states of a finite automaton. If we adjoin another basic module at the top of the specified chain the state of the automaton will not be changed. But it is important to mention that every module adjoined at the top level will lead to a splitting of the given monotonicity classes each into two more narrow ones. That means a continuous division of curve-formed intervals (bands) into smaller ones thus leading necessarily to Weierstrass convergence of exponential chains with increasing chain length.

But if we in contrast adjoin another module at the bottom of the given chain, we get significant state transitions of the considered finite automaton, which means we pass from a specified monotonicity class (defined by the signs of the coefficients) to another one.

As can be seen easily the corresponding state transition occurs according to the following simple rule:

$$s \quad e(s_0,s_1,\ldots,s_{l-1}) = \begin{cases} > e(s,s_0,s_1,\ldots,s_{l-2}) & \text{for } \Pi s_t \cdot s = +1 \\ < e(s,s_0,s_1,\ldots,s_{l-2}) & \text{for } \Pi s_t \cdot s = -1 \end{cases} .$$

Here s symbolically denotes the operator corresponding to the chain extension by adjoining another basic module with the sign s at the bottom level of the given chain.

From this systematization of growth behavior of exponential towers together with the corresponding linear order of coordinate functions it can be immediately derived which solutions of exponential towers are saturation functions, that is, approach a finite value not equal to zero as t tends to infinity.

The phenomenon of saturation essentially only depends on the sign vector $s_0, s_1, \ldots$ of the chain coefficients.

The following sign combinations necessarily lead to saturation functions:

$(+,-)$ eventually followed by a finite number of pairs $(-,-)$, eventually followed by an arbitrary combination;

$(-,-)$ eventually followed by a finite number of pairs $(-,-)$, which then can be followed eventually by an arbitrary combination.

2. A Binary Code Corresponding to Exponential Towers

We are interested in coding the coordinate functions $e(s_0, s_1, \ldots, s_{l-1})$ in such a way that the order relation between the corresponding binary vectors $s_0, s_1, \ldots, s_{l-1}$ is reflected by a growing value of the adjoined code number.

This aim can be achieved quite easily if we introduce for a given a, $a > 1$, the following so-called Exponential Code

$$z = a^{s_0 a^{s_1 \cdots a^{s_{l-1}}}} .$$

This coding stimulates the study in greater detail of the corresponding code mapping determined by the following code generator function:

$$z = a^{sz'}$$

This function defined on the interval $[0,\infty)$ obviously maps for $s = +1$ the interval $[0,\infty)$ on $[1,\infty)$ and for $s = -1$ on $[1,0)$, presenting thus a dichotomy of the basic interval $[0,\infty)$ of the code.

If we iterate this code mapping a finite number of times, we produce apparently a dichotomy tree with smaller and smaller intervals each contracting to length zero as the number of iterations approaches infinity. This is obviously a static model for the Weierstrass convergence we observed for the trajectories of exponential towers with increasing length.

3. Dynamical Chains and Binary Codes

In the case of exponential towers we found the following correspondence between the basic module of such a chain written in integral form

$$x_i(t) = \exp(s_i \mathrm{abs}(K_i),\ Ix_{i+1}(t)) \text{ with } I = \int_0^t$$

and the code generator of the Exponential Code

$$z_i = \exp(s_i, z_{i+1}) \quad .$$

The most important result is that the interval contraction property of the coding procedure according to the dichotomy tree implies the Weierstrass convergence of the monotonicity classes of the exponential tower. This, of course, stimulates the idea of considering arbitrary binary codes and of constructing according to the rule derived above for them the corresponding dynamical chains.

An arbitrary binary code is uniquely determined by a code generator function

$$z = g(s, z')$$

obeying the following conditions:

- s is a binary variable with the values s_1 and s_2
- the functions $g_i = g(s_i, z')$ are monotonic functions both defined on the given basic interval J of the code mapping the definition interval J of the code on intervals J_i, $i = 1{,}2$ defining a partition of J, namely

  $$J_1 \quad J_2 = J \text{ and } J_1 \quad J_2 = 0$$

  (with the possible exception maybe of single points at the boundaries of these subintervals).

Depending on the character of the monotonic branches of the code generator function (increasing or decreasing), we get four different types of corresponding partition with consequences for the linear order of the corresponding binary vectors. The concrete analytical form of the branch functions can be freely chosen. This opens huge possibilities for constructing dynamical chains with similar properties.

If we use the correspondence between the code generator and the basic module of the dynamic chain which we found for the case of the Exponential Towers we get the following dynamical chains

$$x_i(t) = g(s_i, 0(Ix_{i+1})) \quad .$$

Here in general a nonlinear (monotonic) mapping 0 is necessary which maps the interval in which the integral Ix_{i+1} varies into the definition interval J of the code to secure always the applicability of the code mapping. This is on the other hand a suitable advantage to start the construction with a normalized code working for example on the interval [0,1] and to introduce free parameters from the degrees of freedom for choosing the nonlinear mapping 0.

Let us give now some simple examples for this construction.

1. For $g^{-1} = \text{abs}(2z-1)^k$ we get

$$\mathrm{d}x_i/\mathrm{d}t = \frac{s_i}{2k_i K_i^2} x_{i+1}/\text{abs}(x_i^2 - K_i^2)^{(k_i-1)} \quad .$$

2. For $g^{-1} = Kz^k(1-z)^l$ with $K = (k+1)^{(k+1)}/(k^k l^l)$ we get

$$\mathrm{d}x_i/\mathrm{d}t = x_{i+1}/(K^{i^{l_i+1}}(k_i + l_i)^{(k_i+l_i)} k_i^{k_i} l_i^{l_i}) A(x_i)$$

with

$$A(x_i) = x_i^{k_i-1}(K_i + x_i)^{k_i+l_i-1}(k_i K_i - l_i x_i)$$

3. For $g^{-1} = \text{abs}(\ln(z/L))^k/K$ we get

$$\mathrm{d}x_i/\mathrm{d}t = K_i s_i x_i x_{i+1}/(k_i \text{abs}(\ln(x_i/L_i))k_{i-1}$$

which obviously is a generalization of the exponential tower which is the special case with all $k_i = 1$.

4. Time-Series Approximation by Local Models by Solving a Multicriteria Problem

Applying as local models for time-series approximation a sequence of dynamical chains constructed after a certain binary code, we meet the following decision problem which we have to solve in stages. We consider a time-series $x(t)$ on the time-interval $(0,T)$. We have to choose a finite number $Q_1 = N$ of reference points t_i, $i = 1,2,\ldots,N$, in which we are going to construct local models after a certain system concept.

Naturally, we use the system concept for the local models of the above-considered dynamical chains. We apply these models with equal complexity (number of model parameters); the complexity of local models Q_2 is a second objective which should be minimized.

For the purpose of the adaptation of its parameters every local model makes use of measured values from the time-series in a certain neighborhood of the reference point. This information-supplying neighborhood is the third objective Q_3; it should be small, but it is strongly connected with the value of Q_2.

The extrapolation force Q_4 of every local model should be maximal, that is, we want to consider the local model as a substitute for the time-series in an interval around the reference point which should be large. The extrapolation intervals of adjoining local models are admitted for overlapping.

The representation error (mean square error) of every local model in comparison with the given time-series should be minimal. As objective Q_5 we take the overall representation error of all applied local models together as a global objective for modelling the time-series.

Another rather important objective is the data reduction factor Q_6. Let Q_0 be the total number of samples of the time-series. Then the data reduction factor can be defined in the following way:

$$Q_6 = Q_0 / (Q_1 \cdot Q_2) \quad .$$

To use rationally the storage and data communication channels we are interested in as large a value of Q_6 as possible. The main contradiction in this multicriteria task exists between Q_5 (representation error, as small as possible) and Q_6 (data reduction factor, as large as possible).

For the following reasons it seems suitable to apply as local models the above-introduced dynamical chains.

Let us, for a better understanding, perform a mental experment and assume that we adapt local models at all sample points of the considered time-series and that the sample points are taken with a very small tact from the original time-series, which might be even a continuous time-trajectory.

Then, under these conditions, adjoining local models should be in their parameter sets very similar to each other, because the continuous variation of the time-series from point to point should imply also comparable small variations of the parameters of adjoining models. But there are distributed on the (continuous) time-series singular points which necessarily lead to structural changes of the local models, namely, to changes of the corresponding binary vector $s_0, s_1, \ldots, s_{l-1}$. This is in the example of exponential chains for local models the case at all points of the time-series, where a logarithmic derivative taken into account in the local models is changing its sign. Between these singular points (singular in the sense of structural change) the adjoining models show a quite regular behavior caused by the similar properties to binary codes. We can observe the following phenomena under the condition that the parameter fitting process is done with high accuracy.

Passing along the time-series from reference point to reference point quantitative changes of parameters should first inflect the basic models at the top of each chain. After some steps these changes should penetrate deeper and deeper to lower levels of the chains. This tendency continues until we meet the next singular point. These systematic changes in the parameters give us further possibilities for data reduction.

We need not communicate or put into storage the complete parameter sets of all local models, but it is enough to do this with the increments of these parameters if they are larger than a certain threshold. Usually, we need not do this for all parameters but only for some of them, beginning with the highest model level.

In general - see [2] - the efficient solutions of this multicriteria problem also depend on the positions of the reference points. Therefore, we should look for those positions of the Q_1 reference points for which the vector optimization task with the objectives $Q_1, Q_2, \ldots, Q_6$ offers us the most appropriate efficient solution.

References

[1] Peschel, M. and W. Mende: Predator-Prey Models - Do we live in a Volterra World? Akademie-Verlag, Berlin and Springer-Verlag, Berlin-Heidelberg-New York-Wien, 1986.

[2] Peschel M.: Ingenieurtechnische Entscheidungen - Modellbildung und Steuerung mit Hilfe der Polyoptimierung. VEB Verlag Technik, Berlin, 1980.

INTERACTIVE DECISION ANALYSIS AND SIMULATION OF DYNAMIC SYSTEMS

F.Breitenecker
Technical University Vienna
Inst.for Technical Mathematics
Wiedner Hauptstrasse 8-10
A-1040 Vienna

1. INTRODUCTION

The rapid development of computers, especially microcomputers greatly increases the scope and capabilities of computerized decision support and decision analysis systems for various processes and for various applications. Usually these systems are based on (mathematical) models of the process.
In the last few years most of the mathematical modellers have come to the conclusion that mathematical models for decision support and decision analysis must be built up and used interactively (see e.g.Wierzbicki(1984), Peschel and Breitenecker (1984)). That means, that the user has to be involved interactively at all levels of the process (shown in fig.1).
This fact requires advanced information processing systems (AIP systems) for decision support and decision analysis. In case of dynamic systems there exist tools for simulation, which can be seen as analysis of the system in the time domain. But these simulation languages (SLs) have only few features for further analysis and decision support (based on the simulation before).
In the following therefore it is shown, how interactive analysis, decision support and decision analysis can be implemented within SLs using a modern methodological concept for SLs. Three applications are given, a hydro-energetic system, population dynamics and physiological processes.

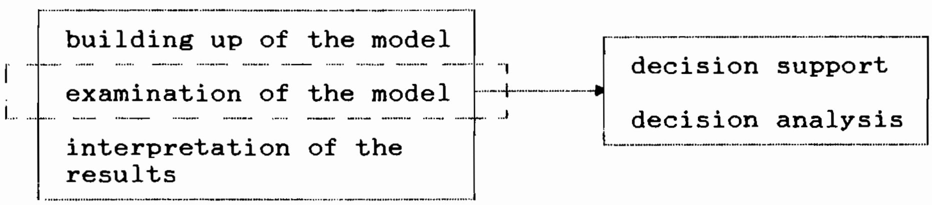


Fig.1: Levels of analysis and decision support

2. FEATURES OF SIMULATION LANGUAGES

The development of SLs (and simulation packages) can be seen as a first step of AIP for decision analysis for a certain class of processes, as mentioned before. Using a SL the user can perform parts of the examination of the model (= simulation of the model in the time domain) in a comfortable and interactive manner, so that SLs are able to perform a part of the interactive decision analysis (fig.1, dashed line). But the building up of a model (model change) and the interpretation of the results has to be done 'by hand'. Furthermore, usually advanced features for model examination are missing.
To sum up, SLs seem to be an unsatisfying tool for implementation of interactive decision analysis and decision support in models for dynamic processes; software for such goals seems to require another basis.
This conclusion holds only in parts. Some SLs can be extended up to a decision support system, if the SL is an 'open' one, that means, additional features can be added to the language. Adding an additional feature extends the SL with an additional algorithm for examining the model of the dynamic process. Usually SLs offer only basic algorithms for examining a model, which are the simulation run (integration of the system governing differential equations), parameter variation and documentation (plots, etc.).
An examination of the model, in the following called an 'experiment' with the model,is for instance a simulation run, a phase plot, changing a time constant.
Following a new methodology for future SLs proposed for instance in Breitenecker (1983b) and Breitenecker, Solar(1986) a concept of a SL should be based on the separation of the levels 'model' (MO), 'method' (ME) and 'experiment' (E). The so-called methods are to be seen as procedures for examination of a model, an experiment is defined then as application of a certain method on a certain model (E = ME(MO)). In today's SLs usually no distinction is made between method and experiment because in an interactive 'simulation session' only one model is available.
This new methodology introduces now the level of methods, which can be defined,linked and changed independently from the actual model. This concept allows to extend the language with very complex features, for instance with a method for a certain stability analysis or for a statistical analysis of data which were sampled in a simulation run. Furthermore, this new concept allows to use more than one model within a simulation session.
Figure 2 summarizes the today's and future structure of a SL.
Due to the standard for SLs (CSSL standard 1968) at each of the two levels model (MO) and experiment (E) macro features have to be available. In the new methodological concept this feature is required also at the level of methods (ME). These macro features are a very powerful tool in SL. It should be noted that method macros can be 'simulated' in today's simulation languages by tricky programming of experiment macros.
For the goal of interactive decision analysis within a SL the method 'model generating' is very important. Using this feature would allow to generate a new model (using the model data base of the actual model) for certain further examinations.

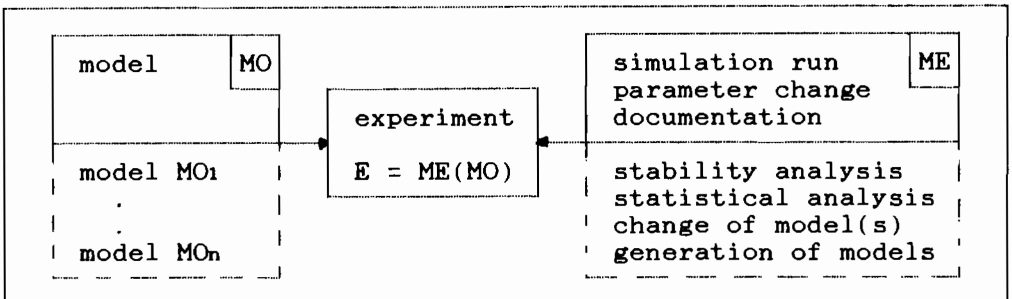


Fig.2: Features of today's SLs (full lines) and future SL (dashed lines)

3. THE SUPERMACRO - A SPECIAL EXPERIMENT MACRO

The interactive decision support and decision analysis in dynamic systems using SLs requires the following tasks:

* interactive generation and change of models
* simulation runs, parameter variations
* generation of special model-dependent procedures (or methods) for investigating the system
* interpretation of the results
* decision support based on interpretation of results
* 'automatic' decisions
* recursive performance of the tasks

Following Breitenecker (1983) a special experiment macro, a so-called supermacro, can be used in today's (and also future) SLs for 'simulating' and/or implementing a 'procedure' (a complex method) which allows to perform the above mentioned tasks.
This supermacro technique proposes to program an experiment macro which is able to perform the following tasks:

* automatic definition of the model equations and models MO_i (an experiment macro E_{model} invokes the method ME_{model} generating the desired models MO_i or changing the preconfigurated models MO_i)
* specification of special purpose methos ME_j for simulating and analyzing the models MO_i (an experimental macro E_{method} performs the method ME_{method} generating modules or combining preconfigurated modules resulting in methods ME_j for analyzing and handling the models MO_i)
* Initialization and/or invokation of experiments E_{ij} ($=ME_j(MO_i)$) performing the generated methods ME_j with the generated models MO_i
* Help function and entry points for/to each submodule of the supermacro (in order to restart the macro with any arbitrary subexperiment)

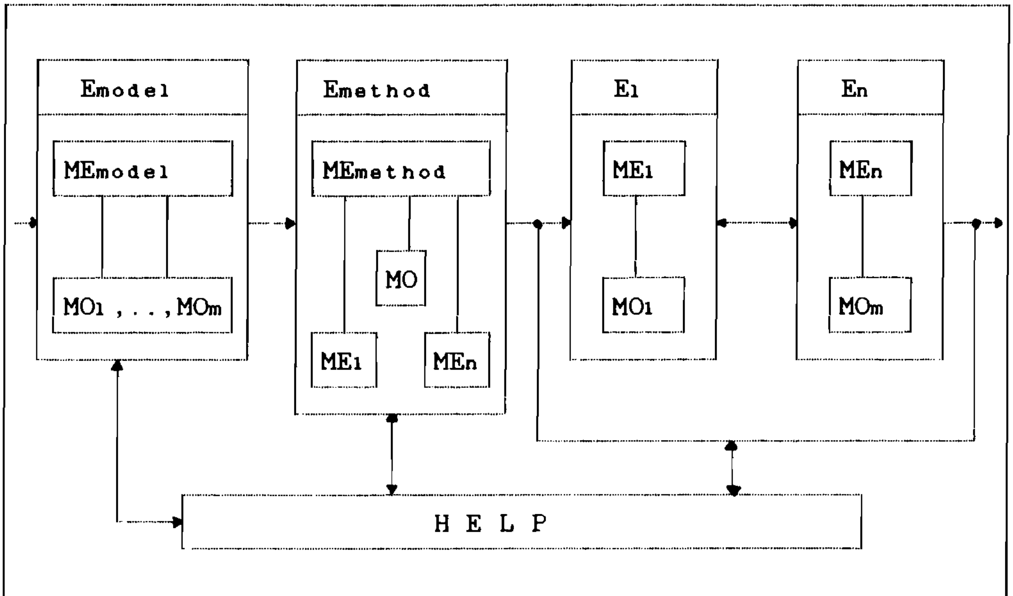


Fig.3: Structure of a 'supermacro'

The outlined concept of 'supermacros' can be implemented in today's simulation languages only in parts. The difficulties are the handling of method specification, separate compiling and recursive definitions.
The implementation in interpreter-oriented SLs can be done using the features of the kernel structure. There subroutines in a host language (usually FORTRAN) can be added to the kernel, which perform for instance the generation of new models (because the model data base is available also on the host language level). Implementations using this 'supermacro'-technique are given in Breitenecker (1983b) where the SL HYBSYS, a hybrid interpreter-oriented language developed at the Technical University Vienna (see Solar et al.(1982) is used.
In compiler-oriented SLs (there the description of the model is compiled into an 'object' (model) instead of generating a model data base) difficulties arise especially in realizing experiments which should generate models or which should change models. There either preprogrammed models can be loaded (but these have usual 'maximal' dimensions) or a sort of a precompiler can be used which generates a source program to be used as input for the model definition part of the SL. The experiments generating or loading methods can be realized by extending the main program generated by the SL: this main program usually consists of a loop which passes the program executing the runtime commands and the program executing a simulation run; within this main program (usually a FORTRAN main program, as in case of ACSL) now anything can be added, for instance optimization programs, statistic data analysis, etc.

4. DECISION ANALYSIS AND SUPPORT AS 'SUPERMACRO'

Using the fore-mentioned supermacro- technique the modules necessary for interactive decision analysis and decision support based on models for dynamic processes can be implemented in today's and future simulation languages.

The advanced information processing system for decision support (based on simulation) has to consist of the following tasks:

* <u>experiment module:</u> an experiment module or help function (E_{exp}) guides the user interactively through all parts of the AIP system for decision analysis and decision support

* <u>input module:</u> an input module (experiment E_{model}) generates interactively (in dialogue with the user) the model or/and the model data base due to the process to be investigated for certain decisions

* <u>method module:</u> a method module (experiment E_{method}) loads and/or generates appropriate methods for analyzing the model and interpreting the results, for decision support based on simulation and/or other analysis (stability analysis,..) giving the experiments E_{res}, E_{dec}

* <u>result module:</u> a result module (experiment E_{res}) activates and performs methods for analysing and interpreting the results for instance of a simulation, of a stability analysis, of a Monte-Carlo study,..

* <u>decision module:</u> a decision module (experiment E_{dec}) activates and performs methods and algorithms for giving interactively proposals and hints for certain decision corresponding to the model and tries to answer certain specific questions using the results of simulation (E_{stan}), of result interpretation (E_{res}),.....

* <u>standard module:</u> the standard module (experiment E_{stan}) offers the standard features of the simulation languages which are necessary for the problem to be investigated; this standard module E_{stan} can either be only a help function guiding the user through the necessary standard runtime commands of the SL or the module is a 'standlone module' which offers interactively only the commands necessary for the analysis of the problem in an interactive way so that the user need not know anything about the syntax of the SL

Figure 4 summarizes the structure of this advanced information processing system for decsion support and decision analysis based on analysis and simulation of the process.
As outlined in section 4 this AIP-system based on the supermacro-concept can be realized and implemented in the today's SLs only more or less completely.
In interpreter-oriented SLs this structure can be 'simulated' in using the fact that the kernel of the language can be extended by subroutines (written in a host language) resulting in 'additional' commands (which may call other commands in a

recursive way). The fore-mentioned experiments are then loaded and linked together with the standard features of the SL which itself often are experiments built up with macro technique from basic experiments.
In compiler-oriented SLs the model is a precompiled object (instead of a model data base). Consequently difficulties arise in 'adding' additional programs (performing the decision support). In this case the main program of the SL (which is generated by the precompiler of the SL and is usually a loop passing the program executing the runtime processor and the program executing a simulation run) has to be extended by additional optional programs. A further difficulty is the fact, that only the names and values of the model variables are available, but no information about the model structure (in case of a model data base information about the model structure is available in a comfortable way).

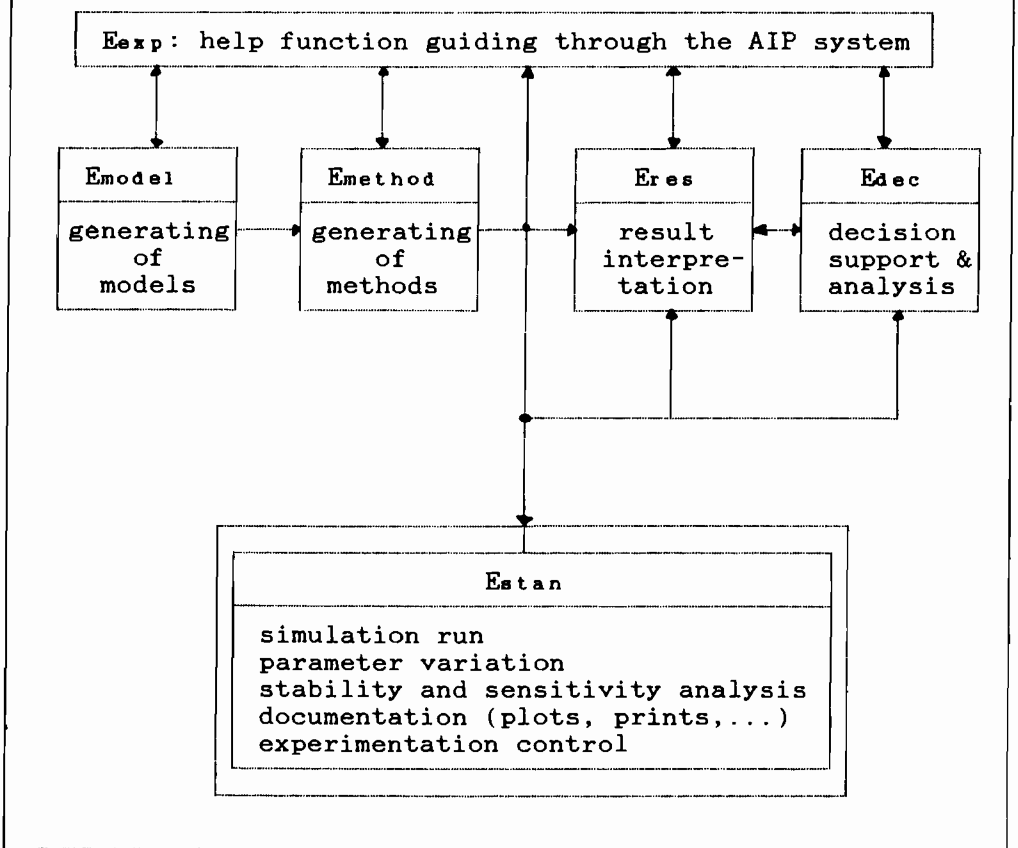


Figure 4: Structure of an AIP system for decision support and decision analysis in a simulation language

5. APPLICATIONS

In this section three applications of the proposed AIP system for interactive decision support and decision analysis based on simulation of dynamic models are discussed.
In the first application, a hydro-energetic system, the user wants informations and decision support about necessary outflows from reservoirs in order to maximize the costs of the system. The second application, population dynamics, tries to find interactively suitable models for a certain population dynamic and to give proposals for suitable external actions in order to change for instance the growth behaviour. In the third application models for physiological processes are considered; there identified model parameters (subject to measured time series) and physiological parameters are used in order to get decisions about diagnosis and therapy.

5.1 Hydro-energetic system

Basis of the AIP system is a continuous model of a hydro-energetic multipurpose system consisting of reservoirs, power plants, rivers, water-users and pumping stations.
With few simplifying assumptions this system is described by state (bilance) equations for the reservoirs (obtained by the continuity principle; see Breitenecker, Schmid (1984)):

$$\dot{s}_i(t) = z_i(t) - a_i(t) - d_i s_i(t), \quad i=1,..,n$$

with given initial values , time horizon and bounds. There s_i, z_i and a_i denote storage, inflow and outflow of the i-th reservoir, the term $d_i s_i$ 'measures' the evaporation. The inflow consists of natural inflows, inflows from upstream reservoirs, iflows from pumping and inflow from upstream users. The total outflow consists of outflow to other reservoirs, outflows to users and outflows to pumps. The outflows to other reservoirs a the controls of the system.
Now the AIP system tries to analyze the dynamic behaviour of this hydro-energetic system in order to maximize or minimize costs such as benefit of energy production, flood control, recreational use, pumping costs and withdrawal of water.

The AIP system for this purposes consists of six tasks:

* 'EQUAT': automatic definition of model equations according to the users input (number of reservoirs, pumps, water users and the flow of water between them) (= E_{model})
* 'MEINIT': generating and initializing modules for analyzing the system (optimization prozedures,..) (= E_{method})
* 'OPTI' optimizes costs subjects to interesting parameters and parametrized controls (= E_{res})
* 'VOPTI' performs vector (poly-) optimization of different contrary cost (energy production, recreational use) subject to interesting parameters and parametrized controls (= E_{res})
* 'COCOST' compares graphically different costs (optimized and not optimized) in order to give hints for possible control actions (= E_{dec})
* 'CHCON' chooses controls in interactive dialogue with the user in order to get 'better costs' (= E_{dec})

This AIP system (summarized in fig.5) was implemented using the fore-mentioned hybrid simulation language HYBSYS, where the modules are supported by standard features of HYBSYS (E_{stan}) and a Help function guides the user through the system (E_{exp}).

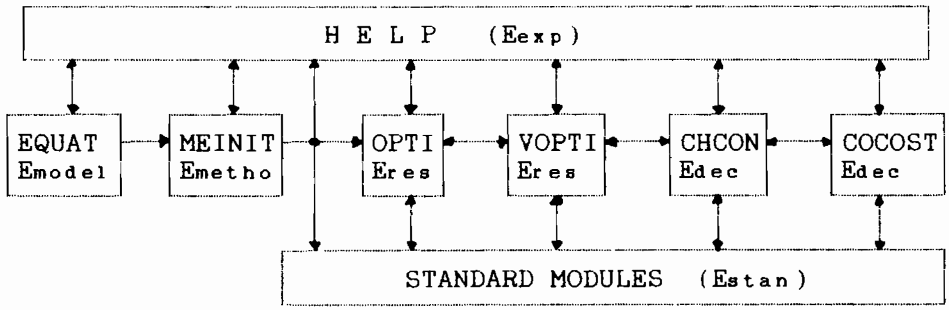


Fig.5: AIP system for a hydro-energetic system

5.2 Population dynamics

This AIP system for decision analysis and decision support is based on the Volterra equation describing the behaviour of population growth in the time domain:

$$\dot{x}_i(t) = \sum_j ((g_{ij}x_j + h_{ij}f_j)\, x_i + k_{ij}u_j)$$

There $x_i(t)$ denotes the i-th population, the variables f_j are external influences and u_j are control variables; the coefficients g_{ij}, h_{ij} and k_{ij} are wheighting factors.

The AIP system was (partly) implemented in HYBSYS and performs the following actions:

* 'EQUAT': interactive generation of the model in dialogue with the user (number of populations, controls,..) (=E_{model})
* 'MEINIT': generation of methods for analyzing the system (generation of linearized models for further use, generation of sensivity equations, ...) (E_{method})
* 'STAPOI': calculation of stationary points (equilibrium points) (E_{res})
* 'LOSTAB': local stability analysis using methods depending on the dimension of the model (using well known results for instance for dimensions 2 and 3) (E_{res})
* 'GLSTAB': global stability and sensitivity analysis using nonlinear methods and iterated simulation runs (E_{res})
* 'GRSTAB': stability analysis for a set of different parameters using graph-theoretical methods giving for instance hints for wheighting factors for controls (E_{res}, E_{dec})

* 'STSTAB': analysis of the structural stability of the system (giving hints for chaotic behaviour) (E_{res}, E_{dec})
* 'MUTUA': check, whether on a certain level (in a certain equilibrium point) mutuations may occur or not (depending on parameters (E_{res}, E_{dec})
* 'CHCON': choice of controls in order to change equilibrium points or resulting in exploding growth or resulting in dying out (E_{dec})
* 'SIMSTO': stochastic simulation in the time domain with appropriate documentation (stochastic disturbances)

The modules of this AIP system are supported by the standard features of the simulation language. Furthermore a module is planned, which transforms any other description of growth (using differential equations, for instance hyperlogistic growth laws) into the fore-mentioned Volterra equations (see Peschel, Breitenecker (1984); Peschel, Mende (1986)).

5.3 Physiological processes

This interesting AIP system for decsion support and decision analyis in physiological processes is based on the so-called 'controlled compartment' - approach for modeling and simulating physiological processes (Breitenecker (1985)).
Considering the well known modeling approaches 'concept fitting using control theory' and 'compartmental modeling' both have natural advantages. Compartmental modeling reflects the structure of the process very well, control theory works with the perhaps most natural principle, with the feedback principle. But both techniques have disadvantages, compartmental modeling results in an amount of model parameters, model building based on control theory (using for instance input- output relations) establishes very often variables and parameters without physiological meaning. The 'controlled compartment'- approach tries to combine the advantages of both approaches: the system is decomposed into as few subsystems (generalized compartments) as possible, each subsystems is described by methods of control theory (input/output behaviour).
This approach is very successful in case of 'active' systems, that means, the behaviour of the system under certain excitations (inputs) is to be analyzed. One advantage furthermore is, that only few model parameters are to be identified corresponding to measured individual time series (of excitations and reactions); due to the nature of the modeling procedure these model parameters are time constants, gain factors, etc.). Model building is performed in the AIP system depending on the available time series (E_{model}).
After identification of the model parameters (E_{res}) the behaviour of the system excitated by other inputs can be studied for individual persons. This prediction of physiological behaviour allows to replace dangerous and expensive labour experiments and gives hints for diagnosis and therapy (E_{dec}).
Furthermore correlations between model parameters and physiological parameters (age, wheight, resting blood pressure, fitness, etc.) derived by statistical methods (parametric and nonparametric statistics, cluster analysis) allows to characterize the physiological behaviour by (static) individual parameters giving information about kind of sickness, degree of sickness, etc. (E_{dec}).

This method of analyzing physiological systems has prooven very successful in the following systems: heart rate and blood pressure during ergometric load (Breitenecker (1985)), glucose production of livers during stress hormones in vitro (Troch et al. (1985)), blood glucose concentration during stress hormones in men suffering on diabetes and in sound men (Komjati et al.(1985)), pathological blood pressure behaviour after load in men with coarctation of the aorta (Breitenecker,Kaliman(1984)). AIP systems for these processes were implemted in HYBSYS using again the structure E_{model}, E_{method}, E_{res}, E_{dec}, E_{stan} (standard simulation features) and E_{exp} (help function).

REFERENCES

Breitenecker F.(1983a). Simulation package for predator-prey systems. Proc.Int.Conf.'Simulation of Systems '83', Prague, 3/1 3:213.

Breitenecker F.(1983b).The concept of supermacros in today's and future simulation languages. Mathematics and Computers in simulation, XXV:279-289.

Breitenecker F., Schmid A.(1984). Decision support via simulation for a multipurpose hydroenergetic system. In M.Grauer, A.Wierzbicki (Eds.), Interactive Decision Analysis, Lecture Notes in Economics and Mathematical Systems,Springer, 227-235

Breitenecker F., Kaliman J.(1984). Results in simulating pathological blood pressure behaviour after treadmill test in patient with coarctation of aorta. In A.Javor (Ed.),Simulation in Research and Development, North Holland, 173-185.

Breitenecker F.(1985). Modeling of physiological processes by the 'controlled-compartment'- approach. In A.Sydow, R.Vichnevetsky (Eds.), Systems Analysis and Simulation 1985, Mathematical Research vol.28/II, Akademie-Verlag, Berlin, 434-438.

Breitenecker F., Solar D. (1986). Models, methods, experiments - modern aspects of simulation languages. Proc. 2nd European Simulation Congress, Antwerp, Sept.1986. To appear.

Komjati M., et al.(1985). Contribution by the glycogen pool and adenosine 3',5'-monophosphate and cyclip AMP release to the evanescent effect of glucagon on hepatic glucose production in vitro. Endocrinology vol.116, no.3, 978-986.

Peschel M., Breitenecker F., Mende W.(1983). On a new concept for the simulation of dynamic systems. Informatik-Fachbericht 71, Springer, 93-98.

Peschel M., Breitenecker F.(1984). Interactive structure design and simulation of nonlinear systems under multiobjective aspects with the Lotka-Volterra approach. In M.Grauer, A.Wierzbicky (Eds.), Interactive System Analysis. Lecture Notes in Economics and Mathematical Systems vol.229, Springer, 63-76.

Peschel M., Mende W.(1986). The Predator-Prey Model: Do We Live in a Volterra World ? Akademie-Verlag, Berlin.

Solar D., Berger F., Blauensteiner A.(1982). HYBSYS - an interactive simulation software for a hybrid multiple-user system. Informatik-Fachbericht 56, Springer, 257-265.

Troch I., et al.(1985). Modeling by simulation of hepatic glucose production in vitro. In A.Javor (Ed.), Simulation in Research and Development, North Holland, 163-171.

Wierzbicki A.(1984). Interactive Decision Analysis and Interpretative computer intelligence. In M.Grauer, A.Wierzbicki (Eds.), Interactive Decision Analysis. Lecture Notes in Economics and Mathematical Systems vol.229, Springer, 2-20.

PATTERN FORMATION BY NATURAL SELECTION WITH APPLICATIONS TO DECISION SUPPORT

Hans-Michael VOIGT
Academy of Sciences of the G.D.R.
Mathematics and Informatics Division
DDR-1199 Berlin-Adlershof, Rudower Chaussee 5

1. INTRODUCTION

One of the most impressive principles of selforganization of complex systems is given by the principle of natural selection as it was presented by Charles Darwin by his fundamental work "On the origin of species by means of natural selection" and its extensions.

Simplified this principle explains the evolution of a system to higher stages with respect to its function and structure by mutation and selection.

The question arises how to use this principle for the structural design of complex technical systems effectively.

For the layout design of printed circuit boards and integrated circuits we want to try to answer this question expecially for the partitioning of large graphs.

2. PARTITIONING OF GRAPHS

The partitioning of graphs and hypergraphs, resp., is one of the problems which has to be solved in the design of printed circuit boards and integrated circuits. This problem gets more and more difficult to solve because of the increasing complexity of the integrated circuits to be designed. Therefore combinatorial methods are prohibitive. As a suitabele method to overcome these difficulties the use of the principle of natural selection seems to be an effective heuristics.

The problem of graph partitioning may be stated as follows: Given is an undirected graph $G = (V,E)$ with the edge set E and the vertex set V. The edges have the weights $G_{ij} = G_{ji}$, $G_{ij} \geq 0$. The vertices should be assigned to partitions of the partition set P. With x_{ik} the membership of vertex V_i to the partition P_k is denoted. Then the partitionung matrix C is given by

$$C = (X,GX)$$

with $X = x_{ik}$ the matrix of membership values.
Here C_{kk} is the fuzzy number of weighted edges within the partition P_k and C_{kl}, $k \neq l$, the fuzzy number of weighted edges between the partitions P_k and P_l.

The partitioning problem of graphs can then be formulated as an vector optimization problem with

$$C_{kl} \dashrightarrow \max \quad \text{for } k = l$$

or equivalently

$$C_{kl} \dashrightarrow \min \quad \text{for } k \neq l.$$

From this a number of special cases based on the used utility function may be derived where the following are especially interesting for the above given paritioning problem

$$\sum_k l_k \, C_{kk} \dashrightarrow \max \qquad (+)$$

$$\prod_k C_{kk}^{l_k} \dashrightarrow \max \qquad (++).$$

Eq. (++) means for $l_k = 1$, that the C_{kk} should maximized and should have at the same time approximately equal size. This case occurs in the design of integrated circuits based on standard cells and gate arrays.

3. REPLICATION

To speak of mutation and selection in a general sense we use the notion of "replicators" as it was introduced by Dawkins.

By "replicators" we mean any units which may be copied and (i) which imply by their nature a certain influence on the probability or possibilty to be copied

(ii) which may be, at least principlallv ancestors of an indefinitely long series of copies.

Replicators in this sense may be alleles in a gene pool, relative numbers within animal populations, selfreproducing macromolecules, strategies of individuals etc.

The replication of these units will be modeled by the following general replicator equation

$$\frac{\dot{x}_i}{x_i} = f_i(x_1,\ldots,x_n) - \emptyset \qquad (x)$$

or

$$\frac{x_i'}{x_i} = \frac{f_i\,(x_1,\dots,x_n)}{\emptyset} \qquad (**)$$

with $\sum x_i = 1$.

In terms of genetics x_i is the proportion of allel i within the gene pool, f_i is the mean fitness of allel i within the gene pool, and $\emptyset$ is the mean fitness of the total gene pool. If we include the possibility of mutations by determined mutation rates this yields a model for natural selection processes.

Obviously stable structures of the evolving system are given by stable attractors - in the most simple case by stable fixed points $\dot{x}_i = 0$ and $\dot{x}_i = x_i$, resp. - which represent a stationary probability or possibility destribution x_i of allel i within the gene pool. This dynamic evolution of stationary structures is driven by competition and cooperation between the alleles.

3.1. LINEAR FITNESS FUNCTIONS

Of special interest are linear homogenous fitness functions f_i. In this case we have

$$f_i(x_1,\dots,x_n) = \sum_j G_{ij}\, x_j = (Gx)_i$$

and

$$\emptyset = \sum_i x_i \sum_j G_{ij}\, x_j = \sum_i x_i\, (Gx)_i = (x,Gx) \;.$$

The G_{ij} are interaction coefficients between alleles i and alleles j. The replicator equation with linear fitness functions has the foolowing properties, which are important for the further analysis:

(i) Eq. (*) is invariant under the transformation $G_{ij} := G_{ij} + a_j$

(ii) Eq. (**) is invariant with respect to the fixed points under the transformation $G_{ij} := G_{ij} + a_j$

(iii) For G real, symmetric and $G_{ij} \geqslant 0$ the following relations for the mean fitness of the total gene pool hold $\emptyset \geqslant 0$ (*) and $\emptyset' \geqslant \emptyset$ (**) .

(iv) For G real, symmetric and $G_{ij} \geqslant 0$ the eigenvalues of eq. (*) are real, i.e. no limit cycles or other dynamic attractors occur.

3.2. REPLICATION RATE CONTROL

For the control of the structure formation process we introduce variable selfreplication rates $G_{ii} = u_i$.

By means of these replication rates a scalarization of the global state of structuring is possible. $u_i \dashrightarrow -\infty$ yields a uniform cooperation between the replicators and $u_i \dashrightarrow \infty$ a uniform competition. Between these two extremal cases further states of structuring occur. Especially interesting are homogenous selfreplication rates $u_i = u$.

4. DIVERSIFIED REPLICATION

Related to the previous sections we habe for every partition P_k a replicator equation of the form

$$\frac{\dot{x}_{ik}}{x_{ik}} = \frac{(Gx_{\cdot k})_i}{C_{kk}} - e_i \qquad (!)$$

or

$$\frac{x'_{ik}}{x_{ik}} = \frac{(Gx_{\cdot k})_i}{C_{kk}e_i} \qquad (!!)$$

with $\sum_k x_{ik} = 1$.

Corresponding to the properties of the replicator equation (*) and (**) eqs. (!) and (!!) have the following properties:

(i) Eqs. (!) and (!!) are invariant with respect to the fixed points under the transformation $G_{ij} := G_{ij} + a_j$ This property will be used for negative replication rate control $u < 0$ such that $G_{ij} \geqslant 0$.

(ii) In analogy to property (iv) of the replicator equation (*) for eq. (!) no limit cycles or other dynamic attractors occur. This property ensures a convergence of the diversified replicator equation.

(iii) For G real, symmetric and $G_{ij} \geqslant 0$ corresponding to property (iii) of the replicator equation (*) for the diversified replicator equation (!) the follo-

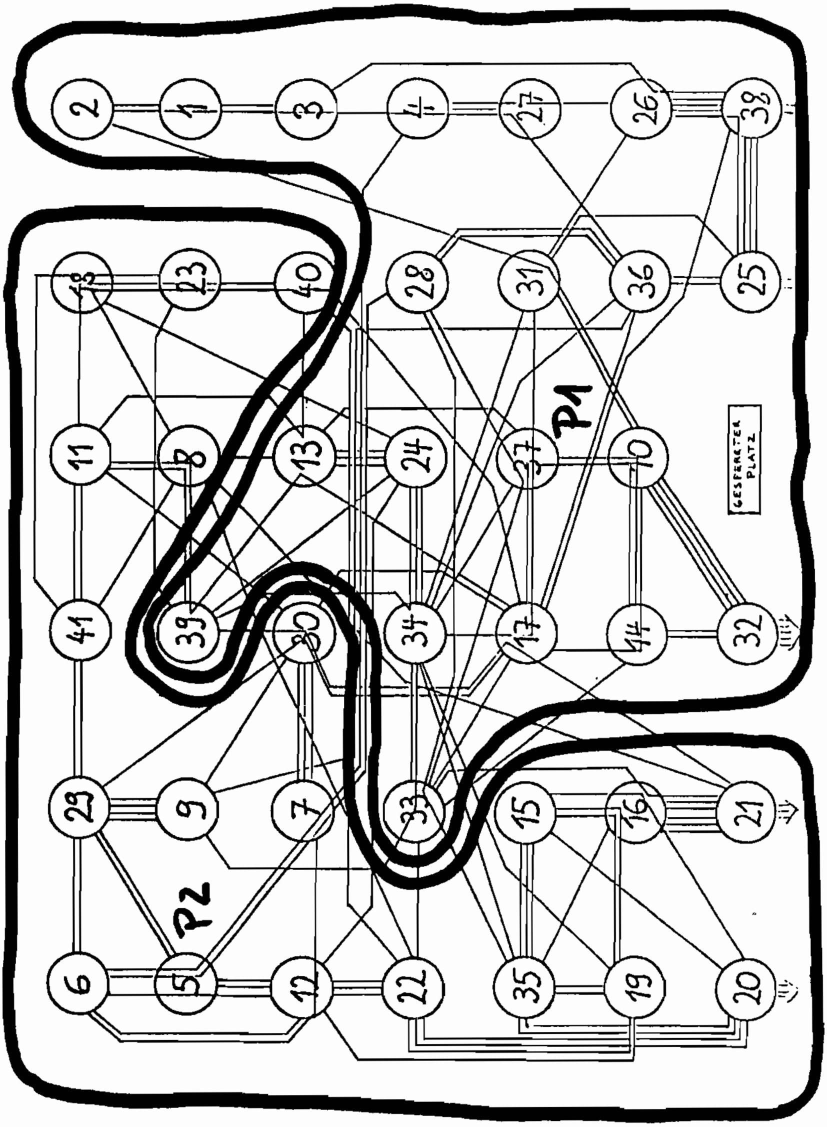


FIGURE 1 B41A, PNS, Partition sizes 20/21, 20 total cuts

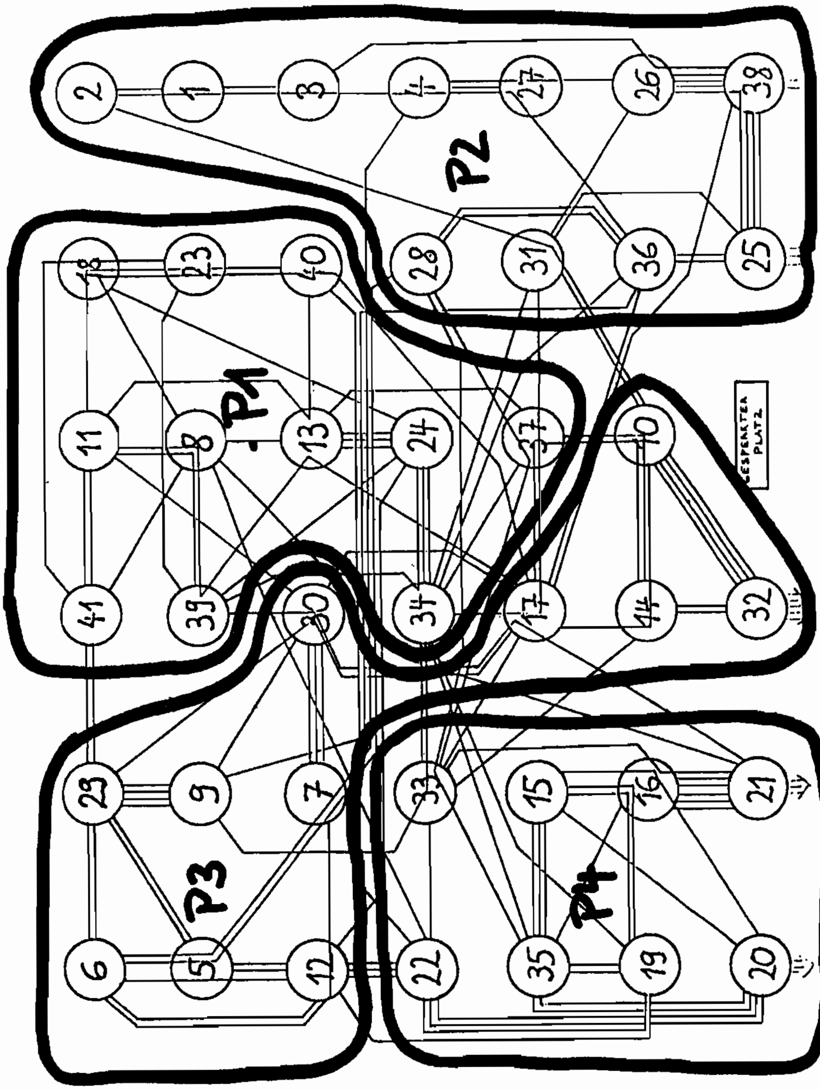


FIGURE 2 B41A, PNS, Partition sizes 11/11/11/8, 38 total cuts

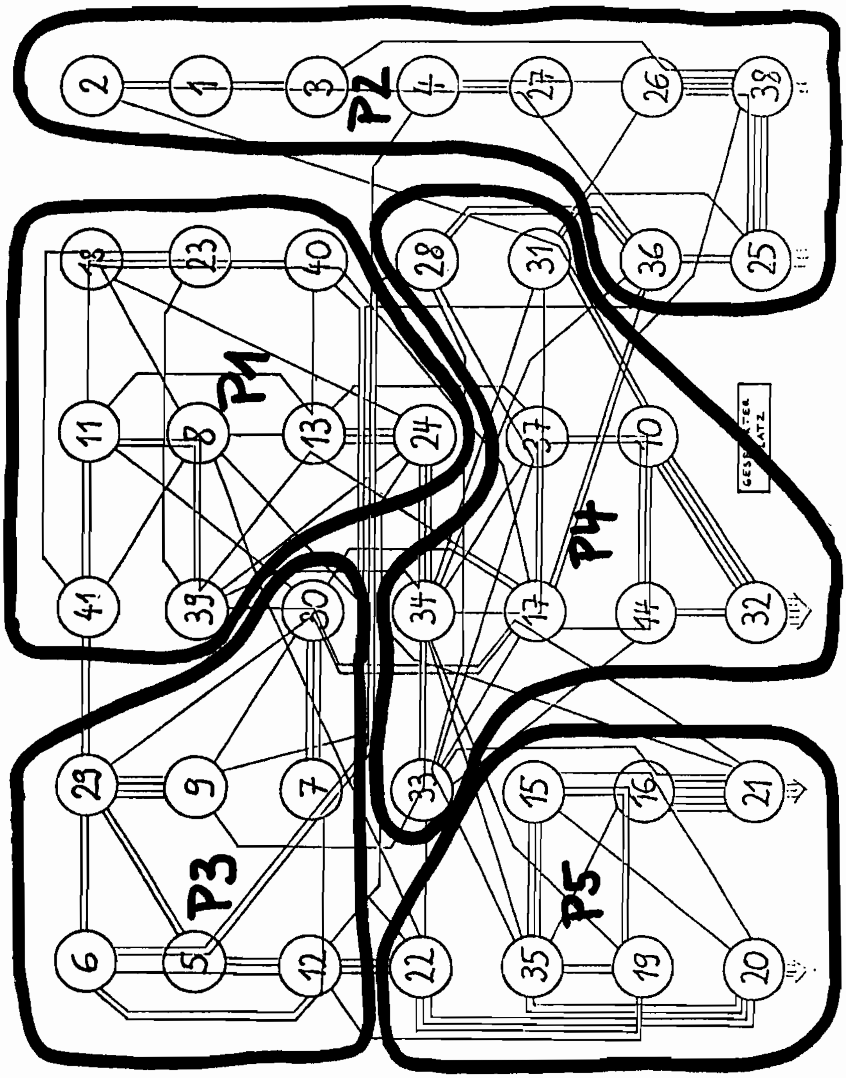


FIGURE 3 B41A, PNS, Partitions sizes 9/9/9/7/7; 39 total cuts

TABLE 1

# of partitions	method	size of partitions	cut potentials	total potential cuts	control parameters	# of iterations
1	2	3	4	5	6	7
2	KOLOSS	19/22	21	21	$T_o = 5$	40 807
		20/21	20	20		30 733
	UTA	18/23	18	18	EV = 1.5, 3.0	37
	PNS	20/21	20	20	u=-1.0,-0.9, -0.8,-0.5,-0.2	172
		19/22	21	21	u=-1.5,-1.0,-0.5	114
4	KOLOSS	10/11/10/10	36	47	$T_o = 4$	44 747
	UTA	8/11/10/12	35	37	EV=1.3,1.5,3.0	81
	PNS	11/11/11/8		38	u=-1.0,-0.5,-0.3, -0.2,-0.1	229
5	KOLOSS	9/9/7/9/7	42	44	$T_o = 5$	42 046
		9/9/7/9/7	40	47	$T_o = 4$	41 076
	UTA	10/7/8/7/9	36	38	EV = 3.0	22
	PNS	9/9/9/7/7		39	u=-4.0,-3.0, -2.0,-1.0,-0.5	165

wing relation holds

$$\left(\prod_k c_{kk} \right)^{\bullet} \geqslant 0.$$

The same holds for the eq. (II) with

$$\prod_k \dot{c}_{kk} \geqslant \prod_k c_{kk} \quad .$$

The equality is valid for the fixed points of eqs. (I), (II).

Eq. (II) extented by a mutation rate control m

$$\dot{x}_{ik} = \left(x_{ik} \frac{(Gx_{\cdot k})_i}{c_{kk}} + m \right) / e_i \qquad (III)$$

was implemented on a computer as a method for graph partitioning because equ. (II) full fills the extremal equ. (++) corresponding to property (iii).

Some results for an industrial example are shown in fig. 1, 2, 3.

Some comperative results for the same example with other methods are shown in the table 1. KOLOSS represents a thermodynamic method, UTA a fuzzy "thermodynamic" method, and PNS the partitioning by natural selection. These methods and results are explained in more detail in Voigt (1985).

Acknowledgment

The author would like to thank Professor M. Peschel for the steady stimulation during the development of this work.

REFERENCES

Brady, R.M. (1985). Optimization strategies gleaned from biological evolution. Nature, vol. 317: 804 - 806.

Eigen, M., Schuster, P. (1979). The Hypercycle, a Principle of Natural Selforganization. Springer, New York.

Ewens, W.J. (1979). Mathematical Population Genetics. Springer, New York.

Hofbauer, J., Sigmund, K. (1984). Evolutionstheorie und dynamische Systeme. Paul Parey, Berlin.

Kirkpatrick, S., Gelatt Jr., C.D., Vecchi, M.P. (1983). Optimization by Simulated Annealing. Science, vol. 220: 671 - 680.

Peschel, M., Mende, W. (1985). Do we live in a Volterra world? Springer, New York.

Voigt, H.-M. (Ed.) (1985). Structure Formation based on Selforganization Principles with Applications to Elekctronics Design. IIR-Informationen 1 (1985) 12, Acad. Sci. GDR.

Zeeman, E.C. (1980). Population Dynamics from Game Theory. In: Lect. Notes 819, Springer, New York.

RANK PRESERVATION AND PROPAGATION OF FUZZINESS IN PAIRWISE-COMPARISON METHODS FOR MULTI-CRITERIA DECISION ANALYSIS

F.A. Lootsma
Department of Mathematics and Informatics
Delft University of Technology
P.O. Box 356, 2600 AJ Delft, The Netherlands

ABSTRACT

We consider pairwise-comparison methods to rank and rate a finite number of stimuli. The decision makers express their preference ratios on a category scale: their responses are restricted to a set of categories labelled with a narrative degree of preference. We put these qualifications on numerical scales with geometric progression, and show that the rank order of the stimuli is preserved when the base number varies. If the preference ratios are expressed in fuzzy numbers with triangular membership functions, we find that the triangles usually exhibit a particular form of symmetry: the parameters satisfy the golden-section rule so that the logarithms have membership functions with isosceles triangles. Under the additional condition that the preference ratios have a uniform degree of fuzziness, we establish the membership functions of the decision criteria, the alternatives and the final scores, and we show that their degree of fuzziness depends on the hierarchical decision level only.

1. PAIRWISE-COMPARISON METHODS

In a method of pairwise comparisons (David (1963), Saaty (1980)) *stimuli* are presented in pairs to one or more decision makers. The basic experiment is the comparison of two stimuli S_i and S_j by a single decision maker who is requested to express his preference (if any) for one of the two. We assume that the stimuli S_i have respective values V_i $(i = 1, \ldots, p)$ on a numerical scale such that $\Sigma V_i = 1$. The purpose of the experiments is to estimate these values. Usually, the decision maker is requested to estimate the ratio V_i/V_j, and in order to assess the preference ratio we use a category scale instead of a real magnitude scale. The conversion is straightforward, requiring only that the *responses* be restricted to a set of categories with a narrative degree of preference. Thus, the decision maker is merely asked to make *qualitative* statements such as

no preference for S_i or S_j,
some preference for S_i over S_j,
strong preference for S_i over S_j ,

whereafter numbers are assigned to the respective categories in order to translate the judgements into estimated preference ratios r_{ij} (the estimates of V_i/V_j). Because we assume that a consistent decision maker satisfies the transitivity property $r_{ij} = r_{i\ell}\, r_{\ell j}$, these numbers should exhibit a geometric progression. Comparing S_i versus S_j we therefore deviate from the scale of Saaty (1980), and we set

$$r_{ij} = \exp(\gamma\ \delta_{ij}) \tag{1}$$

with

$$\delta_{ij} = \begin{array}{ll} 0: & \text{no preference,} \\ 2: & \text{some preference for } S_i, \\ 4: & \text{strong preference for } S_i, \\ 6: & \text{very strong preference for } S_i. \end{array}$$

The intermediate values 1, 3, and 5 are given when the decision maker hesitates between two adjacent qualifications. Of course, we have

$$\delta_{ij} = \begin{array}{ll} -2: & \text{some preference for } S_j, \\ -4: & \text{strong preference for } S_j, \\ -6: & \text{very strong preference for } S_j. \end{array}$$

Many experiments have been conducted by Légrády et al. (1984), Lootsma et al. (1984), and Kok et al. (1985) to find workable vàlues for the scale parameter γ. In practice, we use a *normal* scale (with $\gamma = ½$) and a *widely stretched* scale (with $\gamma = 1$), for reasons which are explained in the cited papers.

In general, we are dealing with a decision-making committee, and we have to allow for the possibility that some members sometimes abstain from giving their opinion. That is the reason why we employ *logarithmic regression* (de Graan (1979)) instead of the *eigenvalue method* (Saaty (1980)). We let D_{ij} stand for the set of decision makers who judged S_i versus S_j, and N_{ij} for the cardinality of D_{ij}. We take r_{ijk} to denote the estimate of V_i/V_j by the k-th decision maker. Now we approximate the vector $V = (V_1,\ldots, V_p)$ by the normalized vector $\bar{v}$ which minimizes

$$\sum_{i<j} \sum_{k \in D_{ij}} (\ln r_{ijk} - \ln v_i + \ln v_j)^2 = \tag{2}$$

$$= \sum_{i<j} \sum_{k \in D_{ij}} (q_{ijk} - w_i + w_j)^2 ,$$

where $q_{ijk} = \ln r_{ijk}$ and $w_i = \ln v_i$. Observing that $q_{ijk} = -q_{jik}$ and $N_{ij} = N_{ji}$, we can write the associated normal equations as

$$w_i \sum_{\substack{j=1 \\ j \neq i}}^{p} N_{ij} - \sum_{\substack{j=1 \\ j \neq i}}^{p} N_{ij} w_j = \sum_{\substack{j=1 \\ j \neq i}}^{p} \sum_{k \in D_{ij}} q_{ijk} . \tag{3}$$

These equations are dependent (they sum to the zero equation). Taking the vector $\bar{w}$ to denote a particular solution of (3), we can write the components of the general solutions as $\bar{w}_i + \eta$, and we approximate V_i by

$$\bar{v}_i = \exp(\bar{w}_i + \eta) / \sum_{i=1}^{p} \exp(\bar{w}_i + \eta) = \exp(\tilde{w}_i) / \sum_{i=1}^{p} \exp(\bar{w}_i) . \tag{4}$$

Obviously, normalization is sufficient to remove the additive degree of freedom in solutions of (3).

2. RANK PRESERVATION

Let us now use the property that we are working on numerical scales with geometric progression (see formula (1)) and substitute $q_{ijk} = \ln r_{ijk} = \gamma\ \delta_{ijk}$ into (3) to obtain

$$w_i \sum_{\substack{j=1 \\ j \neq i}}^{p} N_{ij} - \sum_{\substack{j=1 \\ j \neq i}}^{p} N_{ij} w_j = \gamma \sum_{\substack{j=1 \\ j \neq i}}^{p} \sum_{k \in D_{ij}} \delta_{ijk} . \tag{5}$$

Introducing $\bar{w}(1)$ to denote a solution of (5) for $\gamma = 1$, we can write any solution for arbitrary γ in the form

$$\bar{w}_i(\gamma) = \gamma[\bar{w}_i(1) + \eta] .$$

Thus, we approximate V_i by

$$\bar{v}_i(\gamma) = \exp[\gamma\ \bar{w}_i(1)] / \sum_{i=1}^{p} \exp[\gamma\ \bar{w}_i(1)]\ ,$$

an expression which immediately shows that the rank order of the $\bar{v}_i(\gamma)$ does not depend on γ. Moreover, the experiments by Légrády et al. (1984) and Lootsma et al. (1984) demonstrate that the $\bar{v}_i(\gamma)$ are not very sensitive to the scale parameter γ when it varies over the range of values (including ½ and 1) that seems to be acceptable in the social sciences.

In summary, we have the result that scale transformations do not produce rank reversal in a *single-level* comparison of stimuli. In hierarchical decision making, however, where we compare decision criteria (level 1) and alternatives under each of the criteria (level 2) in order to obtain final scores for the alternatives, rank preservation cannot be guaranteed. The two-level budget-allocation problem in Légrády et al. (1984) shows that the rank order of the final scores may change, although the scale sensitivity remains low.

Even on a single level, however, the vexing problem of rank preservation has not completely been solved. Saaty and Vargas (1984), for instance, were concerned with three methods for approximating the V_i, starting from a *given* matrix of preference ratios so that the effect of scale transformations is ignored. They demonstrate that the results of the eigenvalue method (Saaty (1980)), the least-squares method, and our logarithmic regression method may indeed exhibit rank reversal in simple cases where it should clearly not occur. The observation that the calculated rank order is method-dependent may also be found in Légrády et al. (1984) who compared the results of logarithmic regression and the Bradley-Terry method. And let us finally point at the unresolved complications of adding or dropping a stimulus: approximating the values of the remaining stimuli one may also run up against the phenomenon of rank reversal (Saaty, private communication).

3. FUZZY PAIRWISE COMPARISONS

The method described so far has a particular drawback. The decision makers are supposed to supply *deterministic estimates* of the preference ratios, but mostly they only have a *fuzzy* notion of them. Hence, the approximations to the V_i ($i = 1, \ldots, p$) suggest an accuracy which is out of proportion. This glaring deficiency of the method prompted Van Laarhoven and Pedrycz (1983) to propose a modified version allowing the decision makers to supply *fuzzy estimates* of their preference ratios (see also Lootsma (1985)). For reasons of simplicity, they expressed the estimates in fuzzy numbers with *triangular* *membership functions*, and they modified the algebraic rules of Dubois and Prade (1980) to simplify the subsequent calculations. Thus, they asked the k-th decision maker to supply three estimates of V_i/V_j: an estimate which is taken to be the *modal value* r_{ijkm} locating the top (with value 1) of the triangle, and two estimates, the *lower value* r_{ijkl} and the *upper value* r_{ijku}, to locate the support of the triangle. The resulting membership function of the fuzzy number $\tilde{r}_{ijk}$ is represented by the triple (r_{ijkl}, r_{ijkm}, r_{ijku}). With a negligible error, the authors take

$$\tilde{q}_{ijk} \stackrel{\text{def}}{=} \ln \tilde{r}_{ijk}$$

to be a fuzzy number with a triangular membership function as well. The three parameters q_{ijkl}, q_{ijkm}, q_{ijku} satisfy the relations

$$q_{ijkl} = \ln r_{ijkl}, \text{ etc.}$$

Basically, the authors propose to approximate the values V_i of the respective stimuli S_i by fuzzy numbers $\tilde{v}_i$ with triangular membership functions. The procedure is as follows. Let $\tilde{w}_i = \ln \tilde{v}_i$, generalize (3) so that

$$\tilde{w}_i \sum_{\substack{j=1\\j\neq i}}^{p} N_{ij} - \sum_{\substack{j=1\\j\neq i}}^{p} N_{ij}\, \tilde{w}_j = \sum_{\substack{j=1\\j\neq i}}^{p} \sum_{k\in D_{ij}} \tilde{q}_{ijk} \quad , \qquad (6)$$

and calculate the lower, modal and upper values of $\tilde{w}_i$ from the linear systems

$$w_{im} \sum_{\substack{j=1\\j\neq i}}^{p} N_{ij} - \sum_{\substack{j=1\\j\neq i}}^{p} N_{ij}\, w_{jm} = \sum_{\substack{j=1\\j\neq i}}^{p} \sum_{k\in D_{ij}} q_{ijkm} \,, \qquad (7)$$

$$w_{il} \sum_{\substack{j=1\\j\neq i}}^{p} N_{ij} - \sum_{\substack{j=1\\j\neq i}}^{p} N_{ij}\, w_{ju} = \sum_{\substack{j=1\\j\neq i}}^{p} \sum_{k\in D_{ij}} q_{ijkl} \,, \qquad (8)$$

$$w_{iu} \sum_{\substack{j=1\\j\neq i}}^{p} N_{ij} - \sum_{\substack{j=1\\j\neq i}}^{p} N_{ij}\, w_{jl} = \sum_{\substack{j=1\\j\neq i}}^{p} \sum_{k\in D_{ij}} q_{ijku} \,. \qquad (9)$$

Obviously, the modal values appear separately in the system (7) only. Because $q_{ijkm} = -q_{jikm}$ and $N_{ij} = N_{ji}$, we can show that the equations sum to the zero equation. Hence, a solution to (7) has at least one (additive) degree of freedom. We have a similar result for the lower and upper values, to be solved jointly from the system (8)-(9): because $q_{ijkl} = -q_{jiku}$ and $q_{ijku} = -q_{jikl}$, the equations sum to the zero equation, so that a solution of (8)-(9) has at least one (additive) degree of freedom.

In general, the systems (7) and (8)-(9) have exactly one degree of freedom (violations of the rule are extensively discussed in Boender et al. (1985)). A fuzzy approximation to the value V_i of stimulus S_i is represented by the triple

$$(\exp(w_{il}),\ \exp(w_{im}),\ \exp(w_{iu})) \ ,$$

and to enforce uniqueness, Van Laarhoven and Pedrycz (1983) introduced the *normalized* fuzzy approximation

$$\tilde{v}_i = \left(\frac{\exp(w_{il})}{\sum_{i=1}^{p} \exp(w_{iu})} \ , \ \frac{\exp(w_{im})}{\sum_{i=1}^{p} \exp(w_{im})} \ , \ \frac{\exp(w_{iu})}{\sum_{i=1}^{p} \exp(w_{il})} \right) \ .$$

The proposed method still has two short-comings. First, the triple (w_{il}, w_{im}, w_{iu}) obtained by solving the systems (7) and (8)-(9) does not necessarily satisfy the order relations

$$w_{il} \leq w_{im} \leq w_{iu} \ , \tag{10}$$

so that it does not always represent a fuzzy number with triangular membership function. Second, the decision makers have to supply much more information than in the deterministic case, and they are not always willing to do so.

4. A UNIFORM DEGREE OF FUZZINESS

The responses of the decision makers are in fact limited to a finite number of categories (no preference, some, strong, or very strong preference). Hence, the deterministic estimates r_{ijk}, written as

$$r_{ijk} = \exp(\gamma\ \delta_{ijk}) ,$$

are also restricted to a finite, usually small, set of values ($\delta_{ijk} = 0, \pm 1, \ldots, \pm 6$). Boender et al. (1985) observed that the fuzzy estimates $\tilde{r}_{ijk}$ are subject to similar limitations. Mostly, they exhibit a particular form of symmetry: the three parameters of the respective membership functions can be written as

$$r_{ijkl} = \exp[\gamma\ (\delta_{ijk} - \alpha_{ijk})],$$

$$r_{ijkm} = \exp[\gamma\ \delta_{ijk}] , \tag{11}$$

$$r_{ijku} = \exp[\gamma(\delta_{ijk} + \alpha_{ijk})],$$

with positive, integer-valued *spread* α_{ijk}. The membership functions of the associated q_{ijk} are accordingly *isosceles triangles*. What simplifies matters even more, is the observation that practically all α_{ijk} are equal in a given situation: equal to 2 (a spread of two scale steps) if the decision makers express a significant degree of fuzziness.

In the case that (11) holds we attempt to solve the equations (6) in terms of fuzzy $\tilde{w}_i$ with the same type of symmetry as the $\tilde{q}_{ijk}$. Observing that

$$q_{ijku} - q_{ijkm} = q_{ijkm} - q_{ijkl} = \gamma\ \alpha_{ijk} ,$$

we write

$$w_{iu} - w_{im} = w_{im} - w_{il} = \gamma\ \beta_i \tag{12}$$

with positive spread β_i. The modal values w_{im} are solved from the system (7). Subtracting (8) from (9) we find that the β_i satisfy the system

$$\beta_i \sum_{\substack{j=1 \\ j\neq i}}^{p} N_{ij} + \sum_{\substack{j=1 \\ j\neq i}}^{p} N_{ij}\ \beta_j = \sum_{\substack{j=1 \\ j\neq i}}^{p} \sum_{k\in D_{ij}} \alpha_{ijk} . \tag{13}$$

Now, restricting ourselves to the situation that the spreads α_{ijk} are equal to a *uniform* spread α, we reduce (13) to

$$\beta_i \sum_{\substack{j=1\\ j\neq i}}^{p} N_{ij} + \sum_{\substack{j=1\\ j\neq i}}^{p} N_{ij}\, \beta_j = \alpha \sum_{\substack{j=1\\ j\neq i}}^{p} N_{ij} . \tag{14}$$

Assuming that $\beta_i = \beta$, $i = 1, \ldots, p$, a reasonable assumption in the given situation, we can immediately conclude from (14) that $\beta = \frac{1}{2}\alpha$. This is an extremely simple result, which also explains the success of Van Laarhoven's and Pedrycz' method (1983). They propose to solve the systems (7) and (8)-(9) directly, although the resulting solution may violate the order relations (10). Violations are unlikely, however, because in practical applications the systems (8)-(9) and (13) are perturbations of (14). A fuzzy approximation to the value V_i of the stimulus S_i is now given by

$$(\exp(w_{im} - \tfrac{1}{2}\alpha\gamma),\ \exp(w_{im}),\ \exp(w_{im} + \tfrac{1}{2}\alpha\gamma)),$$

and the normalized approximation can be written in the form

$$\tilde{v}_i = \frac{\exp(w_{im})}{\sum_{i=1}^{p} \exp(w_{im})} \quad (\exp(-\alpha\gamma),\ 1,\ \exp(\alpha\gamma)). \tag{15}$$

We may clearly take the factor

$$\frac{\exp(w_{im})}{\sum_{i=1}^{p} \exp(w_{im})} \tag{16}$$

to be the deterministic estimate of V_i when fuzziness in human judgement is ignored (see formula (4)). The fuzzy factor   represented by the triple

$$(\exp(-\alpha\gamma),\ 1,\ \exp(\alpha\gamma)) \tag{17}$$

makes it a fuzzy estimate. Formula (15) may be applied when the degree of fuzziness is of the same order of magnitude for all judgements.
Usually we distinguish

no fuzziness:	$\alpha = 0$,
moderate fuzziness:	$\alpha = 1$ (one scale step upwards and downwards),
significant fuzziness:	$\alpha = 2$ (two scale steps upwards and downwards).

Substitution of these values in (15) yields an effective tool for a simple form of *sensitivity analysis*. No additional information, such as lower and upper values of the estimated preference ratios, is required. We only ask the decision makers to estimate their *uniform* degree of fuzziness. Obviously, the approximations to the V_i will exhibit exactly the same degree of fuzziness.

5. PROPAGATION OF FUZZINESS

Let us now concern ourselves with the typical problem of multi-criteria decision analysis: the evaluation of a finite number of alternatives under conflicting criteria. The usual procedure is to approximate the values C_i of the criteria (the stimuli at decision level 1) and the values A_{ij} of the alternatives (the stimuli at decision level 2) under the respective criteria. Our objective is to calculate final scores s_j enabling us to rank and rate the alternatives. Taking c_i and a_{ij} to denote the approximations in question we may set

$$s_j = \sum_i c_i \, a_{ij} \,, \tag{18}$$

an operation motivated by the present author (1985) under the condition that the approximations are the result of a pairwise-comparison method using *ratio* estimates (and this is exactly the situation wherein we find ourselves, see sec. 1).

Suppose now that we have used the pairwise-comparison method of secs. 3 and 4, yielding fuzzy approximations

$$\tilde{c}_i = c_{im} \, (\exp(-\alpha\gamma),\ 1,\ \exp(\alpha\gamma))$$

$$\tilde{a}_{ij} = a_{ijm} \, (\exp(-\alpha\gamma),\ 1,\ \exp(\alpha\gamma))$$

to the values of the stimuli just mentioned. Of course, the symbols c_{im} and a_{ijm} stand for the modal values of the fuzzy numbers $\tilde{c}_i$ and $\tilde{a}_{ij}$. Following the algebraic rules of Van Laarhoven and Pedrycz (1983) we obtain the fuzzy final scores

$$\tilde{s}_j = \sum_i c_{im} \, a_{ijm} \, (\exp(-2\alpha\gamma),\ 1,\ \exp(2\alpha\gamma)), \tag{19}$$

an expression which clearly shows the propagation of fuzziness in hierarchical decision making. These results can easily be generalized to comprise the case that we have a uniform degree of fuzziness at each level separately, but not necessarily the *same* degree at all levels.

We conclude this paper with the critical note of Boender et al. (1985): whose fuzziness did we consider here? As we have seen in solving (14), the spread β of the approximations to the values of the stimuli is of the same order of magnitude as the spread α in the judgemental statements. From the viewpoint of the decision makers who were involved in the deliberations this is correct: they realize that the approximations constitute a *compromise solution*, which is not fuzzier than the original preference ratios themselves. For an outside observer, however, the picture is entirely different. Hearing the conflicting opinions of the experts in the decision-making committee, he may conclude that the *truth* is still far to seek. This viewpoint has not been studied in the present paper.

REFERENCES

[1] Boender, C.G.E., Graan, J.G. de, and Lootsma, F.A., Pairwise-Comparison Methods using Fuzzy Numbers on Ratio Scales with Geometric Progression. Report 85-24, Department of Mathematics and Informatics, Delft University of Technology, Delft, Netherlands, 1985.

[2] David, H.A., The Method of Paired Comparisons. Griffin, London, 1963.

[3] Dubois, D., and Prade, H., Fuzzy Sets and Systems: Theory and Applications. Academic Press, New York, 1980.

[4] Graan, J.G. de, Extensions to the Multiple Criteria Analysis Method of T.L. Saaty. National Institute for Water Supply, Voorburg, Netherlands, 1980.

[5] Kok, M., and Lootsma, F.A., Pairwise-Comparison Methods in Multi-objective Programming, with Applications in a Long-term Energy-planning Model. European Journal of Operational Research 22, 44-55, 1985.

[6] Laarhoven, P.J.M. van, and Pedrycz, W.,A Fuzzy Extension of Saaty's Priority Theory. Fuzzy Sets and Systems 11, 229-241, 1983.

[7] Légrády, K., Lootsma, F.A., Meisner, J., and Schellemans, F., Multi-criteria Decision Analysis to aid Budget Allocation. In M. Grauer and A.P. Wierzbicki (eds.), Interactive Decision Analysis. Springer, Berlin, 164-174, 1984.

[8] Lootsma, F.A., Performance Evaluation of Non-linear Optimization Methods via Pairwise Comparison and Fuzzy Numbers. Mathematical Programming 33, 93-114, 1985.

[9] Lootsma, F.A., Meisner, J., Schellemans, F., Multi-criteria Decision Analysis as an Aid to Strategic Planning of Energy Research and Development. Report 84-02, Department of Mathematics an Informatics, Delft University of Technology, Delft, Netherlands, 1984. To appear in the European Journal of Operational Research.

[10] Saaty, Th.L., The Analytic Hierarchy Process, Planning, Priority Setting, Resour ce Allocation. McGraw-Hill, New York, 1980.

[11] Saaty, Th.L., and Vargas, L.G., Inconsistency and Rank Preservation. Journal of Mathematical Psychology 28, 205-214, 1984.

LEARNING IN DECISION MAKING

Robert GENSER
Federal Ministry for Public Economy and Transport, KoVpol
Operngasse 20 b/3, A-1040 VIENNA, Austria

1. INTRODUCTION

The purpose of decision making can be classified in following way.

Detection of signals or of specific states, like the decision on a mechanical equipment, if it may have been got dangerous because of wear out. This is the domain of application for optimal filtering - see f.e. (Sawaragi et al. 1967), (Willsky 1976), and (Manetsch 1984) -, identification (Isermann 1980), diagnosis systems (Pau 1981), and expert systems - see f.e. (Barr et al. 1982), (Hayes et al. 1983), (Fikes et al. 1985), and (Hayes 1985).

Setting of actions, like in control or operational research. The problem to find the optimal renewal strategy for an equipment would belong to this area. Optimizing methods and evolution strategies are used for the solution of such problems, see f.e. (Rau 1970), (Tillman et al. 1981), and (Inagaki et al.1978).

Selection of alternatives and setting of priorities, as it is the problem for procurement of equipments for example. Ranking methods are applied, see f.a. (Roy 1977), (Baas et al. 1977), and (Korhonen et al. 1980).

The way of decision making depends also on the time horizon in consideration. It can be distinguished between the area of strategic planning in the long-range and operational planning in the short-range level. Strategic planning has to deal with non-terminating ill-structured, ill-defined problems. Operational planning has to handle well-structured problems with specific decisions at defined points in time.

Decision making for the selection of alternatives and for setting of priorities in the level of strategic planning has importance in the field of large scale transportation systems. Those aspects are considered in the following sections.

2. PROBLEMS IN DECISION MAKING

In case of large scale long-range problems the decision making process depends on:
- experience,
- data,

- simulation possibilities, and
- optimization methods.

Experience in the sense of skill will be stressed later. Of course, the impact of experience is given in case of customs, usage or course of dealing as it is called in the field of law and regulations. Morever it is understood at present that the information of documentation can be grasped in context only.

In Figure 1 the structure is shown in which decision making is embedded. It is a dynamical process with feedforward

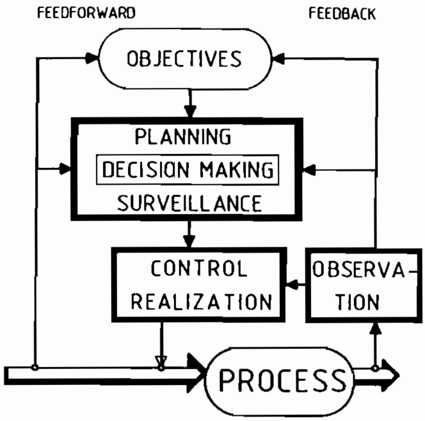


FIGURE 1 Structure of a decision making process

and feedback loops for the information flow. Even the dynamical behaviour of objectives itself has to be considered. Feedforward stimulates also the variation needed for processes with evolutionary strategies; the selection will be caused by the feedback.

Decision making by help of electronic data processing dealt with quantitative information and fixed criteria functions in the beginning. Then it was tried to consider qualities and objectives with different or even conflicting aspects by help of weighting functions. The classical quadratic criteria function of control theory which considers energy consumption, deviation or error, and time is an example for this approach. Number systems have been introduced for dealing with qualities like it was done in the field of reliability for example. Dynamic processes, change of states or of parameters have been mastered. Methods had been developed for processing of probabilities or stochastic data. But possibilities or fuzzy information require other structures for processing as it is suitable for dealing with deterministic or stochastic data.

As it is common for the solution of complex problems an engineering approach is required. But in such a way the approach to solve a given problem depends on the actual situation and comprises not only algorithms but also the organization, the ergonomical aspects etc.

Decision making in transportation planning had been developed in the same way. But it had have given also input for stimulating new approaches, e.g. (Roy et al. 1977). Of course, many efforts were given to tackle problems by fixing weights to the different objectives and by applying an all-over criteria function, see f.e. (Snizek 1985), and (Meier et al. 1985). But this attempt causes following problems.

For example in case of procurement the weights have to be fixed beforehand for more than 300 attributes for specifying complex systems, if transparency for the procurement is demanded. But any change, as it can be because of unexpected qualities of tenders f.e., would cause changes of weights. But it would be difficult to recognize which changes of weights are suitable or which effects would be entailed. Especially the interdependence of weights can hide the real intentions. In case of using weights and an all-over criteria function as usually, the distinction, if all has the same importance or all is very important is lost.

The other disadvantage of such an approach is that the individual opinion of group members and the changing of these opinions about the importance of single attributs is suppressed.

Moreover it has to be recognized that the information available is given in a fuzzy way. The objectives of decision makers or actuators depend on states, information available etc. If an objective is not reached for long time then this goal can be become very important, f.e. hunger. But it can be also switched to neglection by accustoming. The way how to solve a decision problem has to consider the change of information, see Figure 2. The processing should be done according to the state of information.

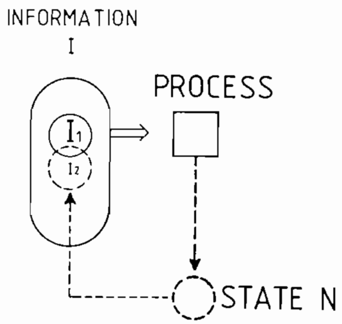


FIGURE 2 Change of information

The complexity and flexibility of objectives should be demonstrated at an example taken from the General Conception of Transport in Austria (Halbmayer 1982), see Figure 3. It was tried to distinguish between long-range (fundamental), medium range (guiding), and operational objectives. The structural aspects like given infra-structure, geographical situation etc. should be covered by the operational objectives.

Fundamental objectives are less prone to conflicts because

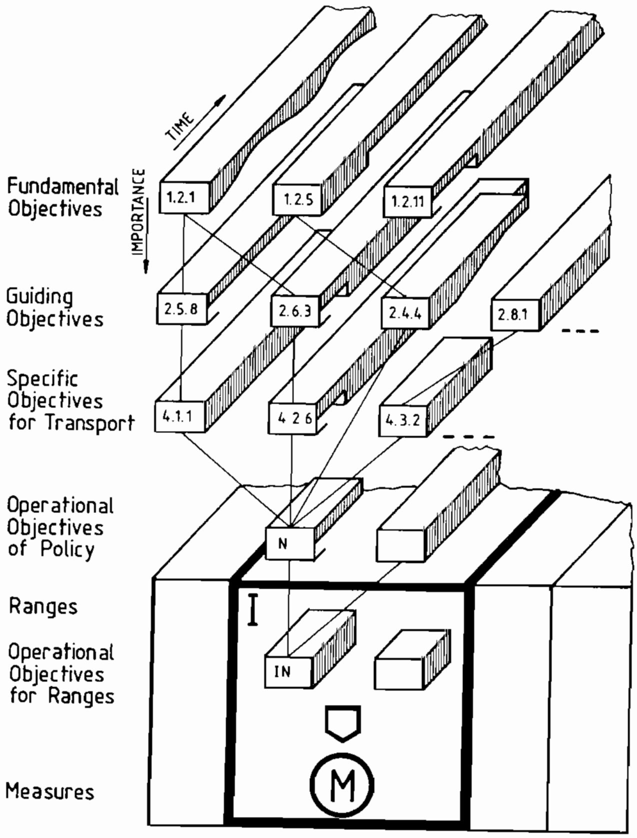


FIGURE 3 Example for the structure of the objectives at the GVK-Ö

of soft (fuzzy) information. The conflicts arise in the detail. But in this level the complexity will be reduced by limited range of time, space or function. Moreover, such conflicts can be reduced in many cases by the objectives with consensus of the higher levels. In a given situation not only one level of objectives has to be taken in consideration. It has to be moved from top to down and vice versa. The importance of single objectives may change and new objectives may be created in lower levels. Considering the fundamental objectives 1.2.1. (living condition should be most equal for population...), 1.2.5 (safety for public should be high...), and 1.2.11 (accessibility to transportation systems should be given ...) as well as the guiding objectives, originating from other fields, like 2.5.8 (physical and psychical effort should be considered...), 2.6.3. (disabled should be considered ...), 2.4.4 (high dependability or reliability should be gained...), and 2.8.1 (use of energy and of resources should be rational ...) together with the specific objectives for the field of transportation 4.1.1 (quality ...), 4.2.6 (disadvantage for user...), and 4.3.2 (energy and resources...) the operational objective N (accessibility for disabled in normal and abnormal situation ...) was developed. This implies the operational objective IN (to prefer ramps instead of escalators...) and the measure M (new regulation for construction of station's buildings) in the planning level for the range I (new stations for commuter's transport).

Figure 4 shows to which extent it is rational to comprehend and cluster the different objectives. This approach assumes that in concrete cases given practically, most of the individual objectives of the different levels can be neglected or combined. If this comprehension would be too coarse, in case of conflicts f.e., then one or more categories have to be splitted up again, as it is indicated in Figure 4. Because of

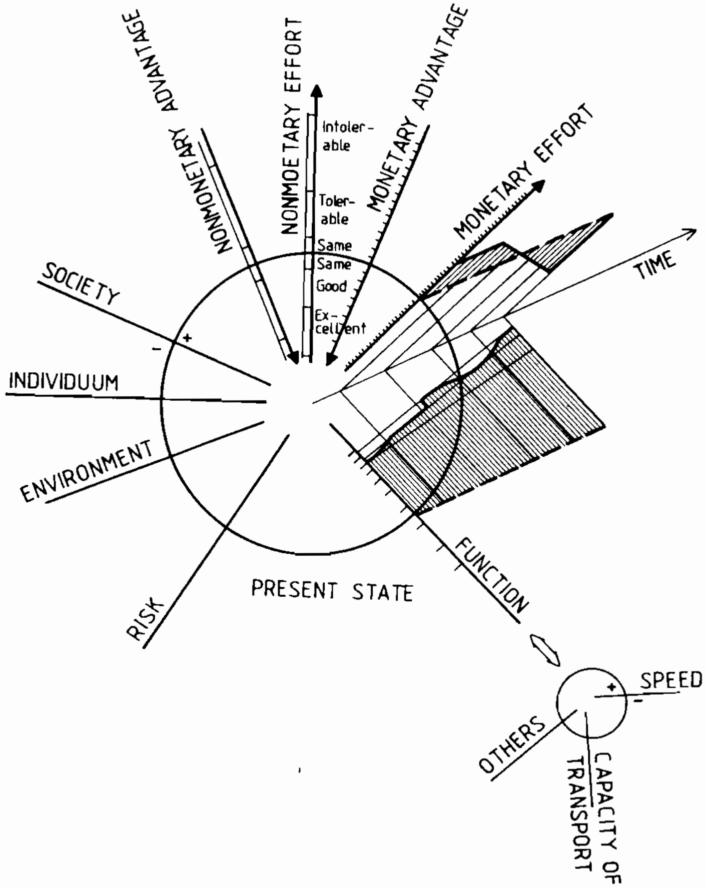


FIGURE 4 Evaluating space for categories of objectives

the dynamical behaviour of a system which should be investigated concerning the effects of different measures, but also by the dynamics of objectives itself, the examination has to be done over the time. The change of the value for the monetary expenses and of the result for the functional category with possible variances is shown over the time in Figure 4 as an example.

This short insight into the problems given in decision making for strategic planning in case of large scale transportation systems points out that the dynamical behaviour of the process and the environment of fuzzy information has to be taken in consideration.

Dynamical processes require dynamical planning. This means it has be planned also for changes and, what is important, a feedback for learning has to be provided.

Fuzzy information can not be improved by introducing probability functions rather a feedback loop has to be provided for handling the possibilities. If an information is missing then it has to be investigated, if this information has any influence or can be neglected. Sensitivity analysis may be applied for example. But if this missing information could have great influence on the result of decisions then steps have to be taken to get this information, e.g. by test or pilot project, by simulation etc. It is needed to have the ability to learn. Ignorance should not be hidden by probabilities assumed.

One step decision making is not possible for complex problems generally. It has to be strived for a decision making process which has dynamical abilities for handling these problems. Such a conception has to comprise not only algorithms, but also the organization.

3. LEARNING

As it is common in complex information processing, the handling of fuzziness requires a dynamic process with feedback. But this loops are needed for learning. This is in difference to feedback loops in simple control systems. In step by step decision making it is strived for getting access to usefull information and it is considered the freedom of actions according possible states. The learning in decision making processes has following purposes:
- selection of data needed (Genser 1969)
- optimizing the handling and use of data
- recognizing objectives and their qualities or behaviour
- recognizing change of objectives
- optimal use of possibilities for corrections and adaptions
- optimal control of decision making process, considering rationality and improvements.

It is important for understanding the objectives to recognize the influence and importance of objectives. The aspiration released by a member of a decision making group should be recognizable. Recognizing the change of objectives requires to grasp the process of consensus and to consider the impact of reality in respect to the different aspirations etc.

Learning implies to get experience, to store information and to make accessible this experience. As shown also in (Tsypkin 1973), (Csibi 1975), (Houston 1976), (Saridis 1977), and (El-Fattah et al. 1978) learning can be done by:
- storage of patterns or boundaries,
- copying,
- conditional coordination,
- trial and error,
- searching by optimization,
- teaching,

- grasping or
- accustoming.

The learning process has influence for better understanding and handling of the decision making process. It allows to recognize the impact of decisions taken. It helps to recognize the interdependence of objectives as well as the aspirations of control environment like decision making group, acting commitee, and public opinion. The mutual understanding of persons involed is improved by accustoming. The process is speeded up by retrieval of similar cases (Genser 1977) and complex algorithms can be avoided, especially nonlinear processes can be treated by linear and simple methods (Genser 1970). In any case an improvement of performance can be expected in the long run. But learning is not only for increasing knowledge. Only learning can improve skill. Skill is required for gaining high performance and for suitable use of knowledge. Even the run - in for mechanical equipments may be considered as learning skill.

Learning requires pattern recognition ability, see f.e. (Tou et al. 1974), (Fu et al. 1976), (Batchelor 1974) or (Ivakhnenko 1970). Pattern recognition is needed for grasping aspirations instead of the common criteria function with weighting. Pattern recognition helps also to grasp real situation or to classify alternatives for example.

4. DECISION MAKING FOR STRATEGIC PLANNING IN TRANSPORTATION

Figure 5 shows schematically the embedding of a decision making support modul. Complex problems have to be solved by a group of decision makers. It is an interactive process. Persons involved have to be trained and got adapted to the background of reality. Data can not be full understood without knowing the context. This limitation of documentation has to

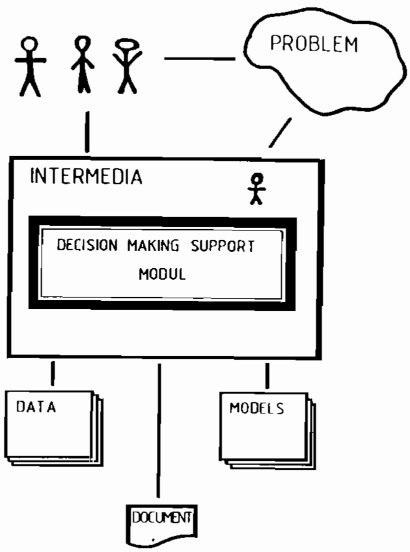


FIGURE 5 The embedding of a decision making support modul

be considered also for the intermedia.

The decision making process should be unterstandable. The persons involved should have the feeling at least, they can control this process.

The results of the decision making process as well as the way how the results have had been recieved have to be documented for improving transparency and understanding, but also for control and learning. No important information should be lost.

A good approach balances the effort for a model in respect to uncertainty: the higher uncertainty the more simple and the easier understandable the models should be.

A good decision making support system can be achieved, if at least the organization, the presentation of information, the ways of interactions, the algorithms, and the documentation for learning and control have had been well selected.
It has to be an ergonomical solution
- which makes it easy to recognize missing informations and opinions of decision makers
- which reduces problem by suitable comprehending
- which improves transparency of aspirations and of decision making process itself.

For example the aspect of man-machine interfaces for expert system is also considered in the paper of Kidd et al. (1985). Menu techniques taking into account the novice as well as the skilled user and graphical representations are helpfull of course.

The approach used in SCDAS (Selection Committee Decision Analysis System) by IIASA, see (Lewandowski et al. 1985) and (Wierzbicki 1984), can be developed for the requirements of practice.

5. CONCLUSION

Decision making in case of complex problems needs a decision making group and an interactive decision support system. An engineering approach is required for achieving a decision making process which is suitable for application in practice. This approach has to consider software with algorithms, hardware with medium for data, and organization. It has to be a solution which had taken into account human, machine, and reality. This system should be improvable by learning.

REFERENCES

Baas, S.M., and Kwakernaak, H. (1977). Rating and Ranking of Multiple-Aspect-Alternatives Using Fuzzy Sets. Automatica, 13 (1): 47-58.

Baar, A., and Feigenbaum, E. (Ed.) (1982). The Handbook of Artificial Intelligence. Pitman Books, London.

Batchelor, B.G. (1974). Practical approach to pattern classification. Plenum Press, New York.

Csibi, S. (1975). Stochastic Process with Learning Properties. Springer Verl., Berlin.

El-Fattah, Y., and Foulard, C. (1978). Learning Systems.

Springer Verl., Berlin.
Fikes, R., and Kehler, T. (1985). The role of frame-based representation in reasoning. Communications of the ACM, 28 (9): 904-920.
Fu, K.S., Keidel, W.D., and Wolter, H. (1976). Digital pattern recognition. Springer Verl., Berlin.
Genser, R. (1969). Die Aufbereitung der Information bei lernfähigen Automaten. Elektron. Rechenanlagen, 11 (6): 330-335.
Genser, R. (1970). Prediction of river flow by means of on-line-learning pattern recognition. Preprints IFAC Symposium on Systems Engineering Approach to Computer Control, Kyoto, August 1970, paper 2.8.2.
Genser, R. (1977). Scheduling and Maintenance Planning in Rail Transportation. In H. Strobel et al. (Ed.), Optimization Applied to Transportation Systems. CP-77-7. International Institute for Applied Systems Analysis, Laxenburg, Austria.
Halbmayer, K. (1982): Gesamtverkehrskonzept. Verkehr, 38 (50): 1849-1853.
Hayes-Roth, F. (1985). Rule-based systems. Communications of the ACM, 28 (9): 921-932.
Hayes-Roth, F., Waterman, D.A., and Lenat, D.B. (1983). Building Expert Systems. Addison - Wesley, Reading, Mass.
Houston, J.P. (1976). Fundamentals of learning. Academic Press, New York.
Inagaki, T., Inoue, K., and Akashi, H. (1978). Interactive Optimization of System Reliability under Multiple Objectives. IEEE Transactions, R-27 (4): 264-267.
Isermann, R. (1980). Practical Aspects of Process Identification. Automatica, 16 (5): 575-587.
Ivakhnenko, A.G. (1970). Heuristic Self-Organization in Problems of Engineering Cybernetics. Automatica, 6(3): 207-219.
Kidd, A.L., and Copper, M.B. (1985). Man-Machine interface issues in the contruction and use of an expert system. International Journal Man-Machine Studies, (22): 91-102.
Korhonen, P., and Soismaa, M. (1980). An interactive multiple criteria approach to ranking alternatives. Preprints EURO IV, Cambridge, England, July 1980.
Lewandowski, A., Toth, F., and Wierzbicki, A.P. (1985). A prototype selection committee decision system: System-implementation, tutorial example, and users' manual. International Institute for Applied Systems Analysis, Laxenburg, Austria.
Manetsch, Th.J. (1984). An Approach to Early Warning of Slowly Evolving Crises with Reference to Food Shortage Forecasting. IEEE Transactions, SMC-14 (3): 391-397.
Meier, W., Heimerl, G. (1985). Multikriterielle Beurteilung von Verkehrsinvestitonen. Schienen der Welt, 16 (7): 41-48.
Pau, L.F. (1981). Failure Diagnosis and Performance Monitoring. M. Dekker, New York.
Rau, J.G. (1970). Optimization and Probability in Systems Engineering. Van Nostrand Reinhold Comp., New York.

Roy, B. (1977). Partial Reference Analysis and Decision Aid. In B.D.E. Keeney and R. Raiffa (Ed.), Decision Making with Multiple Conflicting Objectives. International Institute for Applied Systems Analysis, Laxenburg, Austria.

Roy, B. (1980). Selektieren, Sortieren und Ordnen mit Hilfe von Prävalenzrelationen. Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung, 32(6): 465-497.

Roy, B., and Jacquet-Lagreze, E. (1977). Concepts and methods in multicriteria decision making. In H. Strobel et al. (Ed.), Optimization Applied to Transportation Systems, CP-77-7. International Institute for Applied Systems Analysis, Laxenburg, Austria.

Saridis, G. (1977). Selforganizing Control of Stochastic Systems. M. Dekker, New York.

Sawaragi, Y., Sunahara, Y., and Nakamizo, T. (1967). Statistical decision theory in adaptive control systems. Academic Press, New York.

Snizek, S. (1985). Wirkungsanalyse als Entscheidungshilfe. Verkehrsannalen, 31 (1): 13-33, and 31 (2/3): 5-28.

Tillman, F.A., Hwang, Ch.-La, and Kuo, W. (1981). Optimization of Systems Reliability. M. Dekker, Basel.

Tou, J.T., and Gonzales, R.C. (1974). Pattern Recognition Principles. Addison-Wesley, Reading, Mass.

Tsypkin, Y.Z. (1973). Foundations of the theory of learning systems. Academic Press, New York.

Wierzbicki, A.P. (1984). Negotiations and mediations in conflicts II: Plural rationality and interactive decision processes. In M. Grauer at al. (Ed.), Plural Rationality and Interactive Decision Processes, Proceedings Sopron 1984. Springer Verl., Berlin.

Willsky, A.S. (1976). A Survey of Design Methods for Failure Detection in Dynamic Systems. Automatica, 12 (6): 601-611.

INTERACTIVE DECISION ANALYSIS FOR STRATEGIC MANAGEMENT

Ivan Stantchev
University of Economics "K.Marx"
Dept. of Systems Analysis and Management
Sofia, Bulgaria

The analysis of the methods and the included activities within the process of strategic management (SM) of the economic enterprises proves that their realization has some common elements connected with the group creative activities for:

- generation of objectives, goals and alternatives,
- fixing of evaluations according to certain qualities and their agregation,
- ranking of various objects and forming of group opinion,
- finding out common relations and interdependences,
- giving general priorities on the basis of individual preferences,
- decision making on the grounds of several criteria,
- choice of alternatives for action with combined assessments (quantitative and qualitative).

These elements can be considered separate modules in the procedures of realization of the SM methods and are generally concerning the process of decision making in the management of economic enterprises. Then a legitimate question arises:

Is it not possible and necessary to work out standard software programmes to help the management, the specialists, the consultants and the experts in the process of SM methods realization?

Of course, the need of creative thinking for decision making remains, nor will the priority of human intellect be ignored, but the arsenal of instruments at the disposal of the decision makers will be increased, the creative process will be structured and directed in a better way and the quality and effect of the taken decisions will be improved.

Before answering this question we shall relate in short the nature of the general problems connected to the automization of decision making.

The strategically important problems and tasks of economy that are to be regularly solved by the managers either lack sufficient amount of information, or do not have well-specified goals and restrictions, or may not have standard procedures or rules for decision making. The existence of these real limits leads to the formation of a definite type of problem situations in management that are called "ill-structured" or "non-structured" and cannot be solved with the help of the "classic" information systems and models.

In the beginning of the 60-ies Herbert Simon willing to classify the types of decisions, divides the process of decision making into three stages:

- discovering and analysis of the problem,
- generation of alternatives (variants of decision),
- choosing of the best alternative (variant) on the basis of definite criteria.

According to his classification, a "programable" or "structured" decision is the type of decision that allows automatic execution of all the three stages, that is, with a computer. In the opposite case the decision is "non-programable" or "non-structured".

The attempts for automization of the process of solving the complex problems of SM are directed mainly towards the third stage: comparison and choice of variants in accordance with given criteria. In order to overcome the difficulties, arising from the multicriterial choice, it become necessary to work out models and normative rules for decision making which postulates the behaviour of the decision makers (DM) taking into account the theories that prescribe him rationality and effectiveness.

The created models and systems for decision making in the end of the 70-ies, however, showed that there are many natural limits in the information obtained by the DM in the process of solving ill-structured problems with the help of normative methods.

DM may err when:

- they compare the differences of utility criteria and alternatives,
- they choose various deciding rules,
- they determine the weight of the criteria, their interdependence and relations,
- they rank and fix the aspiration levels and restrictions of acceptable assessments,
- they define the probability of various events' outcome, etc.

The systematic mistakes are due to overtrust on personal experience and judgement, influence of the primary information, orientation only according to the representativeness of the phenomenon or event without taking into consideration the aprioral probabilities or the size of the sample on the basis of which judgements are made. Analysis prove that DM are extremely inconsistent and intransitive in the direct evaluation of the utility of alternatives and their comparison.

The reasons for making these mistakes are mostly of psychological character and that is why in the last ten years great attention is paid to the problems for defining of the limits of human abilities in the process of decision making. This research brought about the formation of a descriptive decision theory and the creation of descriptive models, describing and analyzing the process and the decisions taken by the DM. In other words, these models interpret and explain the behaviour of the decision making body or person.

While building the descriptive models, however, must be kept in mind that the behaviour of the DM in the process of decision making is quite complicated from psychological point of view - sometimes the person himself is not completely aware of it.

The main source of information for descriptive models design is the examining of the behaviour of highly qualified experts, managers and other DM by means of specialized methods. Part of these methods by character are standard methods for observation and analysis when preparing projects for organizational development. Others use "trees of decision", structured on the basis of statistics and formal logic analysis of the obtained information.

The diagnostic approach for designing of descriptive models is primarily connected with the opposition of the classic concept for a"rational manager" who is capable of"optimizing" of the solution of a certain problem with the concept for "bounded rationality" when the process of decision making is aiming at "satisfying" of certain conditions and limits for finding of a "good enough" decision.

Apart from the new theoretic studies and results in the field of decision making, the development of computer technology caused the universalization of personal computers application and creation of new possibilities for establishing of local nets to allow collective use of data base and peripheral devices and exchange of information among the separate working places.

The automization of the working places within the management systems allowed not only the routine applications in the field of word processing, collection and processing of data in local regime and control of the execution activity on taken decisions and set tasks, but also to start developing more complicated systems, mainly related to the process of decision making support. The individualization of these systems for specific users that began with the application of descriptive models, found its technical solution through the direct contact between the users and the personal computers. The dialogue was started not only with regard to the data base serving a separate working place, but also with regard to the established "banks" of methods and models for solving definite problem situations and tasks appearing in the everyday activity of the specific manager.

The abovementioned conditions determined the formation of a new concept for automization of management activities through working out of systems for decision making support.

The evolution of the idea for developing of decision support systems (DSS) began in the early 70-ies when Anthony Gorry and Scott Morton from MIT (USA) by referring Herbert Simon's ideas to the existing practical needs of strategic planning and management in the economic enterprises, publish the basic principles of DSS construction.

Making use of the experience in the designing of models for decision making and the existing methodologies for constructing of DSS, the Department of Systems Analysis and Management in the University of Economics "K.Marx" worked out a series of software systems to support the DM and the specialists from the economic enterprises in the realization of the standard modules in the composition of the SM methodology. The present paper gives a short account of some of them.

Software system "Generator". The basic purpose of the SS "Generator" is to automize the process of determining the priority group order of objects with regard to a given criterion or criteria on the basis of the priority individual orders that are generated by the separate experts or DM who take part in determining of the strategies of development.

The SS "Generator" aims at achieving the following:

- create possibility for automization of the periodic interaction of a wide circle of specialists who generate and evaluate information necessary for defining and evaluating of the objects' priorities in the process of decision making,
- systemize and "clear out" the subjective evaluations generated through induction by the experts, simultaneously supplying a maximum true group evaluation,
- found subgroups (coalitions) of experts within the common group of experts on the basis of similarity of their individual evaluations, thus allowing the account of all possible variants for solving of the specific problem.

Within the context of the prooess of interactive decision making by means of SS "Generator" the following problems are solved:

- primary listing of objectives and goals,
- detailed specifying of the onjectives and goals,
- listing of the criteria for evaluation of alternatives,
- listing of alternatives,
- liting of future events.

Basic functions of the SS "Generator" in solving the above problems are:

- to define the priority group order (group priority) of the objects on the grounds of the priority individual order,
- to specify the aposteriority competence coefficient of the experts as a result of their participation in an expert evaluation,
- to determine coalitions of experts, i.e. groups of experts who have generated similar evaluations of the object priority,
- to define the experts comprising each coalition, the objects generated by it, the group priorities and probability of these objects, the competence of the coalition.

The restrictions imposed on the application of the SS "Generator" are as follows:

- restriction of the number of experts - maximum 20,
- restriction of the number of objects - maximum number of generated objects by each expert must not exceed 30,
- restriction of the ways used for evaluation of the object priority - two ways are possible: by weights (with figures from 0 to 10) and hierarchical, i.e. by ranking of the place of each object within the group.

Software system "Structure". By means of the SS "Structure" the structure of the relations of influence among the elements of the examined system is built. These relations of influence in the context of the problem can be:

- relations of dependence - within a system of goals and criteria,
- relations of the contribution - within systems of activities,
- cause and effect relations - within systems of events and others.

The strength of a certain direct relation is measured by a previously given scale. Khe SS "Structure" can work with a high scale constructed on the basis of full figures from 0 to 99.

The results from the work of SS "Structure" contain information about:

- the strength of the relations of influence generated by

- the strength of the direct relations of influence,
- compatible and incompatible direct relations of influence,
- forms of re-filling of the non-corresponding relations of influence by the experts,
- matrics of maximum influence,
- chain of maximum influence for structuring a graph of maximum influence,
- number of the objects influencing a given object.

SS "Structure" is meant to work by individual or group DM. The number of experts is unlimited but is practically dependent on the available operative memory. The number of evaluated objects should not be over 35.

SS "Choice". The basic purpose of the SS "Coice" is to automize the process of evaluation of management decisions under the uncertainty condition regarding a number of criteria formulated in natural language.

On the basis of information about the probability and utility of the given consequences, anticipated as a result of the various alternatives, SS "Choice" helps to define the degree at which each of the studied alternatives satisfys each of the criteria for evaluation that are fixed in natural language.

The number of alternatives that are to be evaluated by SS "Choice" is unlimited. The number of the consequences, respective of a given alternative cannot exceed 1000.

REFERENCES

Stantchev, I.(1982). An Approach to Multiobjectives Decision Making with Original Information. In R. Trappl (ed.), Cybernetics and Systems Research, North Holland Publishing Company.

Stantchev, I.(1983). Fuzzy Structural Modelling - an Aid for Decision Making. In P. Humphreys, O.Svenson and A.Vart (eds.) Analysing and Aiding Decision Processes. Akademiai Kiadö, Budapest.

MULTI-OBJECTIVE DECISION ANALYSIS - THE RIGHT OBJECTIVES?

Eberhard E. Bischoff
Department of Management Science and Statistics,
University College of Swansea, Swansea SA2 8PP, U.K.

1. INTRODUCTION

A number of different authors, over a considerable period of time, have pointed out that there is a discernable lack of applications of MODM techniques in practice (cf. Hwang and Masud 1979; Lockett et al. 1984). Whilst at first glance there appears to be a marked increase in reported applications in the recent literature, a closer examination reveals that many of the papers in fact merely put forward generalised models - i.e. *proposals* for applications - or deal with the solution of highly idealised problems. Certainly, there are not a great many papers which describe applications involving a real decision maker and even fewer where the results obtained in the analysis are reported to have actually been implemented.

The dearth of genuine applications may be explained in part by a lack of awareness by practitioners of the theoretical developments which have taken place. From another perspective one may argue that more attention needs to be given to developing appropriate user interfaces. This paper, however, takes the view that the gap between the theory and practice of MODM is much deeper. It argues that the paradigms on which much of the theoretical work has been based adequately reflect only a small part of the spectrum of decision making situations to be found in practice. More specifically, the thesis put forward is that the prevailing paradigms neglect a major aspect of decision making, in that they do not address the need for constructing a reasoned defence for the final decision. It is suggested that the need to be able to justify a decision - possibly against a range of arguments based on several different concepts of rationality - may be of crucial importance even in situations where the responsibility for the chosen course of action lies with a single individual.

It should be pointed out at the outset that the paper raises more questions than it tries to answer. While it is thus not wholly constructive in nature, some attempt is made to identify areas towards which further research efforts should be directed.

The remainder of the paper is divided into four parts. In the next section the existing paradigms of MODM are described and critically examined. This is followed by a more concrete illustration of the shortcomings of many current approaches. The fourth section somewhat widens the issue and questions the practical usefulness, in many situations, of the notion of a decision maker's preference structure. The implications for further research are discussed in the final section.

2. THE TWO PREVAILING PARADIGMS

Most empirical comparisons of MODM approaches have used the decision maker's confidence in the solution as one of the performance measures for the methods examined (cf. Dyer 1973; Wallenius 1975; Rothermel and Schilling 1984). In the theoretical literature, on the other hand, explicit references to the question of how the decision maker may be convinced that the solution obtained is the one he should implement are very rare indeed. Rather, the usual assumption appears to be that confidence in the results is acquired simply through the use of the method as such.

While this assumption may seem quite reasonable at first glance, its appropriateness depends crucially on what suppositions are made about the decision making situation concerned - more specifically, it depends on the concept of what the decision maker requires from the analysis. Two different paradigms can be identified in the literature.[1]

The first assumes that at the start of the analysis the decision maker has already formed opinions which imply a preference structure over the set of alternatives available, but that he is unable to articulate his preferences in a form that allows the best alternative to be readily identified. It then becomes the purpose of the analysis to ellicit sufficient information about his preferences for the optimal solution to be determined. Some approaches based on this paradigm involve the derivation of a utility/value function (see Keeney and Raiffa (1976) for examples), others circumvent the explicit construction of such a function by applying mathematical programming algorithms directly (cf. Geoffrion et al. 1972; Oppenheimer 1978).

In order for the decision maker to build up confidence in the solution ultimately obtained, it is clearly necessary that he is not over-stretched by the information requirements of the method. Also, of course, the process of translating the decision maker's responses into statements about his preference structure must be logically sound and, at least in principle, understood by him. With these premises given, however, no conceptual inconsistencies would appear to arise within this paradigm from looking upon the methods put forward both as means of identifying the optimal decision and of convincing the decision maker that it is indeed the best course of action open to him.

The same does not apply in the case of the second paradigm which is used in the theoretical literature. This, which has to some extent displaced the former in the more recent work on MODM methodologies, allows for a decision maker who embarks on the analysis with an, at least partially, 'open mind' - i.e. who expects the analysis to help him to build up the crucial parts of a preference structure for the problem. A wide range of methods for tackling this task have been suggested. Some - like goal programming (cf. Ignizio 1976) - provide little more than rules for calculating some scalar measure of the attractiveness of an alternative in relation to a set of parameters which have to be specified by the decision maker. Other methods, such as, for instance, STEM (Benayoun et al. 1971) and SEMOPS (Monarchi et al. 1973), are more structured and incorporate procedures for searching through and/or narrowing down the set of alternatives in question.

The advantages and disadvantages of different approaches have been the subject of considerable discussion and there is strong disagreement between some authors about the relative merits of specific methods (cf. Zeleny 1981; Alvord 1983). The fine details of the arguments involved are of no concern in the present context; the point of importance is the acknowledged fact that different approaches, applied to the same problem, will very frequently lead to different solutions. While in the framework of the first paradigm discrepancies between the results obtained by two different approaches must be

---

[1]The reader may wish to compare the two paradigms described here with Starr and Zeleny's (1977) distinction between outcome-oriented and process-oriented approaches, which is similar, but based on a somewhat different vantage point.

interpreted as deficiencies in at least one of the methods, errors in the way they have been applied, or inconsistencies in the information provided by the decision maker, the second paradigm, as it does not assume the pre-existence of a preference structure, does not allow any such conclusions to be drawn. If, however, the preference structure the decision maker hopes to build up in the course of the analysis may be crucially influenced by the approach used, the notion of confidence in a potential solution which is acquired simply through applying one particular method becomes a very dubious one. This clearly implies that the decision maker must wholly believe in the framework of reasoning underlying the method employed and be prepared to shut his mind to other ways of looking at the problem. Such a limited outlook, though, is a complete antithesis of what is usually seen as the role and purpose of decision analysis. Moreover, it would appear doubtful whether many decision makers vested with responsibility for making choices of some importance would want to rely on an analysis which is based on only one type of perspective.

3. AN ILLUSTRATIVE EXAMPLE

A simple example may help to clarify the points made and also facilitate the further discussion. Table I describes a - completely hypothetical - decision problem involving 3 alternatives and 4 objectives. It is assumed that the objectives are of equal importance and that performance against each can be measured on a scale from 0 to 100, with 0 being the worst and 100 the

Table I The alternatives available

	Obj. 1	Obj. 2	Obj. 3	Obj. 4
Altern. 1	90	90	90	90
Altern. 2	95	95	95	78
Altern. 3	89	93	90	90

best. A concrete interpretation of alternatives or objectives is not strictly necessary for the argument being presented, but the reader may wish to think of the alternatives as students who are candidates for some college prize, and of the objectives as marks in four subject areas in which the students have had the same amount of tuition. The problem can then be thought of as that of deciding which student is most worthy of receiving the prize.

One might consider applying the standard goal programming approach to this problem. It would not be unreasonable to use a target value of 100 for each of the objectives and, as the objectives are assumed to be equally important, to define the weighting factors required as 1. This approach would suggest the choice of alternative 2, and the same alternative would also be picked if one lowered the goals to a more realistic level of, say, 97 or 95. However, if the analysis were left there, several important aspects of the problem would pass unnoticed.

For instance, a case can be made for choosing alternative 1, which provides for a more equitable performance against the objectives than alternative 2. In fact, if instead of the usual goal programming approach a 'minmax' formulation[1] were used, alternative 1 would be obtained as the solution. One may also argue, however, that alternative 3 would be a better

[1] The 'minmax' approach employs the l_∞-norm, as opposed to the l_1-norm, to measure the overall deviation from the goals (cf. De Kluyver 1979).

choice than alternative 1, as a shortfall of only 1 on the first objective is more than compensated by a gain of 3 for objective 2. In other words, a scoring approach with non-linear scales could lead to alternative 3 as the solution.[1]

There are limits to what can be proven by a simplified and totally fictitious example and care must be taken not to stretch if too far. The example clearly demonstrates, however, that by applying several MODM approaches it may be possible to construct rationales for a number of different solutions to a decision problem. Moreover, there may be no obvious way of distinguishing categorically between the inherent merit of one rationale as opposed to another.

Conversely, it is possible that through applying a number of MODM approaches several different rationales can be discovered for the same decision. If a fourth alternative with a contribution of 92 for each of the objectives were added to the list in Table I, then both the standard and the minmax goal programming method would produce this solution and the argument for choosing alternative 3 would also no longer apply. An analysis based on more than one methodology may thus serve to allay scepticism about the merits of a particular solution.

4. THE JUSTIFICATION OF A DECISION

The discussion has so far focussed on aspects related to the decision maker's *own* confidence in a certain solution to his problem. The issues raised, however, are of wider relevance; namely in connection with the problem of *convincing others* that the course of action chosen is indeed the most appropriate one. In this regard the criticism concerns the first of the paradigms described above as well as the second.

The necessity to be able to defend a decision does not only arise in situations involving several decision makers, but also frequently presents itself in cases where the decision is the task of a single person. The purchase of a private car has been a popular example in the literature for a problem with a single decision maker. However, while a batchelor buying a car may not have to consult anyone about his decision, his choice is likely to be challenged in discussions with friends or colleagues. Questions may be raised about the accuracy of the information on which his decision was based, the degree of importance placed on certain criteria, or even the principles he used in arriving at a decision. This, of course, does not affect his right to make the decision for himself, but if he is unable to rebut the arguments put forward there is a possibility that he may be swayed by them and therefore regret his choice.

In other cases the need to justify a course of action arises because the decision, although the responsibility of one person, affects other people as well. One such problem in which the author has recently been personally involved - in the role of the decision maker - is the purchase of a computer system for an academic department. This type of decision problem might seem to fall well within the usual MODM framework, but in the given situation the effect which the choice made would have on different individuals and groups in the department meant that the notion of the decision maker's preference structure was not a very useful one for conceptualising the problem. Rather, after initially attempting to approach the problem from this viewpoint, it soon became clear that the most important aspect was that the chosen alternative had to be defensible with respect to a wide range of arguments which

---

[1]A goal programming approach based on the l_2-norm would also produce alternative 3 as the solution.

might be put forward against it. Put differently, the best alternative could be viewed as the one that was least likely to attract criticism which could not be answered. It would appear that the same is true in many other such situations.

If the decision analysis is to provide assistance in building up a defence for the solution ultimately chosen, it is clearly important that the problem is explored from as many different perspectives as possible. This process can be seen as a kind of sensitivity analysis, but rather than being concerned merely with changes in the parameters within a given methodology, it must extend to an investigation of the sensitivity of a solution with respect to the concept of rationality embedded in an approach. The MODM methodologies presently available, however, do not meet the needs of this type of analysis. Indeed, the notion of the defensibility of - as opposed to the preference for - a decision alternative has received little attention in the literature.

5. IMPLICATIONS FOR FURTHER RESEARCH

The statement that existing MODM approaches are inadequate in terms of aiding the development of a reasoned defence for a decision does not necessarily imply that a completely fresh start is required. However, rather than leaving the user to answer the question of which approach to choose for a given problem, attention should be given to the question of how best to integrate methods so as to enable several different principles of reasoning to be applied simultaneously. An example of how such integration may be accomplished for two particular techniques is given in Bischoff (1984), but much more work is needed in this area.

On the other hand, it appears unlikely that the shortcomings described can be completely overcome merely by combining different approaches. For instance, one possible way of tackling a situation where rationales exist for several distinct solutions - such as the one discussed in section 3 - is to seek to differentiate between the solutions by bringing into play criteria which were not included in the initial formulation of the problem. In other words, the analysis may involve phases in which the problem itself is redefined. A decision support tool should be capable of coping with, and indeed providing support for, such phases and further research should be directed towards this aspect.

Lastly, and perhaps most importantly, more descriptive work is needed on the process by which a person acquires confidence in the appropriateness of a certain course of action. Many questions in this area remain unanswered and their resolution is required in order to channel the theoretical efforts into the right direction.

REFERENCES

Alvord, C.H. (1983). The Pros and Cons of Goal Programming: A Reply. Computers and Operations Research, 10(1): 61-62.

Benayoun, R., de Montgolfier, J., Tergny, J., and Larichev, O. (1971). Linear Programming with Multiple Objective Functions: Step Method (STEM). Mathematical Programming, 1(3): 366-375.

Bischoff, E.E. (1984). A-Posteriori Trade-Off Analysis in Reference Point Approaches. In: Grauer, M. and Wierzbicki, A.P. (Eds.), Interactive Decision Analysis. Lecture Notes in Economics and Mathematical Systems, No. 229. Springer-Verlag, Berlin : 139-145.

De Kluyver, C.A. (1979). An Exploration of Various Goal Programming Formulations - with Application to Advertising Media Scheduling. Journal of the Operational Research Society, 30(2): 167-171.

Dyer, J.S. (1973). An Empirical Investigation of a Man-Machine Interactive Approach to the Solution of the Multiple Criteria Problem. In: Cochrane, J.L., and Zeleny, M. (Eds.), Multiple Criteria Decision Making. University of South Carolina Press, Columbia, South Carolina: 202-216.

Geoffrion, A.M., Dyer, J.S. and Feinberg, A. (1972). An Interactive Approach for Multicriterion Optimization with an Application to the Operation of an Academic Department. Management Science, 19(4): 357-368.

Hwang, C.-L. and Masud, A.S.M. (1979). Multiple Objective Decision Making - Methods and Applications. Lecture Notes in Economics and Mathematical Systems, No. 164. Springer-Verlag, Berlin.

Ignizio, J.P. (1976). Goal Programming and Extensions. Lexington Books, Massachussetts.

Keeney, R.L. and Raiffa, H. (1976). Decisions with Multiple Objectives: Preferences and Value Tradeoffs. Wiley Series in Probability and Mathematical Statistics, New York.

Lockett, G., Hetherington, B., and Yallup, P. (1984). Subjective Estimation and its Use in MCDM. In: Haimes, Y.Y. and Chankong, V. (Eds.), Decision Making with Multiple Objectives. Lecture Notes in Economics and Mathematical Systems, No. 242. Springer-Verlag, Berlin: 358-374.

Monarchi, D.E., Kisiel, C.C., and Duckstein, L. (1973). Interactive Multiobjective Programming in Water Resources: A Case Study. Water Resources Research, 9(4): 837-850.

Oppenheimer, K.R. (1978). A Proxy Approach to Multi-Attribute Decision Making. Management Science, 24(6): 675-689.

Rothermel, M.A. and Schilling, D.A. (1984). A Comparative Study of Three Methods of Elliciting Preference Information. Omega, 12(4): 379-389.

Starr, M.K. and Zeleny, M. (1977). MCDM - State and Future of the Arts. In: Starr, M.K. and Zeleny, M. (Eds.), Multiple Criteria Decision Making. Studies in the Management Sciences, Vol. 6. North-Holland, Amsterdam: 5-29.

Wallenius, J. (1975). Comparative Evaluation of Some Interactive Approaches to Multicriterion Optimization. Management Science, 21(12): 1387-1396.

Zeleny, M. (1981). The Pros and Cons of Goal Programming. Computers and Operations Research, 8(4): 357-359.

MAN-MACHINE DIALOGUE INTERFACES IN DECISION SUPPORT SYSTEMS

Andrzej Lewandowski
International Institute for Applied
Systems Analysis, Laxenburg, Austria

1. INTRODUCTION

This paper is based on experiences gained in the development of a specific methodology in decision support systems, called DIDAS (Dynamic Interactive Decision Analysis and Support).

The DIDAS methodology has been developed by an international group of scientists at the System and Decision Sciences Program of the International Institute for Applied Systems Analysis. Several versions of the decision support software were implemented, many applications have been studied and various versions of a basic methodology were developed. These results are now being summarized in a book by main authors from this group (Wierzbicki, at all., forthcoming).

Practically, all the developed DIDAS-based decision support software reached experimental or pilot stage. The main attention was oriented to methodological and computational problems as well as to internal organization of data and information flows within the system. One of important problems was the development of sufficiently robust and effective solvers, of linear and nonlinear programming type. New methods were proposed and implemented - e.g. HYBRID system (Makowski, Sosnowski, 1985) and MSPN specialized module (Kaden, Kreglewski, 1986).

Parallely, experiments with practical applications of the proposed decision support methodology were performed. The DIDAS and related systems were applied experimentally for solving many practical problems - energy modelling, economics, environmental problems. These experiments - beside the conclusion that the DIDAS approach is an adequate tool for solving practical - problems have shown that current implementations are rather far from a professional or commercial level. The reason of this conclusion is connected with the relatively poor quality of the man-machine interface of the existing implementations.

Recently, special attention is given to user-friendliness of the software. This is because of the fact that even the best methodology and its implementation will be rejected by the user, if the input language is too complicated, not adequate to the class of problems solved or does not fit to user´s style of thinking. The same relates to the presentation of results - if the only available output from the computer are hundreds of pages full of numbers, the system will not be accepted by the potential user. Therefore, development of the proper man-machine interface will decide about real applicability of the DIDAS and similar decision support systems.

This paper gives a presentation of the ideas and methodologies which could be applied in new generation of the DIDAS system. Some of them were implemented on experimental basis with positive result. It seems however, that the problem of man-machine dialogue design is far from final solution and in the author's opinion future works relating to the development of new generation of DIDAS systems should be concentrated mainly in this direction.

2. MAN-MACHINE INTERFACES IN DECISION SUPPORT SYSTEMS

As it was stated above, the problem of proper design of man-machine interface decides about applicability of the methodology and the software being the implementation of this methodology.

Analyzing the existing implementations and current experience with DIDAS decision support system, it can be stated that the following aspects of interaction with the computer require special effort from the user:

- definition of the problem to be solved and communicating this definition to the decision support system,

- interacting with the system during interactive session with the decision support system,

- interpreting results of the interactive session with the decision support system.

Evidently, special software tools supporting all these tasks must be provided by the system. This is still an open question, how to design such tools to make them easy to use and adequate to class of problems solved. Some suggestions relating to problem interface exist already (Lewandowski, 1986). It is however obvious, that during all listed above stages of working with the computer, the user must utilize some tools to enter to the computer his wishes relating to the analyzed problem. Similarly, the computer can response to user's decisions. Therefore, the dialogue interface constitutes one of the most important part of the above mentioned interfaces: problem interface, conceptual interface and output interface. This dialogue interface can be considered as the bridge between the user and other parts of the system.

3. "USER FRIENDLINESS"

As it was mentioned in previous sections, this term became recently very popular between software developers. However, it is difficult to give its precise definition. One of the possible definitions was given by Goldberg (Goldberg, 1984):

> "...The problem of creating a friendly programming environment centers on the kind of help the system provides, and the ease with which we can cause the effect, which we wish to cause...Another way to think about the word "friendly" is that it is a measure of the distance between the things the user thinks about doing and the

> things the user actually can do in the system. There are several places in which this measurement can be taken in the programming environment. One is at the interface between the user's conceptualization of the actual world he wishes to represent and the programming language in which the user must describe this world ... Another place for measurements is at the interface between the programming language and the visual presentation of the language to the user. A third place is at the interface between the visual presentation of the language and the way the user must physically indicate what action should take place..."

More deep investigations were performed by Jones (Jones, 1978). He states, that:

> "...many languages have been developed with English words and near English syntax, but they never fulfill their designers claim. A training course, or a big manual beside one of the terminal, always seem to be necessary...the reason for this is that the language is only a covering despite of its sophistication...".

According to Jones' opinion, the existing problems with interaction are caused by the following facts:

- many words with the same meaning,
- innumerable unstated and even unconscious assumptions,
- meaning strongly dependent on immediate context.

After examination of some commonly used style of interaction, Jones concludes, that:

> "... the search for a near English vocabulary will fail to bridge the gap. Instead, we should be seeking to base our dialogues on the procedures and strategies which human adopt in communicating to each other..."

and proposes the principles of dialogue design.

Similar investigations were done by Gaines and Facey (1975). The following rules of dialogue design were suggested by them:

Rule 1 - Introduce through experience: Interactive systems are meant to be experienced, not talked about. Get prospective users onto a terminal on a related, or model system before discussing their expected relationship to their own system.

Rule 2 - Immediate feedback: Give the user feedback by making an immediate unambiguous response to each of his inputs. This should be sufficient to identify the type of activity taking place.

Rule 3 - Use the user's model: Use a model of the activity being undertaken which corresponds to that of the user, and program the interactive dialogue as if it were a conversation between two users mutually accepting this model.

Rule 4 - Consistency and uniformity: Ensure that all terminology and operational techniques are consistently applied, and uniformly available, through all system activities.

Rule 5 - Avoid acausality: Do not introduce apparently acausal phenomena into the system. Make changes in the system clear consequences of the user's actions.

Rule 6 - Querry-in-depth: Distribute information and tutorial material appropriately through the system to be accessed by the user through a simple uniform mechanism.

The rest of proposed rules are more technically oriented (Rule 7 - 11) and can be found in Gaines and Facey paper. In the newer paper Gaines (1981) proposes 17 rules for dialogue design and formulates the guidelines for implementation ("dialogue engineering"). Rather extensive analysis of various (however very similar) concepts of "user friendliness" was performed by Shneiderman (1980). Similarly to Gaines, he proposes a uniform approach to design the interactive dialogue, taking into account the psychology of the future user of the system. Similar, very concise rules and requirements for dialogue design were formulated by Thimbleby (1980).

The detailed revue of existing concepts of "user friendliness" and comparison of various approaches can be found in INFOTECH'S Report (1981). The other source of information relating to these problems is the special issue of the Byte Journal dedicated to "easy software" design (December, 1983).

More analytical approaches to man-machine interface problem are recently in the center of interests. The works of Dehning, Essig and Maas (1981), Ledgard, Singer and Whiteside (1981), Card, Moran and Newell (1983) are evidently the most advanced, however their influence on the programming practice is not sufficient and certain gaps between their work and the practice still exists.

This is not a role of this paper to build such a bridge. It is necessary however, to remind the basic properties of the decision making problem, as well as the assumptions about the decision maker's behavior, which could be influential on the designing the man-machine interface:

- the decision maker is usually not a computer specialist. Because of this it is necessary to avoid the "informatics jargon" in the dialogue. Especially he does not know any programming language, and it is not fair to ask him about files, streams etc. (one of the known systems requires from the user to specify the body of FORTRAN FORMAT statement in order to enter his data!).

- the decision maker is a human and usually makes mistakes. One of the systems (Wendler, 1985) contains in the manual the following statement:

> ".. The user ... should pay attention to the input of data. A mistake in typing data and prompting necessitates reinitialization..".

Other, rather not very rare response of some decision support system for user's mistake is "core dumped" message generated by the operating system and termination of the program, what results in total loss of previously calculated material.

- the decision maker is not a specialist in the field of mathematical modelling and mathematical programming. Therefore, the jargon related to this discipline also should be avoided. The system which displays the following messages (Sakawa, 1980):

"INPUT TOLERANCE DELTA1=00.01
KUHN-TUCKER CONDITION SATISFIED...
LANGRANGIAN MULTIPLIER = 0.21900488E+00.."

evidently cannot be considered as user friendly.

- the decision maker is a human, and the famous "seven plus-minus two" principle (Miller, 1956) applies to him. It means, that the capacity of the "fast access" memory is limited to about 7 elements; it is not reasonable to expect, that he will be able to remember and process the large amount of information. Because of this limited memory capacity, the decision maker is usually not self-consistent and is not able to deduct conclusions from large amount of computer generated information. Therefore, the user friendly system must support data visualization and analysis.

- the decision maker is adaptable and a learning object, therefore the system interface must reflect the adaptation and learning behavior. This relates not only to the learning about the decision process being analyzed, but also to learning the operation and interfacing to the decision support system.

It is necessary to mention, that the above rules can be considered as the strategic rules, in the contrary to the operational rules given by Gaines and Facey. Both set of rules are orthogonal - i.e. for every strategic rule all operational rules are applicable.

The most important conclusion which follows from the analysis of the current works on man-machine interaction and from the analysis and experience with selected decision support systems are the following:

- the man-machine interface must be designed at least as carefully as the rest of the system (solvers, etc.)

- this design should take into account behavior of the future user as well the class of problems for which the system is dedicated.

The statements above can be treated as trivial ones, however the existing experience shows, that most of developers do not treat them with sufficient attention, developing ad-hoc, poor or not working interfaces thus wasting the effort invested in development of the rest of the system.

4. PROBLEM OF ADAPTATION

Special attention should be paid for proper designing the input interface to the system. The difficulties are caused by the following facts:

- the interface should posses certain level of adaptivity, both to user´s requirements and to various problems for which it can be applied.

- the decision support systems must ensure the operational access to its basic components:

  * the data base with initial data for model building and scenario generation,
  * the mathematical model of the system being analyzed,
  * the solver of the decision support system.

Let us consider the specific aspects of the designing of input interface.

Most of the existing dialogue interfaces are very simple menu - driven systems without any sophistication and without any level of adaptivity. One exception, however rather far from final development, is the interface developed for DIDAS-MM system (Lewandowski, at all., 1985).

Similarly, as the other versions of DIDAS, this system is menu driven. It was observed that, during early stages of system utilization, this is very convenient to the user - it is possible to run the system without going through long manuals. However, after cumulating some experience it is rather tiring for the user to enter many times the same sequence of keystrokes. Therefore a "recording" option was built into the system: the sequence of keystrokes can be stored in the file while entered; since this moment this sequence can be invoked by pressing a single key.

Other similar systems, of general purpose and not oriented to decision support systems are so called "keyboard enhancers" for the IBM-PC - the PRO-KEY (RoseSoft, 1984) and SUPER-KEY (Borland, 1985).

The most detailed analysis of the adaptation problem was performed by Edmonds (1981). He identifies three types of adaptation:

- adaptation by a computer specialist,
- adaptation by a trained user,
- adaptation by any user.

Let us consider specific aspects of the dialogue adaptivity.

4.1 Adaptation by a computer specialist

Special attention should be oriented to "adaptation by a computer specialist". Usually, the interface is implemented in a common purpose programming language (FORTRAN in most cases). Such a program is very difficult for further modifications; practically - adding a new command requires essential rewriting and redesigning of the system. In most cases it can be done only by the author of the particular implementation. In order to make the interface really adaptable, certain level of structuralization is necessary.

The following levels of structuralization are possible:

- high - level language programming

A well structured dialogue interface can be implemented in any high level programming language. The structure of the program must be, however, properly designed in order to make it easy to understand and easy for modifications.

The simplest, but powerful example of such structuralization is the POL (Problem Oriented Language) approach proposed by Finger (1982). He developed the program skeleton, written in BASIC; this skeleton consists of the series of procedures (or rather "modules") performing well defined functions, like numbers and strings parsing, error checking etc. Together with the program, the collection of well defined rules for implementing any given input language is defined by the author. One of the interesting features is the "recording" capability - any sequence of commands can be stored on disk and treated as a new command added to the system. Therefore, the system is expandable (or adaptable) on two levels - on the level of implementator and on the level of system user.

A similar approach was proposed by Arciszewski and Van Gastern (1984). They proposed the P/CL system, similar to POL, but implemented in PASCAL. The PC/L is a general purpose input package which can be easily linked to any program intended to be interactive. The anatomy of a similar PASCAL based system was presented by Seidel (1983). He presents a uniform procedure of defining the dialogue using the transition network formalism and algorithm for converting this network into PASCAL program.

The simple and flexible system named DIALOG, was developed by Negus and others (1981). This system, which is the FORTRAN based one, is a collection of routines, including a main "driver" program, which is used by an application programmer as the user interface to interactive application software. The system routines handle command analysis, data input and editing and provide necessary help. This feature, which is especially important, is the flexibility of the selection of the dialogue level depending on the user's experience. This is achieved by supplying the "occasional" user interface, which is mostly question-answer type, and a "regular" one, which is command driven. The user can select the type of interface required on the beginning of the session.

The APL language was applied for this purpose by Bennasat and Wand (1984). Their structural approach to dialogue interface design, utilising programming constructs similar to decision tables, can be easily implemented in any other high level programming language.

- Dialogue generators

These are specialized programs which in their output generate the program, that implements a designed dialogue.

The most advanced system of this type is the SYNICS system (Guest, 1982, Edmonds and Guest, 1978, Edmonds, 1981). The heart of the system is a table driven tree traverser that uses the tree defining the grammar of the dialogue. SYNICS accepts a set of syntax and semantic rules and then translates the input string performing necessary semantics actions. The grammar tree and corresponding semantic routines are defined using the simple input language.

Similar systems were proposed by Kaiser and Stetina (1982), Gerasimov and Polishchuk (1982).

Another dialogue generation system is the Starburst User Interface (Vandor, 1983). This is probably one of the most flexible menu-oriented dialogue generation systems commercially distributed (by MicroPro). Using this system, complex menu-oriented dialogue can be easily implemented. It has a simple programming language build in, what allows rather sophisticated control of a generated dialogue.

The menu-oriented systems MCIS (Menu Creation and Interpretation System) was designed by Heffler (1982). MCIS is written in the C programming language and can be used in the UNIX environment.

The other system from this family is the Karlsruhe Screen-Based Application Support System (Bass, 1985).

- Compiler generators

These are general purpose compiler generators for defining programming language compilers.

Evidently the simplest representative of this approach is the LANG-PAK system, developed by Heindel (1975), recently modified and extended by Sobczyk (1985). This system consists of a collection of FORTRAN procedures and a FORTRAN written interactive program. The dialogue is designed using the BNF (Backus-Naur) form of the grammar. The interactive program analyses the definition and produces the transition table which is used by subroutines performing the syntax analysis of the dialogue commands. The semantics routines must be written by the implementator, also in FORTRAN.

A similar system in structure and in the principle of operation is the YACC, running in the UNIX programming environment and based on C programming language (Johnson, 1978). The LEX system can be used for programming the lexical analysis part of the compiler generated by YACC (Lesk, 1975).

- Special-purpose languages

Some more advanced interactive systems are equipped in specialized languages making modification of the interactive part of the system easy for a programmer or experienced user. The Dbase III (ADL language) and the Framework (FRED language) are examples of such systems. Extensions to general-purpose programming languages were also proposed; one of the known tools of this type is the BASYS - the dialogue oriented extension of BASIC (Gaines and Facey, 1975).

Evidently, the most advanced project relating to development the specialized programming environment oriented to programming interactive systems is the USE project (User Software Engineering, see Wasserman et. all, 1981 for details and references). The basic and most important result of this project is the PLAIN language (Programming Language for Interaction). This language is an extension of PASCAL, equipped in essential capabilities necessary for creating the interactive systems. These capabilities include (Wasserman, 1981):

- Data base management. Usually, one of the basic operation which must be performed by interactive decision support systems is handling large amount of information.

- String handling. Interactive systems usually involve large amounts of text to be processed, especially all the man-machine dialogue.

- <u>Exception handling</u>. User errors must be taken into account. The system should properly react on such errors.

- <u>Pattern specification and matching</u>. This is necessary to recognize commands entered to the system.

Additionally, the system is equipped in other tools simplifying interactive system building - like the TDI (Transition Diagram Interpreter) for programming the interactive dialogues, the MDS (Module Control System) supporting the system maintenance and TROLL, a relational algebra interface to a compact relational database system (Wasserman, 1982).

The USE project is still in progress - such features like multiple window facility and multitasking facility are available in new versions of the PLAIN system (Wasserman, 1985).

The problem-oriented languages make implementation of a dialogue simple and fast, however some programmers prefer to spend a lot of time working with a general purpose language instead of investing some effort into learning of a new tool. It follows from the author´s experience, that the effectiveness of dialogue oriented languages is worth the time spent to learn; especially if the interface must be changed relatively frequently.

<u>4.2 Adaptation by a trained user</u>

In order to allow a trained user (a non-computer specialist but an expert in a substantive field of application) to modify a dialogue, certain extra facilities need to be provided. Such users need to be able to modify any given dialogue within a certain limit, although for major modifications, it would still be necessary to call in a specialist or "system analyst". The facilities provided need to include commands (Edmonds, 1984) such as:

- change text (abbreviate input or output, expand input or output, change keywords etc.)

- change formats of the screen, input commands etc.

- abbreviate a section of dialogue using some internal properties of the system.

The basic assumption of user adaptivity of this type is that the dialogue system should provide the tools necessary for implementing changes within the system without the necessity of full understanding its internal design. Moreover, such changes should not require any low level programming work. Certain knowledge about the implementation of the dialogue might however be required.

The best example of easily user-adaptable system are the dialogue systems based on the POL methodology. In order to modify the structure of dialogue it is enough to modify some text files containing the information about "macro commands", error messages etc. This requires minimal knowledge about internal system organization, and implementation of the changes requires only the access to the text editor.

Another easily user-adaptable system is the StarBurst system. The other examples presented above do not ensure sufficient level of this kind of adaptivity. In order to modify the dialogue rather advanced knowledge about the methods of grammar definition, compiler generation, programming techniques is necessary. Moreover, the semantics routines ensuring proper servicing of the updated dialogue usually must be written in general purpose programming language. An appropriate interfacing and designing these routines requires rather deep understanding of the program structure.

4.3 Adaptation by an average user

The tools oriented to average user adaptation should be especially simple - as it was mentioned above, the average user is typically not a computer specialist. Practically, the only available tool is the macro facility, which is currently being used in practically all interactive software (including such like text editors, spreadsheets etc.) The "keyboard enhancers" are good representatives of such tools. They are simple enough to be used by average user, but simultaneously sufficiently powerful. It is necessary to mention however, that some of them are "oversophisticated" - they are itself rather complicated interactive programs, and even a computer specialist can spend many hours trying to investigate all their possibilities. It is enough to state that one of the manuals of such software tool (Borland´s SUPER-KEY) contains 200 pages. However, the principle of such system can be used to build in similar mechanisms into man-machine interface in decision support systems. The experience with the DIDAS-MM system and its "recording capability" has shown usefulness of this idea.

The other way of adaptation (or rather tuning the dialogue to the user´s experience) can be achieved by supplying different versions of interface for experienced and non-experienced user. This approach is rather popular - most text processors available on microcomputers can be used in command mode or menu mode. The same concept was used in DIALOG system (Negus et. all., 1981). In this system the user can specify explicitely required level of the dialogue.

5. EXISTING IMPLEMENTATIONS

From an analysis of the state of the art of dialogue design, program interacting etc. and the analysis of current available implementations of interactive systems we can conclude that there exist several software products which are very close to the ideas and recommendations presented in previous sections. The systems described in this section should be treated only as the guidelines and samples for the development of specialized interfaces, oriented to application in the decision support system. It is possible, however, to use some of them as the "front-end" to a simple decision support systems.

Dbase III. This is the general purpose data base manager developed and distributed by Ashon-Tate. The most interesting feature of this system is ASSIST option, which makes the system adaptable to the user´s knowledge and ensures selection of the most convenient style of the dialogue.

The system itself is command driven with a rather complicated input language. The syntax of this language is flexible enough for solving complicated data processing tasks; it can however be difficult for a novice or for a user who is not a computer specialist. Therefore, a special assisting mechanism was built into the system. This mechanism is a menu-driven interface, which can be invoked at any moment of the work. The information supplied by the assistant suggests possible movements and explains available options. In the extreme case, it is possible to perform the whole data processing task using the assistant as menu-driven interface to the system. In this way, the user can select most convenient forms of interaction: on very early stages of system usage the menu approach can be used almost entirely, after cumulating some experience - most of the functions can be performed by using command language.

Another feature which ensures sufficient level of adaptivity - but only for a qualified programmer or rather experienced user - is ADL (Application Development Language). The computing power of this language is comparable to PASCAL (it has a similar syntax), but some special features oriented to data processing are available to the programmer. Especially, all Dbase commands are basic elements of this language and can be directly used as program components. Special tools for dialogue programming, such as screen formatting and menu generation are available.

Framework. This is also the Ashon-Tate product, belonging to the class of integrated software. It consists of the data base manager, spreadsheet, word processor and tools supporting the interaction. According to author's knowledge it is one of the best interactive systems available on the market. It possesses also a very high degree of adaptivity - on all levels mentioned above. Here is the list of features supporting adaptation of man-machine interface:

- average user. The macro facility was built into the system. This makes the redefinition of the keyboard very easy. Moreover, the screen formatting is in the disposal of the user. All the information is presented in windows. Size of these windows, their contents as well as location on the screen can be manipulated by very simple keystrokes (4 function keys and 4 cursor keys are engaged in these functions).

- advanced user, or system programmer. The FRED (Frame Editor, see Rubin, 1984) language is one of the tools available for the user. This language is the mixture of features from LISP, PASCAL and C, and similarly as ADL ensures direct access to all features of the Framework. One of the interesting features of FRED is the fact, that a program can modify itself. One of the basic elements of the language is MENU command, which allows very easy menu generation for a given application. The other useful features are graphic oriented commands.

- system programmer. The Framework ensures easy access to DOS and to programs running under DOS. Therefore, some part of the interactive system can be written in any high level language and integrated with the Framework.

Smalltalk. This is a general purpose programming language which in the author's opinion is most suitable for programming of highly interactive systems though it does not posses sufficient popularity in the computer world. This is not commonly known, that in fact it is the predecessor of famous McIntosh and Lisa style of interaction. It is however much more than a window oriented system. The concept of the language itself is very close to SIMULA with respect to the concepts of procedure, class and process. Because of availability of these features very modular systems can easily be developed. The careful analysis of this language from the point of view of "user friendliness" was done by Goldberg (1984).

6. CONCLUSIONS

As it was stated above, the man-machine interface constitutes the integral part of any useful and applicable decision support systems. This component of the software must be designed, implemented and verified with the same care and attention, which is now oriented to the algorithmic part of the system. Therefore more effort should be assigned for recognizing all the problems relating to this subject.

One of the important tasks is accumulating experience with various approaches to man-machine dialogue. Results of such experiments can be frequently directly utilized for designing the interface - see, for example the methodology by Good and others (1984), where results of experiment are used for this purpose. The other techniques for experiment planning and utilization were proposed by Casey and Dasarathy (1982) and Benbasat and Dexter (1981).

Another group of important problems relates to the analysis of existing software tools for dialogue system design as well as existing commercial interactive systems. One of the possible options is the utilization of "threaded programming languages" like FORTH (Loeliger, 1981), or nonprocedural languages especially designed for non computer specialists, like PILOT (Martin, 1982). These languages are easily expandable and could be used efficiently for dialogue programming. Finally, the properties of other interactive parts of the decision support system must be taken into account when designing the dialogue interface and possible software tools for its implementation.

7. REFERENCES

Arciszewski, H.F.R. and Van Gasteren, E.M. (1984). P/CL: A Flexible Input Processor. Software - Practice and Experience, Vol. 14.

Bass, L.J. (1985). A Generalized User Interface for Applications Programs. Communications of the ACM, Vol. 28, June 1985.

Benbasat, I. and Dexter, A.S. (1981). An Experimental Study of the Human/ Computer Interface. Communications of the ACM, Vol. 24, No. 11.

Benbasat, I. and Wand, Y. (1984). A Structured Approach to Designing Human - Computer Dialogues. Int. J. Man-Machine Studies, Vol. 21.

Card, S.K., Moran, P.T, and Newell, A. (1983). The Psychology of Human-Computer Interaction. Lawrence Erlbaum Ass. Publ., Hilsdale, New Jersey, London.

Casey, B.E. and Dasarathy, B. (1982). Modelling and Validating the Man-Machine Interface. Software-Practice and Experience, Vol. 12.

Dehning, W., Essig, H. and Maas, S. (1981). The Adaptation of Virtual Computer Interfaces to User Requirements in Dialogs. Lecture Notes In Computer Science, Vol. 110, Springer Verlag.

Edmonds, E.A. and Guest, S.P. (1978). SYNICS - a FORTRAN Subroutine Package for Translation. Report No. 6, Man-Computer Interaction Research Group, Leicester Polytechnic.

Edmonds, E.A. (1981). Adaptive Man-Computer Interfaces. In: Computing Skills and the User Interface. Coombs M.J. and Alty J.L. eds. Academic Press (Computers and People Series).

Finger, M. (1982). Problem Oriented Language - A New Method of Input. BYTE Journal, BYTE Publications Inc., December 1982 - February 1983.

Gaines, R.B. and Facey, V.P. (1975). Some Experience in Interactive System Development and Application. Proceedings of the IEEE, Vol. 63, No. 6.

Gaines, R.B. (1981). The Technology of Interaction - Dialogue Programming. Int. J. Man-Machine Studies, Vol. 14.

Gerasimov, N.A. and Polishchuk, V.N. (1981). Development of Programming Support for Adaptive Dialogue Systems. Programming Software and System Programming, Plenum Publishing Corporation, 1983 (Translated from Programmirovanie, No. 4, July-Agust, 1982).

Goldberg, A. (1984). The Influence of an Object-Oriented Language on the Programming Environment. In: Interactive Programming Environments, D.R. Barstow, H.E. Shrobe and E. Sandewall, Eds., McGraw-Hill Book Company.

Good, M.D., Whiteside, J.A., Wixon, D.R. and Jones, S.J. (1984). Building a User-Derived Interface. Communications of the ACM, Vol. 27, No. 10.

Guest, S.P. (1982). The Use of Software Tools for Dialogue Design. Int. J. Man-Machine Studies, Vol. 16.

Hefler, M.J. (1982). Description of a Menu Creation and Interpretation System. Software-Practice and Experience, Vol. 12, 1982.

Heindel, L.E. and Roberto, J.T. (1975). LANG-PAK: An Interactive Language Development System. Elsevier, New York.

INFOTECH - State of the Art Report: User Friendly Systems. G. Murray, ed. Series 9, No. 4, 1981.

Johnson, S.C. (1978). YACC: Yet Another Compiler-Compiler. Bell Laboratories, Murray Hill, New Jersey.

Jones, F.P. (1978). Four Principles of Man-Machine Dialogue. Computer-Aided Design, Vol. 10, No. 3.

Kaiser, P. and Stetina, I. (1982). A Dialogue Generator. Software - Practice and Experience, Vol. 12.

Kreglewski, T. and Kaden, S. (1985). Decision Support System MINE. Problem Solver for Nonlinear Multi-Criteria Analysis. International Institute for Applied Systems Analysis, Laxenburg, Austria.

Ledgard, H., Singer, A. and Whiteside, J. (1981). Directions in Human Factors for Interactive Systems. Lecture Notes In Computer Science, Vol. 103, Springer Verlag.

Lesk, M.E. and Schmidt, E. (1975). LEX - A Lexical Generator. Bell Laboratories, Murray Hill, New Jersey.

Lewandowski, A., Kreglewski, T. and Rogowski, T. (1985). DIDAS-MZ and DIDAS-MM: the Trajectory-Oriented Extensions of DIDAS. In: Theory, Software and Test Examples for Decision Support Systems, A. Lewandowski and A. Wierzbicki, Eds. International Institute for Applied Systems Analysis, Laxenburg, Austria.

Lewandowski, A. (1986). The Problem Interface for Nonlinear DIDAS. International Institute for Applied Systems Analysis, Laxenbyrg, Austria, to appear.

Loeliger, R.G. (1981). Threaded Interpretive Languages. The BYTE Books.

Makowski, M. and Sosnowski, J. (1985). HYBRID 2.1 - a Mathematical Programming Package for Multiple Criteria Dynamic Linear Programming Problems. In: Theory, Software and Test Examples for Decision Support Systems, A. Lewandowski and A. Wierzbicki, Eds. International Institute for Applied Systems Analysis, Laxenburg, Austria.

Martin, J. (1982). Application Development Without Programmers. Prentice -Hall, Inc.

Miller, G.A. (1956). The Magical Number Seven Plus-Minus Two: Some Limits of our Capacity for Processing Information. Psychological Review, Vol. 63.
ProKey 3.0 - Users Guide. RoseSoft, 1983.

Negus, B., Hunt, M.J. and Prentice, J.A. (1981). DIALOG: A Scheme for the Quick and Effective Production of Interactive Application Software. Software-Practice and Experience, Vol. 11, 1981.

Rubin, D. (1984). FRED - More than Just a Macro Facility Within Framework. Computer Language, December, 1984.

Sakawa, M. (1980). A Computer Program for Multiobjective Decision Making by the Interactive Sequential Proxy Optimization Technique. WP-80-77, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Seidel, S.R. (1983). Scanning on the Fly: An Approach to the User Interface. Computers in Education, Vol. 7, No. 3.

Shneiderman, B. (1980). Software Psychology: Human Factors in Computer and Information Systems. Winthrop Computer Systems Series.

Sobczyk, A. (1985). Selected Problems of Implementing the Simulation Languages. Technical University of Warsaw, Institute of Automatic Control (master thesis, in Polish).

SuperKey - The Owners Handbook. Borland International Inc., 1985.

Thimbleby, H. (1980). Dialogue Determination. Int. J. Man-Machine Studies, Vol. 13, 1980.

Vandor, S. (1983). The Starburst User Interface. The BYTE Journal, Vol. 8, No. 12.

Wendler, K. (1985). A Dialogue System for Solving Linear Multiobjective Optimization Problems (DIALOG) - User's Guide. International Institute for Applied Systems Analysis, Laxenburg, Austria.

Wasserman, A.I. (1979). USE: A Methodology for the Design and Development of Interactive Information Systems. In: Formal Models and Practical Tools for Information Systems Design, H.-J. Schneider, ed. Proceedings of the IFIP TC 8 Working Conference on Formal Models and Practical Tools for Information Systems Design, Oxford, U.K., April 1979. North-Holland Publ. Comp.

Wasserman, A.I., Van de Riet, R.P. and Kersten, M.L. (1981). PLAIN: An Algorithmic Language for Interactive Information System. In: Algorithmic Languages, de Bakker and van Vliet, eds. Proceedings of the International Symposium on Algorithmic Languages, Amsterdam, 26-29 October, 1981. North Holland Publ. Comp.

Wasserman, A.I. and Shewmake, D. (1982). Automating the Development and Evolution of User Dialogue in an Interactive Information System. In: Evolutionary Information Systems, J. Hawgood, ed. Proceedings of the IFIP TC 8 Working Conference on Evolutionary Information Systems, Budapest, Hungary, 1-3 September, 1981. North-Holland Publ. Comp.

Wasserman, A.I. and Shewmake, D. (1985). The Role of Prototypes in the User Software Engineering (USE) Methodology. In: Advances in Human-Computer Interaction, Vol. 1, H.Rex Harston, ed. Ablex Publ. Corp.

Solving Discrete Multiple Criteria Problems by Using Visual Interaction

Pekka Korhonen, Helsinki School of Economics,
Runeberginkatu 14-16, 00100 Helsinki

1. Introduction

In this paper we describe the principles and the use of a new interactive method for solving discrete deterministic multiple criteria problems. A more detailed description of the method is found in Korhonen (1986).

Quite a few approaches to solving discrete multiple criteria problems have been developed: the traditional multiattribute utility theory (see, e.g. Keeney and Raiffa 1976), the analytic hierarchy process (see, e.g. Saaty 1980), and the outranking (see, e.g. Roy 1973), interactive programming (see, e.g. Korhonen, Wallenius and Zionts 1984), fuzzy-set (see, e.g. Yager 1980), hierarchical interactive (see, Korhonen 1985) approaches and some others (see, e.g. Hinloopen, Nijkamp and Rietveld 1983).

Our aim is to design a method for finding the most preferred alternative from among a large set of alternatives and that is easy to use, permits the decision maker to examine any part of the efficient frontier he wishes and makes no assumptions concerning the decision maker's underlying utility function. The criteria are assumed to be quantitative, but the approach can be used also in the case of ordinal criteria. Interactive use of computer graphics plays a central role in the method. The approach is a modification of our visual interactive method (see, Korhonen and Laakso (1985)) for solving general multiple criteria problems.

The main idea in the method is to select a subset from among the set of all efficient alternatives and present it in a visual form to the decision maker for evaluation. The decision maker can consider the values of the criteria of each alternative graphically and numerically. The subset is found by projecting a so-called reference direction that reflects the decision maker's preferences on the set of efficient attainable alternatives.

There are several methods to specify a reference direction. The use of marginal rates of substitution for estimating the gradient of the utility function (see, Geoffrion, Dyer and Feinberg 1972) and the Boundary Point Ranking method (see, Hemming 1976) are two examples of useful techniques. We prefer a simple and convenient alternative to use the decision maker's aspiration levels for this purpose: the vector from the current alternative to the point defined by the decision maker's aspiration levels is used as a reference direction.

During each iteration the decision maker chooses the most preferred alternative from the subset. This alternative is taken as the current alternative for the next iteration. The decision maker specifies new aspiration levels for the criteria and this information is used to determine a new reference direction.

The process is repeated until the decision maker is not able to find any better solution than the current one. If the optimality of the final point is desired to be checked, we have to make some assumptions concerning the utility function. If it is assumed to be quasiconcave, we can use the method by Korhonen, Wallenius, and Zionts (1984) for eliminating alternatives on the basis of the decision maker's sequential choices and thus guarantee the convergence, if necessary.

The reference direction can be projected on the set of efficient points by using an achievement function as suggested by Wierzbicki (1980) in his reference point approach. After the decision maker has specified his aspiration levels for the criteria, we find an efficient solution that maximizes the value of the achievement function. When we apply the achievement function to the reference direction, instead of one point we obtain a set of efficient solutions for the decision maker's evaluation.

Our main interest is to develop an interactive decision aid for the decision maker for evaluating alternatives close to his aspiration levels. In our opinion, a full benefit from an interactive approach can be obtained by interactive utilization of computer graphics; visual representation enables the decision maker to evaluate a large set of available alternatives simultaneously. Only nondominated solutions are presented to the decision maker for evaluation.

Besides the use of a visual aid, our approach has three further desirable features. Firstly, the decision maker is free to examine any efficient alternatives he pleases at any moment, i.e. he is not confined to evaluating alternatives with some special properties, nor is his freedom limited by his earlier behaviour during the interactive process. Secondly, we need no specific assumptions concerning the decision maker's underlying utility function. Thirdly, the decision maker is asked to make very simple evaluations at each iteration. He has only to specify his aspiration levels for criteria and then he is asked to evaluate reasonable alternatives by using an illustrative figure.

This paper consists of four sections. In the next section we present some preliminary considerations. In section 3 the details of our approach and an illustrative example are presented. The final section consists of concluding remarks.

2. Preliminary Considerations

2.1 Basic Definitions and Notation

First we introduce some terms used in this paper. By the term **criterion** we refer to the concept which represents a magnitude that is of special interest to the the decision maker. It is the basis for evaluation. A criterion may be **qualitative** or **quantitative**. A quantitative criterion means that the decision maker is able to present his preferences using some cardinal (interval or ratio) scale. We call a criterion qualitative, if the decision maker can only rank alternatives or express ordinal preferences by stating which of a pair of alternatives he prefers most. The following definitions are reasonable for quantitative criteria, only.

An **aspiration level** is a desired or acceptable level of achievement of a criterion. A criterion in conjunction with an aspiration level is termed a **goal**. For example, if the decision maker wants to attain a profit level of at least $1000 he has established a goal, and $1000 is his aspiration level corresponding to his particular goal.

A **reference direction** is any direction describing a preferable change in the values of the criteria compared with an available alternative. By the term **achievement function** we mean a function of the criteria which maps any given point in the criterion space on the set of efficient points.

Let us next introduce some notation to describe our problem. We assume that there is a single decision maker, a set of n deterministic decision alternatives $\mathbf{a}_i$, $i=1,2,\dots,n$ and p criteria, which define an n × p decision matrix whose elements are denoted by a_{ij}, $i=1,2,\dots,n$ and $j=1,2,\dots,p$. Thus each decision alternative is a point in the criterion space $\mathbf{R}^p$. The set of alternatives is denoted by A. Without loss of generality we can assume that all $\mathbf{a}_i$ are efficient.

Two cornerstones of the method described in this paper are to extract a subset of all efficient alternatives in a reasonable way and to present it for evaluation of the decision maker using visual interaction.

2.2 Choosing a Subset

We use a reference direction and an achievement function to generate a subset of the efficient alternatives. A reference direction reflects the desire of the decision maker to improve the values of the criteria and an achievement function picks up attainable and efficient alternatives that bear some relation to the decision maker's aspirations.

There are several ways to specify an achievement function. An important feature is that for each efficient point there exists at least one point in R^p which minimizes an achievement function. The characteristics of achievement functions are considered more detailed by Wierzbicki (1985).

We use the following simple form for an achievement function:

$$f(g,a,w) = \max_{i \in I} (g_i - a_i)/w_i, \quad w_i > \text{or} < 0, \qquad (2.1)$$

where $I=\{1,2,...,n\}$, $a \in A$ is a feasible solution in the criterion space R^p, $g \in R^p$ is an arbitrary point (a so-called reference point), and $w \in R^p$ is a weighting vector. If criterion i is maximized $w_i > 0$, otherwise $w_i < 0$. By minimizing $f(g,a,w)$ for any given g and w we find an efficient solution a (see, e.g. Wierzbicki 1980 and 1985).

Given an initial solution $b \in A$, a reference point g and a weighting vector w we can define an achievement function F for the reference direction $d = g - b$ as follows:

$$F(t,d,b,a,w) = f(b + td, a, w), \ t \geq 0. \qquad (2.2)$$

For each feasible solution a we define a set T(a):

$$T(a) = \{t \mid a = \arg \min_{x \in A} F(t,d,b,x,w), \ t \geq 0\} \qquad (2.3)$$

If T(a) is not empty, it has clearly a minimum value:

$$t^*(a) = \min_{t \in T(a)} t \qquad (2.4)$$

For each b and d we define an ordered index set

$$J(b,d) = \{j_1, j_2, ..., j_k\} \qquad (2.5)$$

in such a way that $j \in J(b,d)$, iff $t^*(a_j)$ exists and for all $k \in J(b,d)$, $k < j$, $t^*(a_k) < t^*(a_j)$.

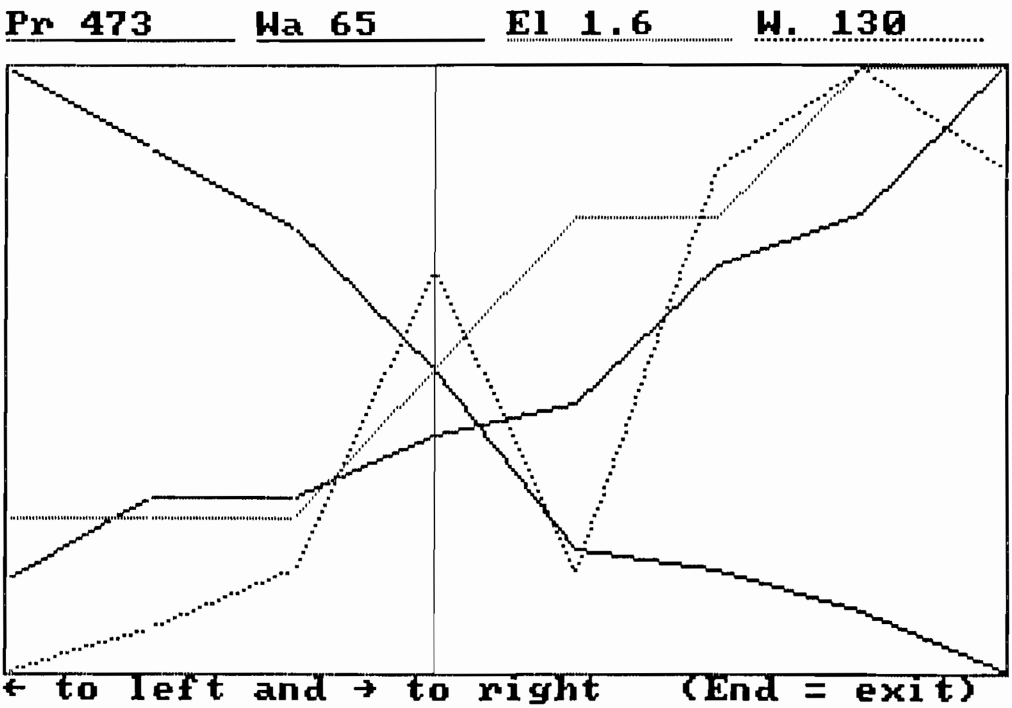
If b is a current solution and g is a reference point the set J(b,d) defines the indices of the alternatives to be presented to the decision maker. In Korhonen (1986) we have described an algorithm for finding alternatives belonging to the set J(b,d).

The index set J(b,d) specifies the subset we use in the method.

2.3 Visual Interaction

The alternatives belonging to the set J(b,d) are presented for evaluation of the decision maker both numerically and graphically. The values of the criteria of the alternatives belonging to J(b,d) are plotted on the screen (on the y-axis). Alternative j_1 (the current alternative) is presented on the left and j_k on the right. The values of the criteria of the subsequent alternatives are connected with a line. Different line patterns or colours can be used for different criteria. The use of colours will enhance the effectiveness of the display. What the graphical display will look like can be seen in Figure 1. The purpose of the display is explained more detailed in the next section

Figure 1. A graphical representation of a subset of alternatives



The cursor (a vertical line) points to the fourth alternative. The numerical values of the criteria for the alternative are shown on the top line of the display. As the cursor is moved back and forth the numerical values of the criteria change, correspondingly.

The scale for each criterion is chosen so that the maximum value is on the top and the minimum value is at the bottom of the display.

This kind of presentation gives the decision maker a possibility to obtain holistic and exact information, simultaneously, on available alternatives in a very convenient way. It seems to be quite easy for the decision maker to choose the most preferred alternative from among the alternatives presented at this display.

3. Development of the Approach

The approach described in this section is a modification of our approach to solving general multiple criteria problems (Korhonen and Laakso 1985). In the discrete case an efficent curve is replaced by a set of efficient alternatives. The approach is implemented on an IBM/PC1 microcomputer under name VIMDA (a Visual Interactive Method for Discrete Alternatives). The dimensions of the problem in VIMDA are: n=500 and p=10. We illustrate the steps of the algorithm in terms of a numerical example using the spreadsheets of the program VIMDA.

Let us assume that we have to help the owner of a laundry to choose a washing machine. Among the criteria, in addition to price, he considers washing time and the consumption of electricity and water to be the most important. He has collected data on four criteria and thirty-three types of machines. The figures are given in Zeleny (1982, pp. 210 - 211). We refer to alternatives with indices.

Step 0. Find an initial (efficient) solution.

Ask the decision maker to specify which of the criteria are to be maximized and minimized (see, Figure 2).

Figure 2. The types of criteria are specified

V I M D A
= = = = =

The types of criteria (max/min):

Criteria	Types
Price	MIN
Wash.tim	MIN
El.consu	MIN
W.consum	MIN

Define weights w_j as follows:

$$w_j = \begin{cases} r_j/\mathrm{abs}(\max_i (r_j * a_{ij})), & \text{if } r_j * a_{ij} <> 0 \\ r_j, & \text{if } r_j * a_{ij} = 0, \end{cases}$$

where r_j is 1, if criterion j is maximized and -1, if it is minimized. (Other definitions are possible, too).

Find an initial alternative by optimizing the value of the last criterion.

In our example all criteria are to be minimized and alternative 2, **b** = (425, 80, 1.5, 110), has the best water consumption (see, Figure 3).

Figure 3. Specification of aspiration levels

V I M D A
= = = = =

The specification of aspiration levels for criteria

The name of a current solution: 2.

Criteria	Lower bounds	Upper bounds	Current values	Aspiration levels
Price	395	595	425	395
Wash.tim	50	80	80	50
El.consu	1.4	1.8	1.5	1.4
W.consum	110	140	110	110

Step 1. Find a reference direction

Ask the decision maker to give aspiration levels for criteria. As we can see from Figure 3, the maximum and minimum values of each criterion are displayed for the decision maker's information.

It seems to be very usual that the decision maker gives ideal values as aspiration levels for criteria at the first iteration. In our example these are (395, 50, 1.4, 110).

The vector from the current alternative to the point defined by the decision maker's aspiration levels is used as a reference direction. In our case

$$\mathbf{d} = (395-425,\ 50-80,\ 1.4-1.5,\ 110-110)$$
$$= (-30,\ -30,\ -0.1,\ -0).$$

Step 2. Find a subset of efficient solutions for the decision maker's evaluation.

In this step we determine the ordered index set $J(\mathbf{b},\mathbf{d})$ as described in section 2.

In our example $J(\mathbf{b},\mathbf{d}) = \{2,21,29,7,8,11,5,26\}$.

Step 3. Find the most preferred solution from among set $J(\mathbf{b},\mathbf{d})$.

The alternatives belonging to set $J(\mathbf{b},\mathbf{d})$ are presented to the decision maker in a visual form as described in section 2.3 and he is asked to indicate which alternative he likes most. The values of the criteria are plotted on the screen using distinct colours and line patterns for criteria. The cursor can be moved from point to point and the corresponding numerical values of criteria are displayed simultaneously.

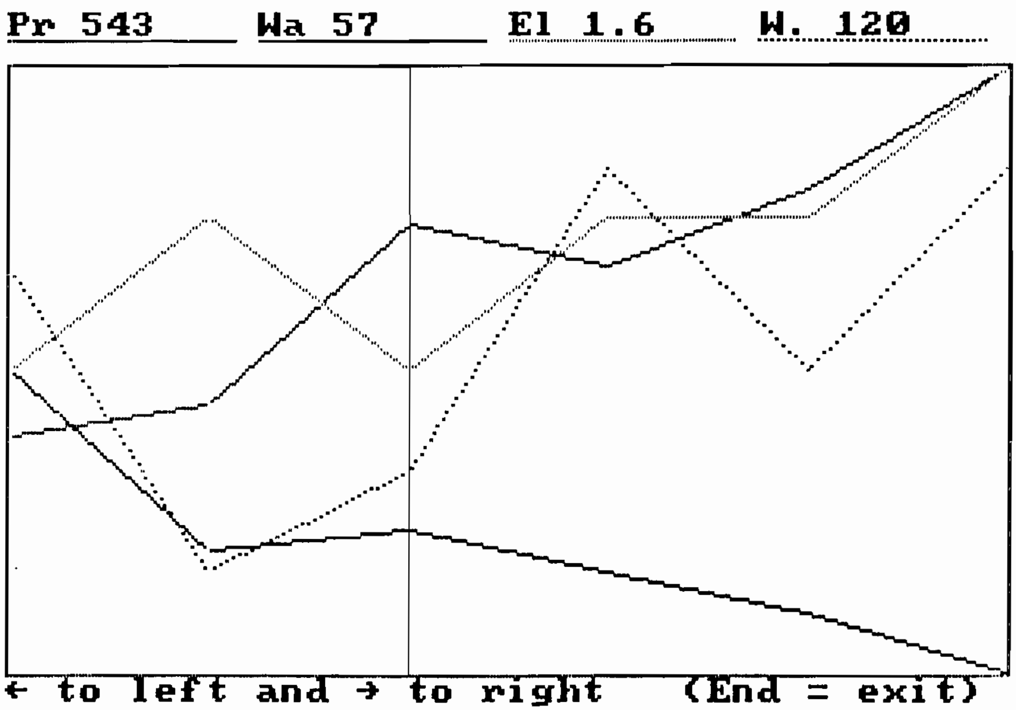
The scale for each criterion j is chosen as explained in section 2.3.

If any improved solution is found in this step or the decision maker is willing to consider other directions, return to step 1, otherwise stop.

The sample display at the first iteration is presented in Figure 1. Let us assume that the decision maker chooses alternative 7 as the most preferred one from among this set. Thus the current alternative for the next iteration is alternative 7.

Let us further assume that the aspiration levels for the next iteration are (450, 60, 1.5, 130). Now the set J(**b**,**d**) = {7,8,19,11,18,26}. The corresponding display is described in Figure 4. If we assume that the washing time is one of the most important criteria, the decision maker might choose alternative 19, **b** =(543, 57, 1.6, 120), as the most preferred alternative from among this set.

Figure 4. The display at the second iteration



If the aspiration levels of the decision maker for the criteria are the same as at the previous iteration the set J(**b**,**d**)= {19,14,8,7,6,3,27,32}. If he still prefer the current solution 19 to the others he may be willing to stop.

In our approach we are not primarily interested to prove the optimality of the final solution. We prefer the satisfaction of the decision maker to proving the optimality, mathematically. If the optimality of the final point is desired to be checked, we have to make some assumptions concerning the utility function. If it is assumed to be quasiconcave, we can use the method by Korhonen, Wallenius, and Zionts (1984) for eliminating alternatives on the basis of the decision maker's sequential choices and thus guarantee convergence. We can utilize all preference information cumulated so far during the process. If more than one alternative remain we can restart the search process using only the remaining alternatives. We have not implemented this feature in our program VIMDA.

4. Concluding Remarks

The approach presented in this paper is not based on unduly restrictive assumptions concerning the decision maker's behaviour, and it is also convenient to use. We neither deal with inefficient alternatives nor impose stringent restrictions on the form of the underlying utility function. The decision maker is free to examine any efficient alternatives he pleases.

We have noticed a very interesting fact when people have used the approach. Some people make cycles when searching the best alternative. There exist some explanations for this phenomenon:

1. If the people have a static utility function, then they have a very imprecise knowledge of their own utility function.

2. The utility function is changing due to learning and "changes of mind" during the interactive process.

3. There is no utility function. The choices of the people depend on the context, in which they are considering alternatives.

Whichever explanation is right, from hence it follows that it is very difficult to find proper rules to restrict the choices of the decision maker on the basis of the assumptions concerning the utility function during the search process. Our objective is to study this phenomenon more carefully in further research.

We have developed our approach assuming the criteria to be quantitative. However, we may also use the approach in the case of ordinal criteria, i.e. when the decision maker is able to only rank alternatives with respect to each criterion. The decision maker can express his aspiration levels in terms of ordinal numbers and the approach tries to find for his evaluation the alternatives reflecting his aspirations. The choice of the subset of efficient alternatives depends, of course, on the weights used, but primarily on the aspiration levels specified by the decision maker. By setting weights in a reasonable way we can help the decision maker to find the alternatives he likes.

Preliminary experiments indicate that the decision makers are interested in the approach. They seem to like colours, figures and spreadsheet interface, and the basic principle of the approach is easy to understand.

References

Geoffrion, A., Dyer, J. and Feinberg, A. (1972), "An Interactive Approach for Multi-criterion optimization, with an application to the operation of an academic department", Management Science, 19, pp.357-368.

Hemming, T. (1978), Multiple Objective Decision Making under Certainty, The Economic Research Institute at the Stockholm School of Economics, Stockholm.

Hinloopen, E., Nijkamp, P. and Rietveld, P. (1983), "The Regime Method: A new Multicriteria Technique", in. P. Hansen (ed.), Essays and Surveys on Multiple Criteria Decision Making, Springer, pp. 146-155.

Keeney, R. and Raiffa, H. (1976), Decisions with Multiple Objectives, Wiley, New York.

Korhonen, P., Wallenius, J. and Zionts, S. (1984), "Solving the Discrete Multiple Criteria Problem Using Convex Cones", Management Science, 30, pp. 1336-1345.

Korhonen, P. (1985), "A Hierarchical Interactive Method for Ranking Alternatives with Multiple Qualitative Criteria", forthcoming in European Journal of Operational Research.

Korhonen, P. and Laakso, K. (1985), "A Visual Interactive Method for Solving the Multiple Criteria Problem", forthcoming in European Journal of Operational Research.

Korhonen, P. (1986), "A Visual Reference Direction Approach to Solving Discrete Multiple Criteria Problems", Working Paper, F-134, Helsinki School of Economics.

Roy, B. (1973), "How Outranking Relation Helps Multicriteria Decision Making", in J. Cochran and M. Zeleny (eds.), Multiple Criteria Decision Making, University of South Carolina Press, Columbia, SC, pp. 179-201.

Saaty, T. (1980), The Analytic Hierarchy Process, McGraw-Hill, New York.

Wierzbicki, A. (1980), "The Use of Reference Objectives in Multiobjective Optimization", in G. Fandel and T. Gal (eds.), Multiple Criteria Decision Making, Theory and Application, Springer, New York.

Wierzbicki, A. (1985), "On the Completeness and Constructiveness of Parametric Characterizations to Vector Optimization Problems", Working paper.

Yager, R. (1980), "A New Methodology for Ordinal Multiobjective Decisions Based on Fuzzy Sets", Decision Sciences ,12.

Zeleny, M. (1982), Multiple Criteria Decision Making, McGraw-Hill.

ADVANCED COMPUTER APPLICATIONS FOR LARGE-SCALE SYSTEMS ANALYSIS

Dr. Kurt Fedra
International Institute for Applied Systems Analysis
A-2361 Laxenburg, Austria

1. INTRODUCTION

The abundance of increasingly affordable computing power lends itself to new and demanding applications. In the environmental field, one of the most demanding problem areas is that of environmental systems analysis, subsuming the areas of policy design, planning and management. These areas and their problems are characterized by their multi- and inter-disciplinary nature, as well as the often dominating importance of political and judgemental elements, as opposed to purely technical, scientific problems. Thus, since the classical, formal approaches to technical problem solving are not strictly applicable, and the people involved are not necessarily technically trained experts but will include elected representatives, interest groups, and the general public, new methods of problem solving, or applied systems analysis, and new methods of communicating scientific and technical information have to be developed.

The basic problem is one of man-machine communication, that is, combining the largely numerical domain of scientific evidence and formal models with the necessarily subjective and largely judgemental domain of perception and evaluation in an interactive, attractive, and educational format. A friendly user-interface with emphasis on symbolic and pictorial representation of information can provide access to otherwise difficult-to-use formal methods. The full exploitation of modern operating systems allows the structuring of individual command environments at various levels of user experience and technical competence, adapting the system to the user's experience.

Decision-support systems also have to be integrated into the user's institutional structure, they have to use the appropriate formats and language, and they need to be built into the established decision-making process to be useful – and used – in the day-to-day operations of the individuals and institutions concerned.

The analysis of large-scale socio-technical and environmental systems by necessity involves a strong subjective and value-dominated human element, which defies formal representation in any generally acceptable way. I therefore argue that only the direct involvement of users, in various phases of the analysis, and interactive methods which give the user an appropriate role, can expect widespread acceptance and use. This direct user involvement, in turn, requires new modes of man-machine interaction. An important aspect in designing these interactive methods is the use of advanced computer technology to make formal methods more accessible to non-technical users.

The new paradigm of man-machine interaction is based on personal, i.e., not shared, micro/minicomputers and interactive graphics as the standard means for user interaction. And a new approach to software development calls for customized and problem-specific rather than general-purpose user environments, composed of highly modular systems of *ad-hoc* code, heuristic rather than algorithmic in nature. This aims at a high degree of flexibility, responsive to the adaptive nature of the problem-solving and decision-making structures we feel are essential in coping with a seemingly ever-growing array of environmental problems.

At this point, it seems appropriate to insert a *caveat*: beware of naive technological optimism. Computers alone are not going to solve anything. And in fact, much can be said against their all too intimate involvement in human affairs (Weizenbaum, 1976). However, this expanding technology could provide a common language and framework for the necessary multi-disciplinary cooperation, or stimulus and focus for new approaches to the solution of both old and new problems. Eventually.

2. METHODS FOR INTERACTIVE ANALYSIS

To build computer-assisted methods of environmental analysis right into the planning and decision-making process (which is about as elusive a concept as the mythical 'decision maker'), some specific features are required. Easy access – a terminal or rather microcomputer workstation on every desk – and easy use are as important as reliability and credibility. The often cited user-friendliness is as important as a transparent and understandable function of any such system. User-friendliness implies a style and language (i.e., jargon and symbolism) of the interaction between model user and the model that is familiar and easy to understand. In addition, costs of both the computer hardware and software, and the manpower for operation and maintenance, must be low relative to the perceived benefits. Finally, the computer-assisted procedures must be compatible with the other tools and methods used in the planning- and policy-making process.

With more and cheaper computer power becoming available, more of this power can be used to improve the user interface, communication and representation aspect. More and cheaper computer power is available to create an interface engineered to support human planning and decision-making procedures without the introduction of a rigid and demanding formalism. Whereas, traditionally, interaction with computers was designed to make things easy and straightforward for the machine – at the expense of the human operator – the advent of abundant and dedicated computer power should allow for a reversal of this approach. Given the vastly increased capabilities of modern hardware, one can afford to be wasteful – from the machine's point of view – to make it easier for people to interact with the machine. For the above problems, this requires that the formal methods, or at least the user interface, are cast into the structure and language of the respective institutional framework, as well as the problem context. The basic principle is to organize information to facilitate judgement.

3. MODEL-BASED DECISION SUPPORT

To provide a useful and generally acceptable problem representation for large-scale socio-technical systems, methodological pluralism is a must: any "model", whether it is a simulation model, a computer language, or a knowledge representation paradigm, is by necessity incomplete. It is only valid within a small and often very specialized domain. No single method can cope with the full spectrum of phenomena, or rather points of view, called for by an interdisciplinary and applied science. Therefore,the selected approach for software system design is eclectic as well as pragmatic. We use proven or promising building blocks, and we use available modules where we can find them.

Real world problems are best described as "a mess", and cannot necessarily be represented by, say, linear algebra or Taylor series expansions. Information about the real world cannot always be organized in sortable, sequential files. For the description of real world problems one would have to use many more conditional constructs (with a usually small set of possible conditions) representing rules, experience and expertise, rather than (differentiable) functional relationships. Conditions for these rules as well as the resulting actions will often have qualitative and symbolic rather than quantitative and numeric character. These rules and symbols are basic elements in the artificial intelligence field, e.g., Barr and Feigenbaum (1981, 1982). In most practical problems, tentative empirical knowledge and expertise is not only more readily available, but also more relevant as a direct reflection of the problem. The usual differentiable functional relationship, in contrast, introduces an arbitrary pseudo precision. As a rule, we know neither the exact kind of most functional relationships nor the necessary parameters (e.g., Fedra 1983).

The alternative or rather complementary approach is based on heuristic (Simon and Newell, 1958) and linguistic modeling (Zadeh, 1973). These approaches are represented in the rule- and knowledge-based approaches of artificial intelligence (Barr and Feigenbaum, 1981, 1982). Heuristic methods concentrate on problem solving, especially the mental operations useful in this process. Heuristics are not necessarily algorithms, or effective procedures of computer sciences, but rather rules-of-thumb, applied to problems, especially practical problems. Realizing the dominant role of interpretation and judgement in more comprehensive problems, and the importance of exploratory or educational as opposed to engineering or design applications of models, these models start from the perceptions of a problem rather than from physical, ecological, or economic theory. Unrestricted by algorithmic constraints, building a model starts with an appropriate form of *knowledge representation.*

The methodological pluralism required for any more complex real world problem again implies that multiple representation paradigms are combined in a hybrid knowledge representation system. A knowledge base might therefore consist of term definitions represented as frames, object relationships represented in predicate calculus, and decision heuristics represented in production rules.

Predicate Calculus is appealing because of its general expressive power and well-defined semantics. Formally, a predicate is a statement about an object:

((property_name) (object) (property_value))

A predicate is applied to a specific number of arguments, and has the value of either TRUE or FALSE when applied to specific objects as arguments. In addition to predicates and arguments, predicate calculus supplies *connectives* and *quantifiers.* Examples for connectives are AND, OR, IMPLIES. Quantifiers are FORALL and EXISTS, that add some inferential power to predicate calculus. However, constructs for more complex statements about objects can be very complicated and clumsy.

In *Object-oriented representation* or *frame-based knowledge representation,* the representational objects or frames allow descriptions of some complexity. Objects or classes of objects are represented by frames. Frames are defined as specializations of more general frames, individual objects are represented by *instantiations* of more general frames, and the resulting connections between frames form *taxonomies.* Each object can be a member of one or more classes. A class has attributes of its own, as well as attributes of its members. An object *inherits* the member attributes of the class(es) of which it is a member. The inheritance of attributes is a powerful tool in the partial description of objects, typical for the ill-defined and data-poor situations real-world applications have to deal with.

A third major paradigm of knowledge representation are *production rules (IF - THEN decision rules)*: they are related to predicate calculus. They consist of rules, or *condition-action pairs*: "if this conditions occurs, then do this action". They can easily be understood, but have sufficient expressive power for domain-dependent inference and the description of behavior.

Translating the perception of the problem and the set of definitions and descriptions of interactions into the executable code of a computer model is a problem of ***knowledge engineering***. I have already briefly referred to the role of the user himself in the development of a specific implementation of such a decision supporting system. In addition to the user, who may or may not be an expert on any of the domains the system covers, the expertise of numerous domain experts is required to make such a system work. It is the task of the designer and developer, the knowledge engineer, to provide a framework and structure for the representation of the experts' knowledge. The knowledge engineer must extract this domain-specific knowledge, formulate it in terms of heuristics or rules, declarations, procedure, etc., and incorporate it into the system.

In any more comprehensive system, expert knowledge or rules are intricately merged with more traditional forms of information representation, i.e., data and algorithms or models. The design and development process thus combines "classical", i.e., mass- and energy-conservation models and optimization methods, with elements of heuristic programming, or rule-based expert systems in artificial intelligence, (see, e.g., Davis and Lenat, 1982). One straightforward possibility is using fuzzy sets for the direct and easy coupling of symbolic and numerical elements (Fedra, 1984). Alternatively, languages with advanced data structures suitable for symbolic computing such as LISP can be used for combining numerical and symbolic entities in one compatible framework.

3.1 A Display-oriented User Interface

What was said above about the way models are usually built with a conservative bias no longer required by today's hard and software, holds certainly true for the way people are supposed to interact with computers. The standard user-interface seems to know only about teletypes. This punch-card style of communication, restricted to alpha-numerical formats, geared towards the batch-processing environment of the past, is hardly suitable for a truly interactive approach and dedicated workstations.

Bit-mapped graphics systems with multiple window capabilities – or more than one parallel output device – allow the structuring of complex displays. This can greatly increase the amount of information communicated and at the same time also enhance ease of understanding (Teitelman, 1977; Meyrowitz and Moser, 1981). Mixtures of alpha-numeric, symbolic, and graphical elements, using such familiar backdrops as maps or flow-chart representations of systems, can be very effective; they do, however, require a considerable amount of design effort (Foley and Van Dam, 1982). Consequently, in designing the model representations, the style of the display, or the visual part of the user interface, has to be considered from the start (Figure 1).

3.2 Decision Support and Expert Systems

Underlying the concept of decision support systems in general, and expert systems in particular, is the recognition that there is a class of (decision) problem situations, that are not well understood by the group of people involved. Such problems cannot be properly solved by a single systems analysis effort or a highly structured computerized decision aid (Fick and Sprague, 1980). They are neither unique – so that a one-shot effort would be justified given the problem is big enough – nor do they recur frequently enough in sufficient similarity to subject them to rigid mathematical treatment. They are somewhere in between. Due to the mixture of uncertainty in the scientific aspects of the problem, and the subjective and judgemental elements in its socio-political aspects, there is no wholly objective way to find a best solution.

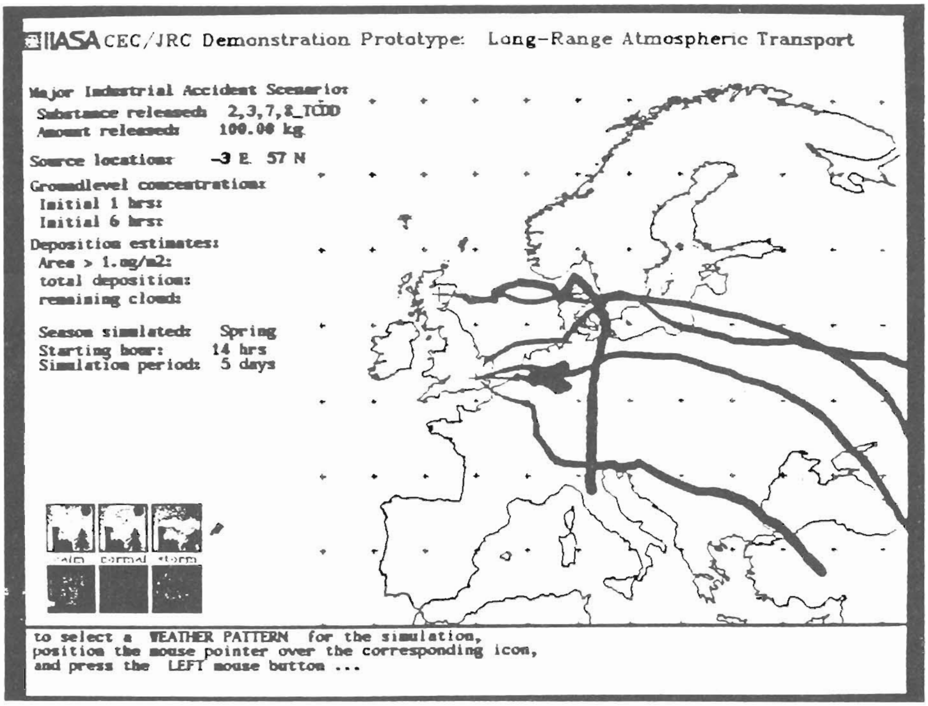


Figure 1a: Specifying weather pattern for atmospheric modeling

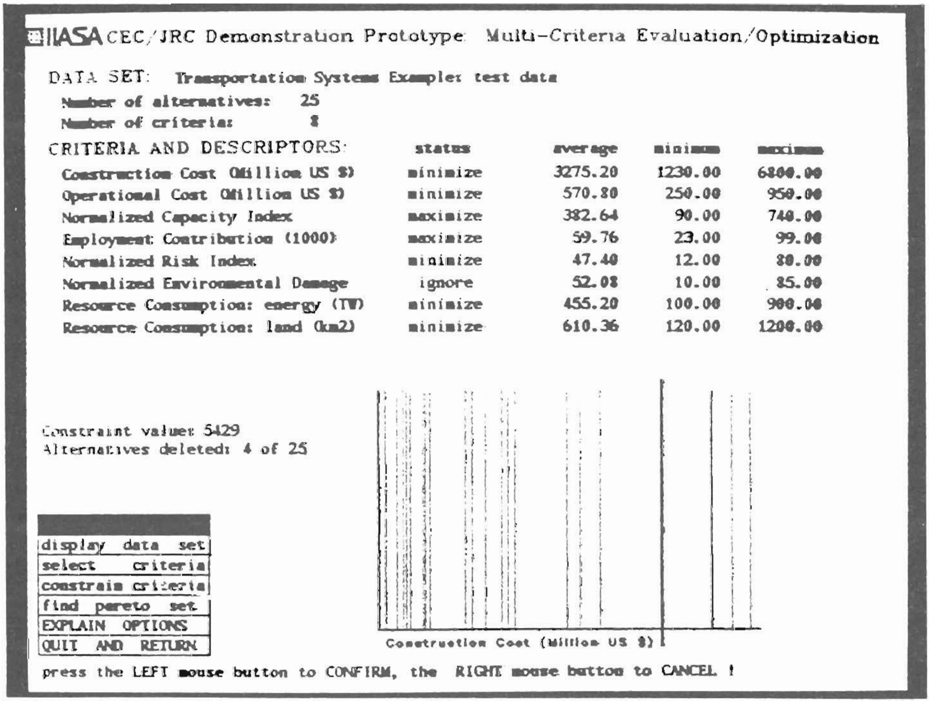


Figure 1b: Multi-criteria data evaluation: constrain criteria

One approach to this class of under-specified problem situations is an iterative sequence of systems analysis and learning generated by expert- or decision-support system use (Figure 2). This should help shape the problem as well as aid in finding solutions. Key ingredients, following Phillips (1984), are the ***Problem Owners***, ***Preference Technology*** (which helps to express value judgements, and formalize time and risk preferences, and tradeoffs amongst them), and ***Information Technology***, (which provides and organizes data, information, and models).

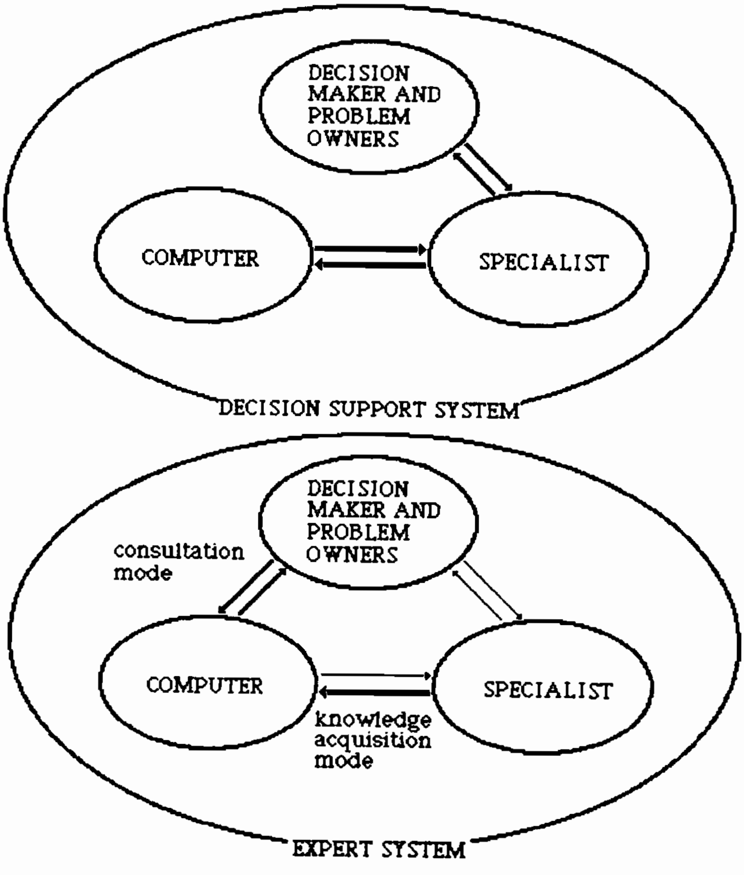


Figure 2: The roles of decision makers, specialists, and the computer: DSS vs the expert systems paradigm.

There is no universally accepted definition of ***decision support systems*** (DSS). Almost any computer-based system, from data base management or information systems via simulation models to mathematical programming or optimization, could support decisions. The literature on information systems and decision support systems is overwhelming (e.g., Radford (1978); Bonczek et al. (1981); Ginzberg et al. (1982); Sol (1983); Grauer et al. (1984); Wierzbicki (1983); Humphreys et al.

(1983); Phillips (1984)). Approaches range from rigidly mathematical treatment, to applied computer sciences, management sciences, or psychology.

Decision-support paradigms include *predictive models*, which give unique answers but with limited accuracy or validity. *Scenario analysis* relaxes the initial assumptions by making them more conditional, but at the same time more dubious. *Normative models* prescribe how things should happen, based on some theory, and generally involve optimization or game theory. Alternatively, *descriptive or behavioral models* supposedly describe things as they are, often with the exploitation of statistical techniques.

Most recent assessments of the field, and in particular those concentrating on more complex, ill-defined, policy-oriented and strategic problem areas, tend to agree on the importance of interactiveness and the direct involvement of the end user. Direct involvement of the user results in new layers of feedback structures (Figure 3). The *information system model* is based on a sequential structure of analysis and decision support (i.e., the relationships shown in the upper part of Figure 3, from Radford, 1978). In comparison, the *decision support model* implies feedbacks from the applications, e.g., communication, negotiation, and bargaining onto the information system, scenario generation, and strategic analysis.

The realism of formal models is increased, for example, by the introduction of *Multiattribute Utility theory* (Keeney and Raiffa, 1976; Bell et al., 1977), extensions including uncertainty and stochastic dominance concepts (e.g., Sage and White, 1984), by multi-objective, multi-criteria optimization methods, and finally by replacing strict optimization, requiring a complete formulation of the problem at the outset, by the concept of satisficing (Wierzbicki, 1983).

Another basic development is getting closer to the users. Interactive models and computer graphics are obvious developments here (e.g., Fedra and Loucks, 1985). Decision conferences (Phillips, 1984) are another approach, useful mainly in the early stages for the clarification of an issue. While certainly interactive in nature, most methods involve a decision analyst as well as a number of specialists (generally supposed to be the problem holders). Concentrating on the formulation of the *decision* problem, *design and evaluation* of alternatives, i.e., the *substantive models*, are only of marginal importance.

Often enough, however, the problem holder (e.g., a regulatory agency) is not specialized in all the component domains of the problem (e.g., industrial engineering, environmental sciences, toxicology, etc., see section 4.2). Expertise in the numerous domains touched upon by the problem situation is therefore as much a bottleneck as the structure of the decision problem. Building human expertise and some degree of intelligent judgement into decision supporting software is one of the major objectives of Artificial Intelligence (AI).

Only recently the area of *expert systems* or *knowledge engineering* has emerged as a road to successful and useful applications of AI techniques. An expert system is a computer program that is supposed to help solve complex real-world problems in particular, specialized domains (Barr and Feigenbaum, 1982). These systems use large bodies of *domain knowledge*, i.e., facts, procedures, rules and models, that human experts have collected or developed and found useful to solve problems in their domains.

Typically, the user interacts with an expert system in a consulting dialogue, just as he would interact with a human expert. Current experimental applications include tasks like chemical and geological data analysis, computer systems configuration, structural engineering, and medical diagnosis (Duda and Gaschnig, 1981; Barr and Feigenbaum, 1981). Expert systems are machine-based intermediaries between human experts (who supply the knowledge in a *knowledge acquisition*

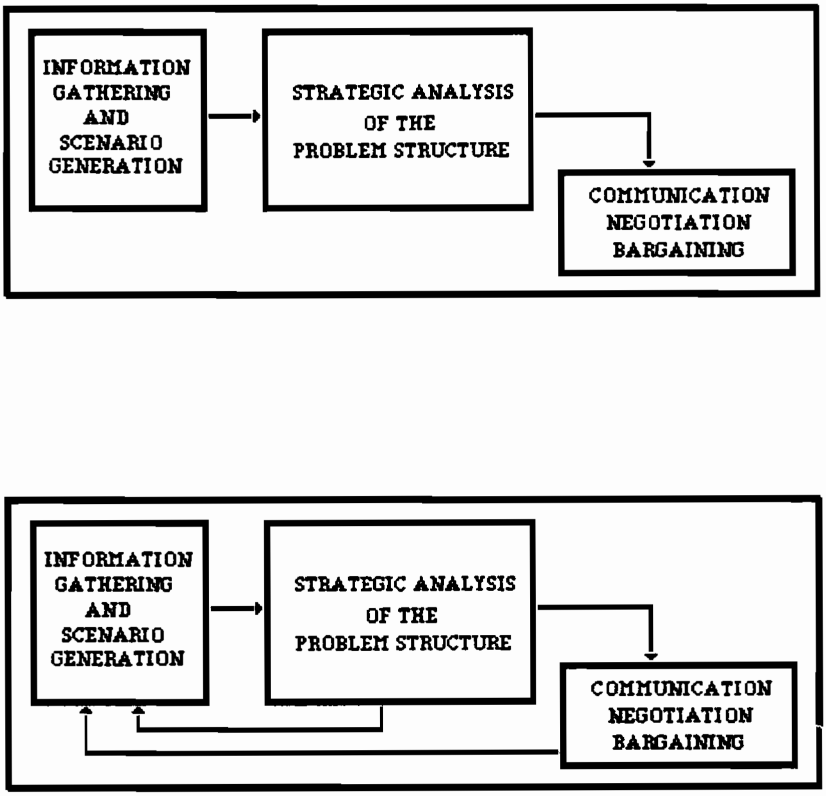


Figure 3: Strategic decision problems: information systems versus DSS approach (partly after Radford, 1978).

mode), and the human user, who seeks consultation and expert advice from the system (*consultation modes*). An important element in the user interface and the dialogue with such systems is their ability to guide the user in formulating his problem, and to *explain* the reasoning used by the system.

The systems described below combine several methods of applied systems analysis, operations research, planning, policy sciences, and artificial intelligence into fully integrated software systems. The basic idea is to provide direct and easy access to these largely formal and complex methods for a broad group of users.

4. DECISION-ORIENTED SOFTWARE FOR THE MANAGEMENT OF HAZARDOUS SUBSTANCES.

Many industrial products and residuals such as hazardous and toxic substances are harmful to the basic life support system of the environment. In order to ensure a sustainable use of the biosphere for present and future generations, it is imperative that these substances are managed in a *safe and systematic* manner. The aim of this project is to provide software tools which can be used by those engaged in the management of the environment, industrial production, products, and waste streams, and hazardous substances and wastes in particular.

The objective of the project is to design and develop an ***integrated set of software tools***, building on existing models and computer-assisted procedures. This set of tools is designed for non-technical users. Its primary purpose is to provide easy access and allow efficient use of methods of analysis and information management which are normally restricted to a small group of technical experts. The use of advanced information and data processing technology should allow a more comprehensive and interdisciplinary view of the management of hazardous substances and industrial risk. Easy access and use, based on modern computer technology, software engineering, and concepts of Artificial Intelligence (AI) now permit a substantial increase in the group of potential users of advanced systems analysis methodology and thus provide a powerful tool in the hand of planners, managers, policy and decision makers and their technical staff.

To facilitate the access to complex computer models for the casual user, and for more experimental and explorative use, it also appears necessary to build much of the accumulated knowledge of the subject areas into the user interface for the models. Thus, the interface will have to incorporate a knowledge-based expert system that is capable of assisting any non-expert user to select, set up, run, and interpret specialized software. By providing a coherent user interface, the interactions between different models, their data bases, and auxiliary software for display and analysis become transparent for the user, and a more experimental, educational style of computer use can be supported. This greatly facilitates the alternative policies and strategies for the management of industrial risk.

4.1 A Structure for the Integrated Software System

The system under design combines several methods of applied systems analysis and operations research, planning and policy sciences, and artificial intelligence into one fully integrated software system (Figure 4). The basic idea is to provide direct and easy access to these largely formal and complex methods for a broad group of users.

Conceptually, the main elements of the system are:

- *an Intelligent User Interface*, which provides easy access to the system. This interface must be attractive, easy to understand and use, error-correcting and self-teaching, and provide the translation between natural language and human style of thinking to the machine level and back. This interface must also provide a largely menu-driven conversational guide to the system's usage (dialog - menu system), and a number of display and report generation styles, including color graphics and linguistic interpretation of numerical data (symbolic/graphical display system);
- *an Information System*, which includes the system's Knowledge and Data Bases (KB, DB) as well as the Inference Machine and Data Base Management Systems (IM, DBMS), which not only summarize application- and implementation-specific information, but also contain the most important and useful domain-specific knowledge. They also provide the information necessary to infer the required input data to run the models of the system and interpret their output. The Inference and Data Base Management Systems (which are at the same time part of the Control Programs and Task Scheduler level) allow a context- and application-oriented use of the knowledge base. These systems should not only enable a wide range of questions to be answered and find the inputs and parameters necessary for the models, but must also be able to explain how certain conclusions were arrived at. For a given application, the data base systems must also perform the more trivial tasks of storing

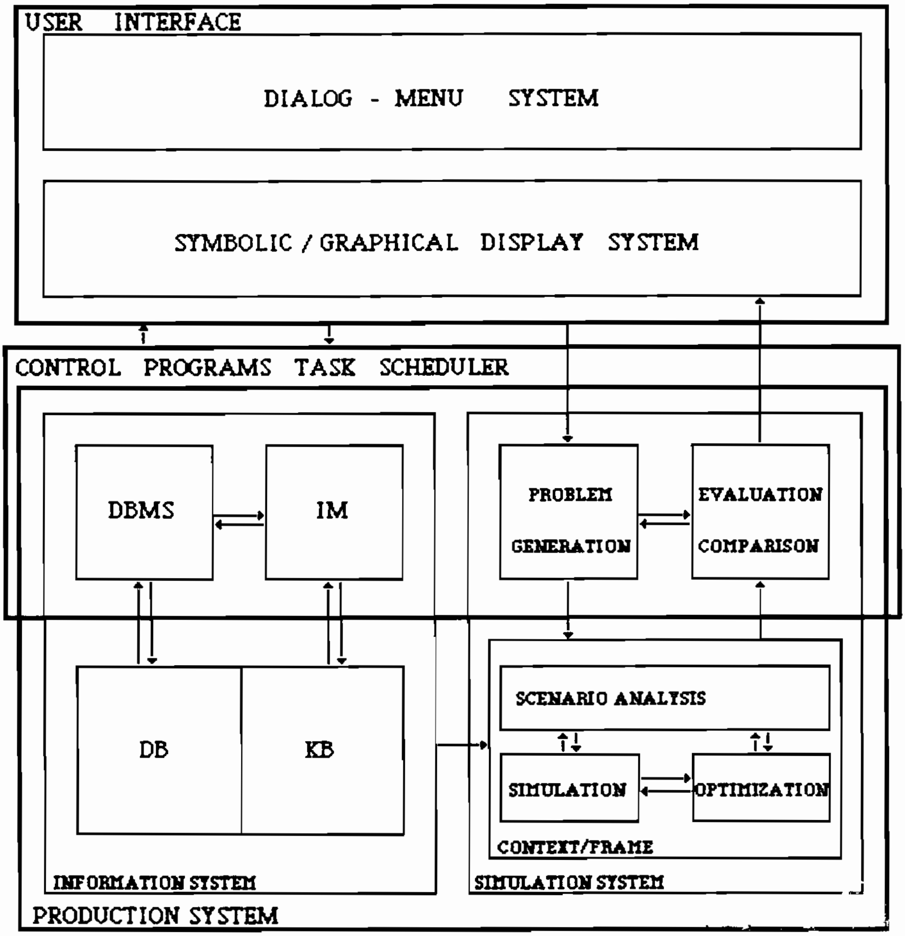


Figure 4: Elements of the integrated software system

and organizing any interim or final results for display and interpretation, comparison, and evaluation;

- *the Simulation System*, which is part of the Production System and consists of a set of models (simulation, optimization), which describe individual processes that are elements of a problem situation, perform risk and sensitivity analyses on the relationship between control and management options and criteria for evaluation, or optimize plans and policies in terms of their control variables, given information about the user's goals and preferences, according to some specified model of the systems' workings and rules for evaluation.

4.2 Components of the Simulation System

The structure and basic elements of the simulation system are shown in Figure 5. The simulation system is always applied to a specific regional context, and the transboundary flows are specified to obtain the necessary material balances.

The system represents a life-cycle approach, that traces substances from their origin and point of release to their impact. For most of these functionally

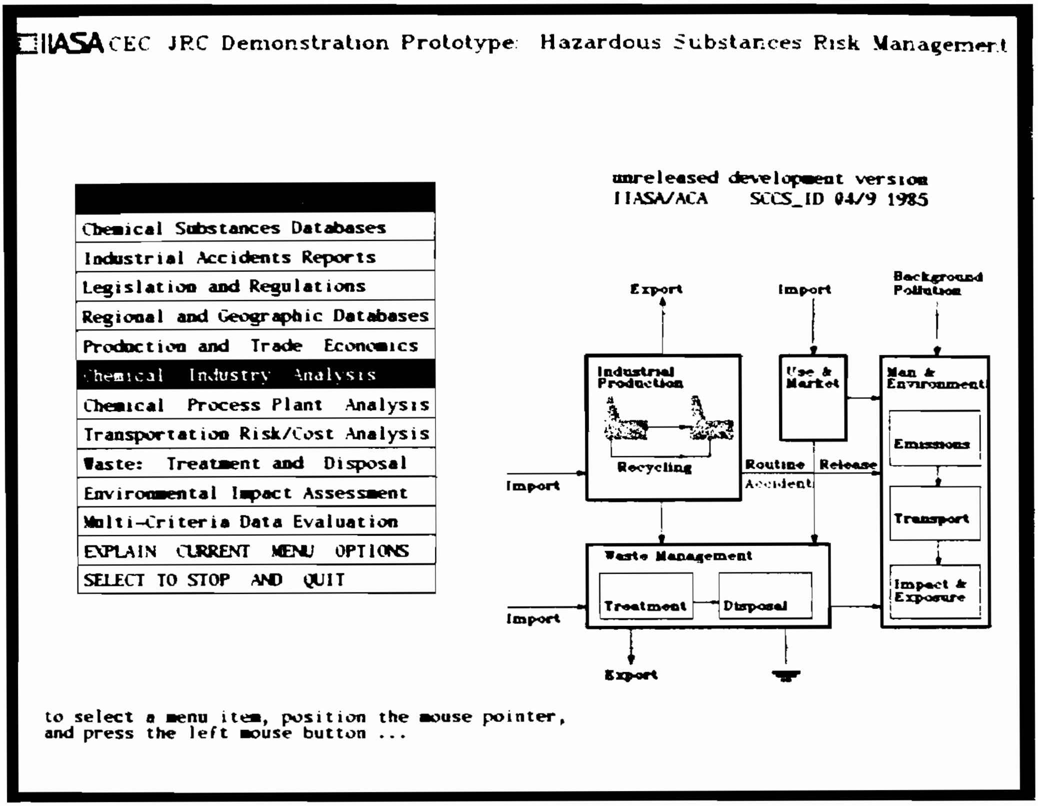


Figure 5: Elements of the simulation system.

specified elements, several models can be used in parallel or alternatively. The selection of the appropriate model(s) depends on the required scope and resolution in time and space, the emphasis on a certain process within a specific problem, and the available data. Wherever possible, the system will select the appropriate model automatically, or switch from one model to another automatically, if, for example, the emphasis changes from a short-term, near-field to a long-term, far-field problem.

The main components of the simulation system are:

1) *The Industrial Production System*, that describes the generation of hazardous substances as products, byproducts, interim products, or wastes of the industrial production process.

2) *Use and Market*, a module that acts as a gateway for the industrial products, diverting them into different pathways according to their use (dispersive or non-dispersive) and waste streams (industrial, domestic). For non-dispersive use, the compartment also serves as interim storage according to the life-time of the product.

3) *Waste Management*; this module simulates treatment and disposal of wastes arriving from either the industrial production or the use/market compartments.

4) *Man and Environment*, a set of models that simulate, starting from the emissions coming from either the industrial production sector, the use compartment, transportation (see below), or the waste management block, the transport of substances through the environment (atmospheric, aquatic, soil, biological pathways), as well as impacts on man and the environment. Since more than 95% of all hazardous waste produced is in liquid form, water resources models play a key role in this system. A detailed description of the models used is given in Fedra (1985b).

5) *Transportation* models interconnecting several of the above blocks. The transportation model estimates costs and risk of various transportation alternatives, and provides input to the emission gateway in the environmental sector.

6) *Evaluation and Interactive Decision-support* is another cross-cutting element that is used for most of the sectoral models. This evaluation comprises monetary as well as non-monetary indicators. For the multi-criteria or selection, a module for discrete optimization is integrated, that permits the selection of preferred alternatives from a set of feasible alternatives generated by any of the models, using the reference point approach (Zhao et al., 1985).

These elements are transparently linked and integrated. Access to this system of models is through a conversational, menu-oriented user interface (compare Figure 1), which employs natural language and symbolic, graphical formats as much as possible.

The system described above can be used in many ways. These modes of operation, however, serve only as design principles. They are transparent for the user, who always interacts in the same manner through the user interface with the system. The system must, however, on request "explain" where a result comes from and how its was derived (e.g., from the data base, inferred by a rule-based production system, or as the result of a model application). The simplest and most straightforward use of the system is as an *interactive information system*. Here the user "browses" through the data and knowledge bases or asks very specific questions.

The alternative mode of use is termed *scenario analysis*. Here the user defines a special situation or scenario (e.g., the release of a certain substance from an industrial plant), and then traces the consequences of this situation through modeling. The system will assist the user in the formulation of these "What if ..." questions, largely by offering menus of options, and ensuring a complete and consistent specification.

All these refinements of the basic information and simulation system however must not complicate the users' interactions with the system. Ease of use, and the possibility to obtain immediate, albeit crude and tentative, answers to problems which the machine helps to formulate in a directly understandable, attractive and pictorial format are seen as the most important features of the system.

5. REFERENCES

Barr, A. and Feigenbaum, E.A. (1981) The Handbook of Artificial Intelligence, Volume I. Pitman, London. 409p.

Barr, A. and Feigenbaum, E.A. (1982) The Handbook of Artificial Intelligence, Volume II. Pitman, London. 428p.

Bell, D.E., Keeney, R.L. and Raiffa, H. [eds.] (1977) Conflicting Objectives in Decisions. International Series on Applied Systems Analysis. John Wiley & Sons. 442p.

Bonczek, R.H., Holsapple, C.W. and Whinston, A.B. (1981) Foundations of Decision Support Systems. Academic Press, New York. 393p.

Davis, R. and Lenat, D.B. (1982) Knowledge-based Systems in Artificial Intelligence. McGraw-Hill Inc. 490p.

Duda, R.O. and Gaschnig, J.G. (1981) Knowledge-based expert systems come of age. BYTE 6:238-281.

Fedra, K. (1983) Environmental Modeling Under Uncertainty: Monte Carlo Simulation. RR-83-28, International Institute for Applied Systems Analysis, IIASA, A-2361 Laxenburg, Austria. 78p.

Fedra, K. (1984) Interactive Water Quality Simulation in a Regional Framework: a management oriented approach to lake and watershed modeling. Ecological Modelling 21 (1983/84) 209-232.

Fedra, K. (1985a) A Modular Interactive Simulation System for Eutrophication and Regional Development. Water Resources Research, 21/2, 143-152.

Fedra, K. (1985b) Advanced Decision-oriented Software for the Management of Hazardous Substances: Structure and Design. IIASA WP-85-18, International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria. 61p.

Fedra, K. and Loucks, D.P. (1985) Interactive Computer Technology for Planning and Policy Modeling. Water Resources Research, 21/2, 114-122.

Fick, G. and Sprague, R.H. jun. [eds.] (1980) Decision Support Systems: Issues and Challenges. Proceedings of an International Task Force Meeting, June 23-25, 1980. IIASA Proceedings Series, Pergamon Press, Oxford. 189p.

Foley, J.D. and Van Dam, A. (1982) Fundamentals of Interactive Computer Graphics. Addison-Wesley, Reading, Massachusetts. 664p.

Ginzberg M.J., Reitman, W.R. and Stohr, E.A. [eds.] (1982) Decision Support Systems, in Proceedings of the NYU Symposium on Decision Support Systems, New York, 21-22 May, 1981. North-Holland, 174p.

Grauer, M., Lewandowski, A. and Wierzbicki, A. (1984) Multiple-Objective Decision Analysis Applied to Chemical Engineering. IIASA RR-84-15, International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria. 40p.

Humphreys, P., Larichev, O.I., Vari, A. and Vecsenyi, J. (1983) Comparative Analysis of Use of Decision Support Systems in R&D Decisions. In: H.G.Sol [ed.], Processes and Tools for Decision Support, Proceedings of the Joint IFIP WG 8.3/IIASA Working Conference. North-Holland, Amsterdam.

Keeney, R.L. and Raiffa, H. (1976) Decisions with Multiple Objectives: Preferences and Values Tradeoffs, New York, Wiley, 1976.

Meyrowitz, N., and Moser, M. (1981) BRUWIN: An adaptable design strategy for Window Manager/Virtual Terminal Systems. Proc. of the 8th Annual Symposium on Operating Systems Principles (SIGOPS), Dec. 1981 Pacific Grove, California.

Phillips, L. (1984) Decision support for managers. In: H. Otway and M. Peltu [eds.], The Managerial Challenge of New Office Technology, Butterworths, London. 246p.

Radford, K.J. (1978) Information Systems for Strategic Decisions. Reston Publishing Co. Inc., VA. 239p.

Sage, A.P. and White, C.C. (1984) ARIADNE: A Knowledge-Based Interactive System for Planning and Decision Support. IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-14, No.1, 35-47.

Sol, H.G. (1983) Processes and Tools for Decision Support. North-Holland, Amsterdam. 259p.

Simon, H.A. and Newell, A. (1958) Heuristic Problem Solving: The Next Advance in Operations Research. Operations Research 6.

Teitelman, W. (1977) A Display-Oriented Programmer's Assistant. Proc. of the 5th International Conference on Artificial Intelligence, 905-915.

Weizenbaum, J. (1976) Computer Power and Human Reason. W.H.Freeman and Co., San Francisco. 300p.

Wierzbicki, A. (1983) A Mathematical Basis for Satisficing Decision Making. Mathematical Modeling USA 3:391-405. (also IIASA RR-83-7).

Zadeh, L. A. (1973) Outline of a New Approach to the Analysis of Complex Systems and Decision Processes. IEEE Transactions, SMC-3,1, 28-44.

Zhao, C., Winkelbauer, L. and Fedra, K. (1985) An Interactive Graphics-based Data Post Processor: An Eclectic Approach to Decision Support. Collaborative Paper, CP-85-xx. International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria. Forthcoming.

ON AN EXPERT SYSTEM FOR THE DESIGN OF CHEMICAL PROCESS SYSTEMS

Klaus Hartmann, Michael Wagenknecht
Technical University "Carl Schorlemmer" Leuna-Merseburg
Merseburg, Otto-Nuschke-Str.
DDR-42, German Democratic Republic

1. INTRODUCTION

Large-scale chemical process systems which produce a great varity of chemical products (more than 75oo organic and several hundred of different anorganic compounds are in commercial production) are built up on a rather small number of different types of equipment (chemical reactors, columns, heat exchangers, pumps etc). The investment which is required for modern large-scale chemical process systems is quite high, the prices for raw material and energy are also increasing. Therefore it seems necessary to synthesize large-scale low-cost ("optimal") process systems.

The synthesis of chemical process system can be defined as an act of determination of the optimal interconnection of processing units (elements) as well as the optimal type and design of the units within a process system (Rudd (1973), Komatsu and Umeda (1973), Hartmann (1974)).

The initial situation of the synthesis can be characterized as follows:

There is a set of available raw material, a set of chemical reactions and physical operations and a set of processing units (apparatus,machines). The task is to select subsets of these sets to satisfy a given set of performance criteria.

The synthesis of chemical process systems is a step-by- step decision process with the following decision layers:

1. Definition of the system objectives
2. Definition of the evaluation criteria
3. Selection of technology which attains the objectives
4. Decomposition of the global systems in interconnected subsystems/tasks
5. Realization of the subsystems/tasks

The problem-solving approach is complicated because of its combinatorial nature. This problem will be illustrated by considering the synthesis of different subsystems.

The process synthesis problem can be decomposed into the following subproblems:

1. Synthesis of reaction path
   e.g.finding the optimal sequence of chemical transformations from available feedstock to given target compound. The num-

ber of alternatives of reaction paths for forming complicated compounds for example a nucleotide sequence (Table 1) is very great (Powers (1973), Govind and Powers (1981)).

TABLE 1 Synthesis of reaction paths-Number of alternative paths for the synthesis of a nucleotide sequence

Sequence length	Number of alternative paths
2	1
3	2
4	5
5	14
6	42
7	132
8	429
9	1.430
10	4.862
20	10^{10}
50	10^{27}
100	10^{50}
500	10^{286}
1.000	10^{575}

2.Synthesis of reactor systems
e.g.to find the optimal reactor types and their interconnection (Hartmann (1974), Kauschus (1979), Hartmann and Kaplick (1985), Chitra and Govind (1985) Anders (1981)).

3.Synthesis of separation schemes
e.g. finding the sequence of separation units that will isolate desired products from given mixtures in the best way. The number of alternatives which can meet the specified performance requirements is also very great (Table 2) (Hendry (1972), King (1971), Tedder (1978), Hacker (1981)).

TABLE 2 Number of alternative structures of distillation sequences

Number of components in the mixture	Number of feasible structures (Main column systems)
2	1
3	2
4	5
5	14
8	429
10	4.862
15	2.674.440
18	$4\ 10^{7}$

4. Synthesis of energy transfer / heat exchanger systems (Masso and Rudd(1969), Nishida, Kobayashi and Ichika (1971), Umeda (1972) , Rockstroh and Hartmann (1975), Linnhoff (1979)) e.g. finding for example the cost minimizing energy recovery network. The number of possible alternatives existing for this task is enormous (Table 3).

TABLE 3 Number of alternative structures in synthesis of heat exchanger networks

Number of hot and and cold streams	Number of feasible structures
4	6
5	720
6	$3.6\ 10^{5}$
7	$4.8\ 10^{8}$
10	$1.5\ 10^{15}$
16	$8.3\ 10^{81}$
20	$2.5\ 10^{101}$
25	10^{101}

Many subtasks / subsystems are characterized by interrelationships between the decision layers and the synthesis procedure is often carried out iteratively.

For a good synthesis procedure it is important to have efficient tools to avoid the complexity of combinatorial problems.

Chemical process systems have to be designed to fullfil different objectives: minimizing investment and operating costs, maximizing reliability and resilience, minimizing air and water pollution etc. To achieve these multiple and conflicting objectives ,it is necessary to make decisions under multiple criteria using methods of decision making and vectoroptimization.

The synthesis stage of basic process design of a chemical process system is characterized by a limited amount of information concerning the objectives, the parameters, the models etc. Besides this lack of information there is a low level of accuracy in the data. There are two types of uncertainty involved in the process design. The major category of uncertainty is due to the fuzziness of the models and parameters of the synthesis procedure.Therefore most of the existing synthesis procedures (Nishida, Stephanopoulos and Westerberg (1981), Umeda (1982)) cannot solve such problems effectively.

Because of this we deal with a simple interactive expert system using the heuristic approach of heat exchanger networks, separation schemes , separation schemes with heat integration and reactor networks.

In the field of chemical engineering there is a large number of rules-of-thumb called heuristic rules which allow the design engineer to generate one or several acceptable system structures quickly.These heuristic rules are based on the experience . of the process and system engineers in designing similar processes and systems. Heuristics are able to reduce the growth rate of the solution tree. These rules will often be in conflict with each other and give contradictory results.In order to resolve these conflicts the heuristics can be combined with learning algorithms.We proposed an another way to resolve this problem which was to combine heuristics with the fuzzy theory approach . (Hartmann (1979, 1981, 1985), Zeising (1982,1984)).

On this basis we developed a small expert system for synthesis of chemical process systems preferable for the stage of process development. This expert system contains quite a sophisticated heuristic device for updating the problem-solving tree under uncertainties of data, models and objectives. This system is shown in figure 1.

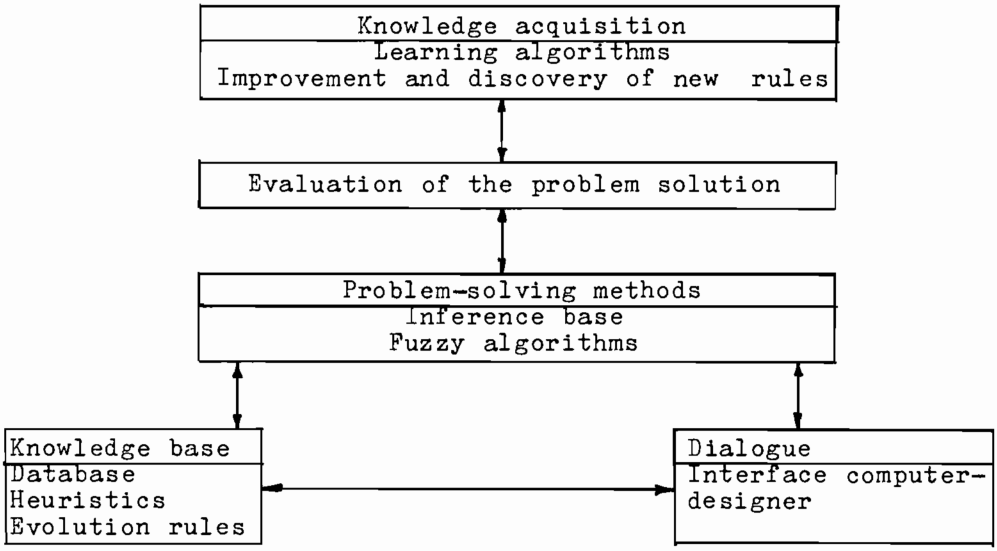


FIGURE 1 Elements of the expert system

The knowledge base is open for any heuristic and evolution rule as well as parameters and weights of membership functions. The designer can easy modify them during the problem-solving process.

The estimation of the objective function(s) which characterize the synthesized system, for example the operating and capital costs, the reliability and so on is carried out during the evaluation procedure.There is also a fuzzy rank-ordering procedure for vectoroptimization problems which is based on pairwise comparison of the importance between the different objective functions (Wagenknecht and Hartmann (1983)).

For the knowledge acquisition we use learning and pattern recognition algorithms. Algorithms for the improvement and discovery of new heuristics are currently being studied.

In the next chapter the main parts of this system are described and finally some results of the application and accumulated experience using this system will be presented.

2. FUZZY ALGORITHMS AND INFERENCE BASE

One of the crucial points in this approach is the transformation of the heuristic rules in the fuzzy theory language and the selection of the best rule.

We assume that the fuzzy heuristic rule is given in the form

$$\text{if } \tilde{A} \text{ then } \tilde{B}$$

where $\tilde{A}$ and $\tilde{B}$ are given fuzzy sets. This fuzzy implication (symbolic expression "$\tilde{A}$ ===> $\tilde{B}$") may be modelled by a fuzzy relation $\tilde{R}$.

For the input $\tilde{A}'$ we have the output

$$\tilde{B}' = \tilde{R} \circ \tilde{A}' \tag{1}$$

where "o" is a suitable composition rule.

If X and Y are basic spaces of the inputs and outputs we can give the equation (1) the form

$$\mu_{\tilde{B}'}(y) = \sup_{x \in X} \min \left(\mu_{\tilde{A}'}(x), \mu_{\tilde{R}}(x, y)\right)$$

where μ is the membership function.

The fuzzy implication "$\tilde{A}$ ===> $\tilde{B}$" e.g.R can be computed as follows

$$\mu_{\tilde{R}}(x, y) = \min \left(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(y)\right) \tag{2}$$

An operator other than the min-operator may be used, but this operator gave good results.

In many cases the heuristics consists of several rules of the type

$$\text{"}\tilde{A}(i) \ ===> \ \tilde{B}(i)\text{"} \ ; \quad i = 1, \ldots, N$$

The rules are connected by the linguistic "or". For this case the relation $\tilde{R}$ can be computed as follows

$$\mu_{\tilde{R}}(x, y) = \max_{1 \leq i \leq N} \min \left(\mu_{\tilde{A}(i)}(x), \mu_{\tilde{B}(i)}(y)\right)$$

Now we describe an important special situation which plays a central role for the application.Let us assume that for the fuzzy inputs it is necessary to realize some instruction or operation.Let $\tilde{A}(i), \ldots, \tilde{A}(N)$ are the fuzzy premise and i(1) ,...,i(N) the instructions which must be executed.If we define the fuzzy sets i(j) in the form

$$\mu_{i(j)}(i) = \begin{cases} 1 & \text{for } i=i(j) \\ 0 & \text{otherwise} \end{cases}$$

we can use the a. m. assumption.

As basic spaces for the i(j) we can choose the set of the natural numbers N (or an suitable subset of them).
R is defined as the Cartesian product X x N as follows

$$\mu_{\tilde{R}}(x,\ i) = \max_{1 \leqslant j \leqslant N} \min\ (\mu_{\tilde{A}(j)}(x),\ \mu_{i(j)}(i)) \tag{3}$$

In technical appplications we often have the case of "crisp" inputs $\tilde{A}'$ and we obtain

$$\mu_{\tilde{A}'}(x) = \begin{cases} 1 & \text{for } x = x(o) \\ 0 & \text{otherwise} \end{cases}$$

and according to the composition rule (1) we obtain for the output $\tilde{B}'$

$$\mu_{\tilde{B}'}(i) = \sup_{x \in X} \min\ (\mu_{\tilde{A}'}(x),\ \mu_{\tilde{R}}(x,\ i) = \mu_{\tilde{R}}(x(o),\ i) \tag{4}$$

If $\mu_{\tilde{B}'}$ achieves its maximum value at $i = i(j^*)$ follows

$$\mu_{\tilde{B}'}(i(j^*)) = \mu_{\tilde{A}(j^*)}(x(o))$$

and for the input x(o) the rule j* is the best one.

If the i(j) are-as in the case of the synthesis of separation schemes-instructions for the separation of a mixture we can construct the following fuzzy algorithm for the synthesis procedure (figure 2).

The algorithm works step-by-step, therefore it is difficult to find a global optimum. But there are rather good "suboptimal" structures in the neighbourhood of this optimum which can be generated and evaluated by our approach.

3. THE FUZZIFICATION OF HEURISTIC RULES

In order to use the described above algorithm it is necessary to present the heuristic rules of the knowledge base as fuzzy implications. We will show this crucial point of the approach by the rules for the synthesis of distillation trains, a special kind of separation scheme.

In general , this synthesis problem can be defined as follows: A multicomponent mixture of given composition and with widely ideal behaviour of the liquid and vapour phases is to separated into certain products of fixed composition (goal products). For the evaluation of these systems we can use for example an objective function like the annual total costs:

$$C_{total} = C_{operating} + C_{capital} = \text{minimum}$$

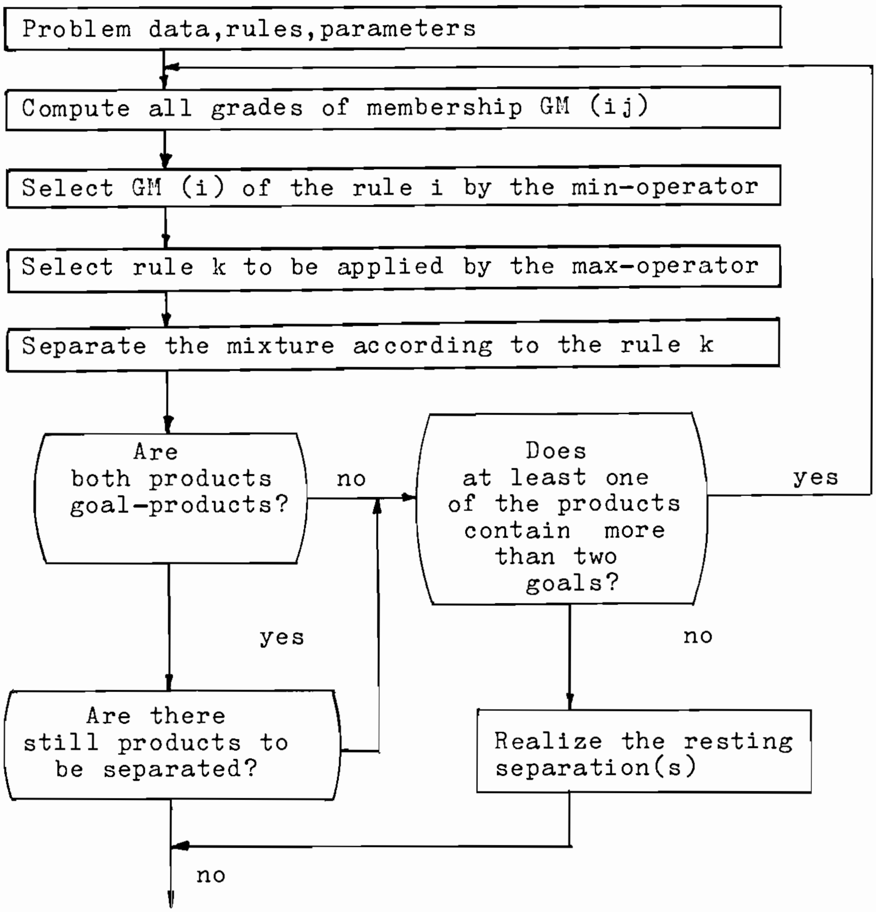


FIGURE 2 Problem-solving algorithm for the synthesis of separation systems

Some of the most common heuristics are:

1. Remove the lightest component first, e.g. remove the component with the lowest boiling point as distillate.
2. Remove the plentiful component first, e.g. the quantitative dominating component in the mixture.
3. Favor 50-50-splits, e. g. separate in such a way,that the molar flowrates of distillate and bottom products are as equal as possible.
4. Save difficult separation until last, e. g. complicated separations are to be realized with the least possible quantities of mixture.
5. Remove corrosive and hazardous components first, e. g. separate such components as soon as possible.

For the synthesis of separation schemes with heat integration we can formulate another important rule

6. Choose the column pressure in such a way as to enable the

heating of some columns with a train by means of the distillate vapours of some others.

It is obvious from the formulation of these heuristics that some rules are contradictory. Each rule implicated a local goal and using the single rules sequently we obtain different structures with local optimum features. Our aim is to formulate such rules which include global goals in the fuzzy implication, e.g. it is necessary to construct multiobjective rules (in some sense).

The first rule ,that is the separation of the lightest boiling component at each stage of the separation process as distillate, shall serve to demonstrate the formation of fuzzy implications. With the help of criteria like concentrations, relative volatilities, boiling temperature differences, demand for special materials for construction and the number of components still contained in the mixture it can be transformed as follows:

(1) Separate as distillate the current lightest boiling component
(a) if the sum of the molar fractions $x(i)$ of the heavier components beginning form $x(3)$ is as low as possible and less than 0.5 and if
(b) the relative volatilities $\alpha(12)$ and $\alpha(23)$ are as high as possible and higher than 2.0 and if
(c) all components requiere the same material for construction as far as possible (influence of the rule (5)) and if
(d) no component, except the first, is present in a dominating quantity and if
(e) the difference of the boiling temperatures is as great as possible and if
(f) the realization of this separation does not violate the rule (4).

As a model of the membership function of the individual criteria we used functions of the type

$$F(x) = 1/(1 + p(1)\, x^{p(2)})$$

or

$$G(x) = 1 - F(x)$$

where $p(1)$, $p(2) > 0$ (figure 3).

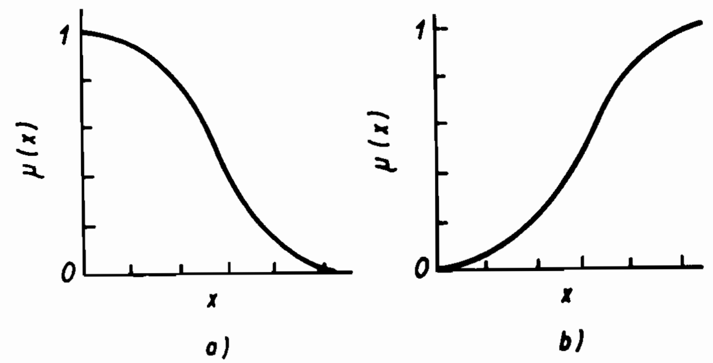


FIGURE 3 Membership functions used for modelling of heuristics a) F(x), b) G(x)

We now show some models of the fuzzy sets of the a.m. rules:

Subrule (a)

For the fuzzy set "as low as possible and less than 0.5" we used the model

$$\mu_a(x) = 1/(1 + 49.353\ x^{7.624})$$

whereby the argument is $\sum_{i=3}^{n} x\ (i)$

Subrule (b)

We used the model of the type G(x)

$$\mu_{b1}(x) = 1 - (1/(1 + 0.005x^{9.638})) \qquad \text{for} \quad \alpha(12)$$

$$\mu_{b2}(x) = 1 - (1/(1 + 0.045x^{7.638})) \qquad \text{for} \quad \alpha(23)$$

whereby $\mu_{b1} \wedge \mu_{b2} = \mu_b$.

Subrule (c)

We evaluate the necessity for using expensive materials for columns due to to corrosive components by 1 and the case of corrosive-free components by 0.If no component is corrosive or only the lightest component which removes according to the rule (1) demands an expensive construction material,the grade of membership is 1, the criteria is met optimally.
In all other cases, the grade of membership function decreases with increasing numbers of corrosive components.The model is as follows:

$$\mu_c(x) = 1 - (1/(1 + 0.0187\ x^{1.6116}))$$

with the argument $\Delta T(12)$.

Subrule (f)

The criterion (f) corresponds to the heuristic rule (4) and states that complicated separations are to be realized with the least possible quantities of mixture.
A separation is recognized to be complicated in this sense if the relative volatility $\alpha(12)$ is lower than 1.5 and still further components are present in the mixture. Also the difference of the boiling temperatures of the key components $\Delta T(12)$ influences the difficulty of this separation. The corresponding membership functions are

$$\mu_{f1}(x) = 1 - (1/ (1 + 0.0393x^{5.8912}))$$

and

$$\mu_{f2}(x) = 1 - (1/ (1 + 0.0187x^{1.6112}))$$

The model of the "last" separation has the membership function

$$\mu_{f3}(x) = 1 / (1 + 10.035x^{1.0532})$$

The argument is (n-2)/N, whereby N represents the number of components in the feed.For n = 2 we have $\mu_{f3} = 1$. This is the case the case of the last separation. The rule can be subdivided into the difficulty of separation and the number of components which are still present in the mixture compared to their original number. Boundary conditions for the fuzzy models are the following ones:

-If the actual separation is the first of an extended separation train it should not be of any difficulty.

-If the actual separation is the last for the multicomponent mixture at hand its difficulty is nearly without importance,because no other separation can be offered at all.

We found that the following membership function is the best one for the subrule (f):

$$\mu_f = [\min (\mu_{f1}, \mu_{f2})]^{1-\mu_{f3}}$$

All other rules are modelled in a similar way.The parameters of these functions are selected with the help of expert knowledge, they are improved by learning. It is worth noting that the synthesized structures are quite robust against the modification of the membership function. Once all grades of membership of the individual criteria are calculated the overall grade of membership of the corresponding rule is determined by the minimum operator, because the rule cannot be evaluated higher than its worst criterion. This pessimistic philosophy is certain to work on the safe side of the self-chosen membership functions.

On the contrary,that competing rule ,which is to be applied, is selected by the maximum-operator. This is , because the rule with the highest grade of membership value represents the best instruction for the actual situation. Weighting exponents assigned to each rule serve to include the insight of the designer again as well as to enable a later learning procedure.

The table 4 shows the step -by-step synthesis procedure for the separation of a mixture consisting of 10 components. For this system we need 9 columns. It can be seen from this table which rule is the best in the different steps.

In the similar way is designed a special system (SYNTHESE),(Zeising (1982)) for the synthesis of heat exchanger systems.Table 5 shows a comparison of the efficiency (annual costs) of systems generated by SYNTHESE with the best known solution from the literature.

TABLE 4 Results of synthesis for the separation schemes of a mixture consisting of 10 components

Column	Number of the lightest component	Grade of membership function of the rule		
		"the lightest first"	" the most plentiful first"	"50-50" split"
1	1	0.095	0	0.234
2	5	0.014	0	0.229
3	2	0.014	0	0.346
4	7	0.088	0	0.219
5	6	-	-	-
6	8	0.031	0	0.175
7	3	0.367	0	0.175
8	9	-	-	-
9	4	-	-	-

"----" no choice (separation end)

TABLE 5 Comparison of heat exchanger systems obtained by SYNTHESE with the best known solution

Number of streams	SYNTHESE	Best known solution
4	19.380	19.571
6	3.695.100	3.695.989
7	70.540	70.284
10	43.984	43.857
20	24.064*	25.293*
25	20.656*	21.982*

* heat exchanger area

Figure 4 shows the heat exchanger network synthesized for a system consisting of 6 cold and 19 hot streams (crude oil distillation).

4. THE SYNTHESIS OF SEPARATION SCHEMES WITH HEAT INTEGRATION (Rathore (1974), Hartmann and Zeising (1981,1982))

After having generated separation schemes (or an another feasible flowsheet) ,the heat duty diagram is suited to giving the designer a general impression of the heat recovery situation.

The streams are ordered in a sequence of decreasing temperature not considering their place in the distillation-train.The following matches are possible without any manipulation of the pressure in the columns (according to the rule(6)):

1)Distillate-bottom matches at atmospheric pressure
2)Distillate-feed matches at atmospheric pressure

By manipulation of the column pressure we can use the following possibility:
3)Distillate-bottom/feed matches at varying pressure.

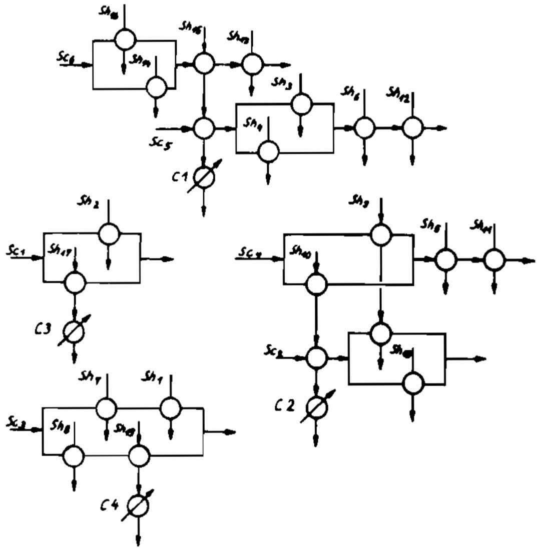


FIGURE 4 Structure of the heat exchanger network for a crude oil distillation

By manipulation of the reflux ratios in the columns we can sometimes also reduce the annual total costs in such systems.Some results for the heat integration of a separation scheme consisting of 6 columns (a feed mixture with 7 components) are shown in the figures 5 and 6 .

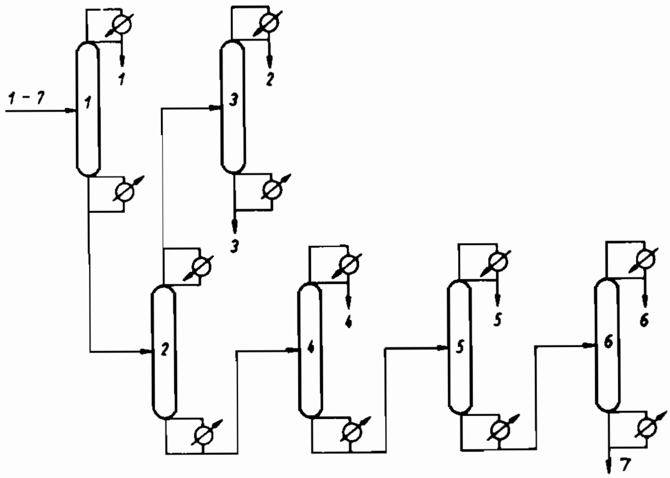


FIGURE 5 Optimal structure of distillation train without heat integration

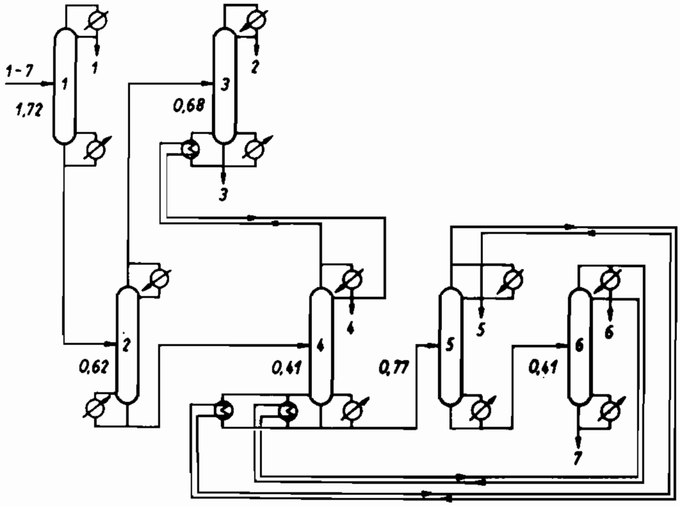


FIGURE 6 Optimal structure of distillation train after heat integration

Table 5 gives an impresssion of the energy saving of the heat integration for two industrial problems.

TABLE 5 Energy saving efficiency by heat integration in distillation trains

Number of components in the mixture	Annual total costs without heat integration	Annual total costs after heat integration	Energy saving %
4	470.231	251.901	58.5
7	940.520	637.397	49.7

5. CONCLUSION

The accumulated experience in using this system for solving industrial tasks shows that this system is a good aid for the designer in the synthesis of new chemical process systems as well as for the improvement of the existing ones.
Furthermore,this approach make it possible to accumulate and to structure knowledge concerning the synthesis procedures in chemical engineering, and makes this knowledge easily accessible for every design-engineer.

6. REFERENCES

Anders, R.(1981).Ein Beitrag zur Struktursynthese heterogener verfahrenstechnischer Systeme. Diss.A.(Ph.D.Thesis).Technische Hochschule "Carl Schorlemmer" Leuna-Merseburg

Chitra, S.P., Govind, R.(1985).Synthesis of Optimal Serial Structure for Homogeneous Reactions. AIChE J. 31, 177-194

Govind, R., Powers, G.J.(1981). Studies in Reaction Path Synthesis.AIChE J.,27, 429-442.

Hacker,I.(1980).Ein Beitrag zur heuristischen Strukturierung von verfahrenstechnischen Systemen .Diss.A.(Ph.D.Thesis) Technische Hochschule "Carl Schorlemmer" Leuna-Merseburg.

Hartmann, K.(1974).Analyse und Synthese verfahrenstechnischer Systeme am Beispiel von Reaktorsystemen .Diss.B (Dr.sc. Thesis).Technische Hochschule "Carl Schorlemmer" Leuna-Merseburg

Hartmann, K.,Dohnal, M.(1979).A Fuzzy Definitional Algorithm and Synthesis Problem. 12.Symposium on Computer Application in Chemical Engineering. Montreux

Hartmann, K.,Kauschus, W.,Dohnal, M., Wagenknecht, M.,Zeising, .G.(1981). Fuzzy Algorithm for Synthesis of Heat Exchanger Networks. CHISA '81,Prague

Hartmann, K.,Zeising, G.(1983).Methoden zur Synthese verfahrenstechnischer Systeme und ihre Anwendung zur Verbesserung der energetischen Guete.Wissenschaftliche Zeitschrift TH Leuna-Merseburg, 25,105-112

Hartmann, K., Kaplick, K.(1985). Analyse und Entwurf chemischnologischer Systeme. Akademie-Verlag, Berlin

Hendry, J.E.,Hughes, R.R. (1972).Generating Separation Process Flowsheets.,Chem.Eng.Prog.69

Kauschus, W.(1979). Strukturoptimierung-eine Methode zur Synthese verfahrenstechnischer Systeme.Diss.B.(Dr.sc.Thesis) Technische Hochschule "Carl Schorlemmer" Leuna-Merseburg.

King, C.J. (1971). Separation Processes. McGraw-Hill Book Co.New York

Komatsu, S.,Umeda, T. (1973).Process Systems Design, Kogyo Chosakai Publ.Co.,Tokyo (Japanese)

Linnhoff, B. (1979).Thermodynamic Analysis of the Design of Process Networks. Ph.D.Thesis, University of Leeds

Masso, A.H.,Rudd, D.F.(1969).The Synthesis of Systems Designs.II.Heuristic Structuring. AIChE J.15,10.

Nishida, N.,Kobayashi,.Ichikawa, A.(1971). Optimal Synthesis of Heat Exchange Systems.Chem.Eng.Sci.,27,1408

Nishida, N.,Stephanopoulos, G., Westerberg, A.W. (1981).A Review of Process Synthesis. AIChE J.,27,321-351

Powers, G.J., Jones, R (1973) .Reaction Path Synthesis Strategy. AIChE J.,19,1204-1214

Rathore, R.N.S.,Van Wormer, K.A.,Powers, G.J.(1974).Synthesis Strategies for Multicomponent Separation Systems with Energy Integration.AIChE J.,20,491.

Rockstroh, L.,Hartmann, K. (1975). Entwurf und Optimierung verfahrenstechnischer Systeme.Chem.Techn.(Leipzig)Teil I 328-332, Teil II 389-392, Teil III 439-442

Rudd, D.F.,Powers, G.J., Siirola, J.J. (1973). Process Synthesis, Prentice Hall, Englewood Cliffs, N.J.

Tedder, D.W.,Rudd, D.F. (1978).Parametric Studies in Industrial Distillation.Part I AIChE J.24, 303, Part II AIChE J.24,

Umeda, T.,Hirai, T., Ichikawa, A.(1972).Synthesis of Optimal Processing System by Integrated Approach.Chem.Eng.Sci. 27,795
Umeda, T.(1982).Computer Aided Process Synthesis. Proceedings of the Symposium "Process Systems Engineering"
Wagenknecht, M.,Hartmann, K.(1983).On Fuzzy Ordering in Polyoptimization.Fuzzy Sets and Systems,11,253-264
Zeising, G.(1982).Struktursynthese verfahrenstechnischer Systeme Scharfe und unscharfe Loesungsmethoden fuer homogene und heterogene Systeme.Diss.A.(Ph.D.Thesis).Technische Hochschule "Carl Schorlemmer"Leuna-Merseburg.
Zeising, G.,Wagenknecht, M.,Hartmann, K.(1984).Synthesis of Distillation Trains with Heat Integration by a Combined Fuzzy and Graphical Approach.,Fuzzy Sets and Systems 12, 103-115

The Object-Oriented Problem Solving Environment ARK-Concept and First Experience

Werner Gabrisch, Manfred Grauer and Lothar Pollmer*

Technical University Leuna-Merseburg
Gausaer Strasse
GDR-42-Merseburg
and
*Academy of Sciences of the GDR
Division of Mathematics and Informatics
Rudower Chaussee 5
GDR-1199-Berlin

1. Introduction

In the field of MCDM (Multiple Criteria Decision Making) we have now the situation that there exist several different conceptual approaches with a similar number of computer implementations (software packages). One of the authors took part in this development by contributing to the DIDASS- methodology (Grauer, Lewandowski and Wierzbicki, 1984). But we feel that now there is not so much a need for extending these existing software packages (supporting decision making processes), nor for writing new and more complex programs. But we should integrate the needed parts from the quantity of good software to a new problem solving environment which could represent a new level of quality in this area. This integration process should have the target to support all stages of a problem solving process like modeling, simulation, optimization and evaluation of solutions. This view is substantiated by the statement of T. Winograd about programming systems (Winograd, 1979): "We get swamped by the complexity of large systems, lost in code written by others, and mystified by the behavior of our almost debugged systems".

The paper presents an attempt to overcome some of the difficulties described above by presenting a concept and first experiences with a problem solving environment. This environment evolved in the field of computer-aided design (modeling, simulation, optimization) of chemical engineering systems. But the experiences gained seem to be valuable also for the analysis of decision processes in nontechnical systems.

2. The problem

In (Winograd, 1979) is the current situation in the field of developing complex programming systems characterized. We will repeat this here because this describes also the starting point for our work.

(I) Computer are not primarily used for solving well-structured mathematical problems or data processing, but instead are components in complex systems.

(II) The building blocks out of which systems are built are not at the level of programming language constructs. They are "subsystems" or "packages", each of which is an integrated collection of data structures, programs, and protocols.

(III) The main activity of programming is not the organization of new independent programs, but in the integration, modification, and explanation of existing ones.

Especially the third statement seems to be important for a next step in creating new systems to support the analysis of decision problems.

We started from the observation that one of the main problems in software development is the communication between independent moduls. The usual way to build up these connection is to rely on the underlying semantical structures. This means that mostly the transfer of data is used to connect software. Going this way it seems to be not possible to solve the problems of integration of complex software systems in the above mentioned sense. To overcome this we should make use of the communication pattern of human beings. They communicate using the semiotic content of an information. Our approach here is mainly based on this idea. Solving this problem we have to implement on the computer a semiotic model of the real world processes under analysis. In this we have to define then a semantic model of all objects we will work with.

The communication between independent software moduls must be based on such an internal model of real processes. In the following this approach will be demonstrated by a simple example: the optimization of a real world object given by its mathematical model. This is presented in Figure 1.

tasks		objects
optimization	–	mathematical model
	–	objective function(s)
	–	bounds
	–	constraints
	–	variables
	–	parameters
modeling	–	set of equations (math. model)
	–	engineering units
	–	mass streams
	–	materials
	–	variables

Figure 1. The problem of optimization of a technical system and the corresponding objects.

The communication between a mathematical model of a technical system and its method of optimization in a problem solving environment must be based on joint elements from the both sets of objects. For the case presented in Figure 1 these are the mathematical model and the variables. Another feature of this kind of description is that it has a hierarchical structure, e.g., each task can be understood as a subtask of an higher ordered one. In the above discussed case modeling and optimization are subtasks and by this objects of the task computational methods of applied mathematics. So we can distinguish between tasks the functioning of which is defined by a software module and on the other hand there are tasks which are defined by their subtasks.

Inputs and outputs of subtasks belong to the internal state of a higher order system which presents a task in itself. The internal state of each task can be described by its set of objects. To indicate the type of objects the notion category is used. This is then the entry point in the hierarchical system of notions. Each of the notions consists of a name and a set of properties, where the properties are represented by a name and a type. The type defines the feasible set of numerical values which correspond to the given property. Further on, let us assume that we call the real world objects (technical or nontechnical systems, processes) realizations. The given definitions are illustrated by the task modeling of a chemical engineering system (heat exchanger) in Figure 2.

Category			
Notion			Realization
property	type		attribute
unit, u			w1
input,i	stream		(str1, str2)
output,o	stream		(str3, str4)
type	heat exchanger		
heat exchanger, hex			
area,a	eng.unit in m^2	100	
K-value	eng.unit in $kW.m^2/{}^oK$	0,37	
column			
. . .			
mixer			
⋮			
stream		str.1	str.2
material		water	air
temperature	eng.unit in ^oK	300	400
pressure	eng.unit in MPa	.1	.2
flow rate	eng.unit in kg/s	10	3
material			
⋮			

Figure 2. Example for the hierarchy of notions for the description of a technical system (heat exchanger).

3. The concept of ARK and its use

In the literature the term programming environment is currently often used with different meanings for different people. In (Blair, Malone and Mariani, 1985) a programming support environment is defined as an all- enveloping system which covers the complete software life cycle-activities such as requirements analysis, design, code, test and installation - with maintenance repeating cycles of the previous steps. This understanding is close to the understanding of an operating system. A somewhat different broader meaning is given in (Barstow, Shrobe and Sandewall, 1984). They understand interactive programming environments as computer-aided design systems for software. On this context we would like to describe a ***problem solving environment*** as a software system which supports decision processes by integration of the modeling, simulation, optimization and evaluation stages. Such a system cannot be built up from scratch but evolves with the growing abilities of computer technology (both software and hardware). Our own experiences are gained from the development of a system for modeling the steady-state behavior of chemical engineering systems. This system was then completed by adding a database management system to it to describe physical and chemical properties of the materials involved. The next step was to add graphic facilities to visualize information. Currently the integration of optimization methods is finished.

Based on the definitions given in the last chapter we will now describe the structure of ARK. The categories were specified by the corresponding tasks, which are characterized by their objects. The order is defined by the prototype through the objects of a task and is independent from a realization of a task. This means that objects can be defined more than one time in different tasks. So, for instance, a technical system (unit process) can be defined on the one hand by its graphical presentation (technological scheme) and on the other hand by its economic behavior (cost function). Figure 3 is a presentation of this hierarchical structure of ARK and his physical capabilities.

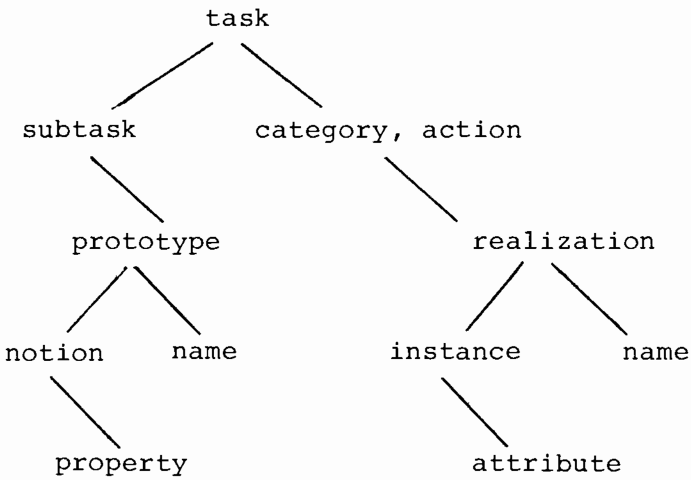


Figure 3. The physical capabilities of ARK (the left side demonstrates the hierarchical order and the right side the specifications).

About the user of ARK we have the following model. There are three types of them. First there is the computer specialist or software expert, then there is the problem specialist or systems analyst and thirdly there is the decision maker or daily user who deals very often with very specific problems.

The structure of ARK as presented in Figure 3 is only of interest for the computer specialist. He defines the prototypes to connect his own software with the whole system. He uses the categories to fit into the systems structure. The problem specialist connects the software moduls by connection of tasks and definition of objects. He uses the implemented programs as tools from the whole system. When this is done the decision maker works in the defined structure and analyses or solves his own problem by changing the attributes. This understanding of work with the problem solving environment makes it possible to view each type of the three users at the same time as an expert users and as a novice.

To work with the objects we developed an input-language. Each command has the following form:

category.name property 1 = attribute 1, property 2 = attribute 2, . . .

category - name of the category of the object (prototype)

name - name of the object (realization)

property - property of a selected notion

attribute - numerical value of the property.

Using this language the objects of the example in Figure 2 could be defined in the following way:

unit.w1 a=100m^2, type=hex, i=(str1, str2), 0 = (str3, str4), k-value = .37

stream.str1 temperature=300K, pressure=.1MPa, flow rate=10 kg/S, material=water

stream.str2 temperature=400K, material=air, flow rate=3kg/S, pressure=.2MPa

The objects defined in this way can be changed by the user at any time. The following command changes the material from air to water and deletes the numerical value for pressure:

stream.str2 material=water, pressure=

For the unskilled user a help function is provided in ARK, which works with a menu-like technique. The question mark is the input to the help function. We will demonstrate it by selecting an optimization method:

? task

taskname = opti

type = ?

select one of the following options:

optimization

integration

linear equations

zero finding methods

differential equations

:

optimization method = ?

random search

polytope method

variable metric

generalized reduced gradients

⋮

value of the stopping criteria = 1.0E - 02

name of the objective variable = obj

name of the independent variables = (ts1, ps1)

name of the mathematical model = example

starting point = (400, .2)

upper bounds = (600, 1)

⋮

For the use of the problem solving environment it is only needed to know the meaning of the provided options to specify the own problem. As in the above example if the user wants to solve the problem which consists in optimization of a technical system, he should not know the mathematical model. This could be provided by other users, e.g. the expert in modeling.

Let us assume the user wants to maximize the concentration of the third component of the stream st4. The optimization method would be selected as shown above and the mathematical model is in the database under the name "example". The user has to load it by

load task = example

and can then connect it with the optimization method. First he has to go into "example" by:

task.example

and then he provides the variables for the optimization by:

stream.st1 t=%=ts1, p=%=ps1

stream.st4 conc=(,, % = obj).

During the session the user can modify interactively the mathematical model or the optimization problem, for instance the change from maximization to minimization can be done by the simple statement:

obj = - obj .

If the user wants to run the specified optimization problem, the statement is:

solve task = opti .

The solution can be either stored, or numerically or graphically interpreted.

4. About some conclusions

The current version of ARK is implemented in assembler due to some software limitations. Currently under work is an extented version in C. We hope to have ARK then portable.

It seems important to mention here that the problem solving environment ARK is able to handle parallel processes. This means it has a feature comparable with background and foreground processes in UNIX. This allows the user to interrupt and restart processes, to change parameters and objects at any time even in runtime. Based on this feature of parallelism currently is work underway to use ARK as an operating system for a multiprocessor system and for running a local network.

On the other side we are trying to improve the system by including some of the advancements of Artificial Intelligence into ARK. A possible direction to do this seems to be given by symbiotic (knowledge-based computer support) systems (Fischer, 1983).

We understand this paper as one report about the continuing activities in extending and developing support systems for analyzing complex decision situations. This is by no means an end or a finished product.

References

[1] Barstow, D.R., Shrobe, H.E. and Sandewall, E. (1984): Interactive Programming Environments, McGraw-Hill, 1984.

[2] Blair, G.S., Malone, J.R. and Mariani, J.A. (1985): A Critique of UNIX, Software-Practice and Experience, Vol. 15 (12) 1125-1139, Dec. 1985.

[3] Fischer, G. (1983): Symbiotic, Knowledge-based Computer Support Systems, Automatica, Vol. 19, No. 6, pp. 627-637.

[4] Grauer, M., Lewandowski, A. and Wierzbicki, A (1984): DIDASS-Theory, Implementation and Experiences in Interactive Decision Analysis, Springer- Verlag, 1984.

[5] Winograd, T. (1979): Beyond Programming Languages, Comm. ACM 22, 7 (July 1979), pp. 391-401.

AN INTERACTIVE (OPTIMIZATION-BASED) DECISION SUPPORT SYSTEM FOR MULTI-CRITERIA (NONCONCAVE) CONTROL PROBLEMS

R. Straubel and A. Wittmüß
Central Institute for Cybernetics and Information Processes of the Academy of Sciences of GDR
DDR 1086 Berlin, German Democratic Republic

1. INTRODUCTION

With the aim of the computing PARETO-optimal strategies for (non-academic) real complex systems or processes - inside of a given time horizon T - we have developed an interactive decision support system characterized by the following aspects:

The use of the support system demands a minimum pre-knowledge on the background of the implemented algorithms of the decision theory. That means that the user can interactively generate his subjective preference structure respecting (non cooperative) different objective criteria based on computed results representing his own (scientific) subject. Therefore there were not only implemented normative decision-theoretical procedures, but the user gets descriptive decision support at each step of the program system, too. The number of decision situations (i. e. the number of questions put by the computer directed to the user) should be as small as possible.

The other aspect of the above mentioned restriction is : each desired input interactively given by the user is immediately tested according to its actual feasibility. If any input is not feasible the user is informed on the actually valid tolerance range for the input and he can choose a pre-programed value as an example.

In order to guarantee the application of the program system on real processes the implemented algorithms allow to compute both cases of non-convex/concave criteria and non-connected control regions. Continuous and/or integer nonlinear optimization problems can be solved.

2. NECESSARY PREPARATION FOR USING THE DECISION SUPPORT

Before the program system can be used one has to prepare a compiled subroutine in a certain form including an algorithm for the description of the process model which should be controlled, a formulation of the different criteria Q_i (i=1,..,n) depending on the vector $\underline{u}(t_i)$ of control, $\underline{z}(t_i)$ of state variables as function of the time steps t_i (i=1,.. ..,T), and a definition of the external model parameters $\underline{a}(t_i),\underline{b}(t_i)$.

It is assumed that the model is described in the form of a time discretized ensemble of subautomata with transition functions of the state vector

$$\underline{z}(t_{i+1}) = f(\underline{z}(t_i),\underline{z}(t_i-m),\underline{x}(t_i),\underline{u}(t_i),\underline{a}(t_i)) \tag{1}$$

and the output of the subautomata

$$\underline{y}(t_i) = \underline{g}(\underline{z}(t_i),\underline{u}(t_i),\underline{b}(t_i)) \tag{2}$$

where the interaction of the automata is

$$\underline{x}(t_i) = H\,\underline{y}(t_i) \tag{3}$$

with H as an interaction matrix. Additional restrictions can be explicitely considered in the form

$$\underline{u}_l(t_i) \leq \underline{u}(t_i) \leq \underline{u}_u(t_i) \tag{4}$$

and implicitely considered in the form

$$\underline{r}(\underline{u}(t_i),\underline{u}(t_i-k),\underline{z}(t_i),\underline{z}(t_i-j),\underline{a}(t_i),\underline{b}(t_i)) \geq \emptyset \tag{5}$$

where the indices u,l mean the upper and lower bounds of the vector $\underline{u}$, and k,m,j mean time delays, respectively.

For a user-friendly definition of the changeable external parameter vectors $a(t_i),b(t_i)$ there is a program based on a man-machine dialogue in natural English language. This program automatically generates a compileable input subroutine in a certain form used in the decision support system.

3. THE STRUCTURE OF THE PROGRAM SYSTEM

Considering descriptive decision theoretical aspects we classify the different wishes of a decision maker into three classes :

- (i) he wants to get a survey on the time behaviour of the process model,
- (ii) he wants to check the feasibility of some numerical values for special model variables (or relation between them) designed by him,
- (iii) he wants to compute the "best" PARETO-optimal decision alternative with respect to his own subjective preference structure which is (perhaps)to be generated simultaneously.

In agreement with these wishes our decision support system consists of three levels :

(i) a level for learning by simulation games,
(ii) a level for checking desired target values,
(iii) a level for computing PARETO-optimal alternatives.

These three levels are shortly described in the following (see also table 2).

3.1 Learning level

On this level we implemented three types of games. The first type represents a simple simulation of the process model for the whole time horizon T. Before the game is started one has to assign each control variable of the process model to anyone of the different (until 5) players. The type of the

TABLE 1 Regimentations of the different game types

Game type	Control variable	Time step 1...2...3...4... ...T-1...T	Involved players
1	1 2	o o o o o o	player 1 player 2
2	1 2	o o o x x x	player 1 computer
3	1 2	o o o x x x x x x x x x	player 1 computer

distribution of information of given players'input and of the computed results among the players after each step of the game has also to be chosen.

Then starting with the first time point in a chosen sequence each player has to determine the numerical values for his "own" control variables following his individual tactics.

The next two types of games are restricted to two players only where one player is represented by the computer itself which is computing its "own" control variables relating to one (or more, if a rank order for the objectives can be formulated) assigned objective/s. In the second game type, the computer only considers a reduced time horizon (from time point 1 until the actual point of the game step). In the third type, at each game step all those values of the control variables for the whole time horizon T which are not fixed earlier by player 1 are computed.

The different regimentations of the game types after the third time step (and for two control variables only) are illustrated in table 1 where o mens a component of a control variable determined "by hand" only and x is a computed component.

Repeating the input for any time point the first player (user) can iteratively improve his strategy on the basis of the computed results using the second game type from an operational viewpoint, and using the third type for a long term planning, respectively.

3.2 Level of checking desired targets

If a user wants to test the feasibility of a set of numerical values for any model variables (as state, input, output, and control variables or relations between them) it can be done in two ways :

On principle a qualitative one means the computation of the so-called utopical point represented by the individual optima for each assigned model variable (or relation between them). By this the upper and lower bounds of these model variables give a qualitative information on their feasibility range. Therefore firstly we implemented on this level a closed program for the possible computation of the utopical point. For this run it is possible to use different optimization procedures (see chapter 4).

On the other hand very often the user looks for a control strategy by which a certain number of model variables Q_i which are playing the role of objectives on the next level of the program system should attain simultaneously (at the end of the given time horizon) anticipated target values T_i with $i=1,..,n$. To solve this "inverse problem" we implemented a so-called Target method :

Here the actual maximum difference between a model variable $Q_i(T)$ and its target values T_i is minimzed using always the optimization procedure that is simulating the evolution process (see chapter 4) :

$$\min_{\underline{u}} \max_{i} \left| Q_i(T) - T_i \right| \tag{6}$$

If a feasible solution can be found (inside of a changeable range of accuracy) the resulting strategy $\underline{u}(t)$ is represented by the center of an n-dimensional sphere in the space of the Q_i. Inside of this sphere there are laying all values T_i. The radius of the sphere is depending on the preassumed accuracy.

If no feasible strategy for the chosen targets T_i can be found by the Target method there is implemented annother procedure searching a feasible strategy. But now the given targets

T_i are playing the role of upper and lower bounds of its Q_i, respectively :

for all i=1,.. ..,n holds

$Q_i(T) \geq T_i$ if the maximum value of $Q_i(T)$ is desired and

$Q_i(T) \leq T_i$ if the minimum value of $Q_i(T)$ is desired.

3.3 Level of PARETO - optimal solutions

In general the chosen targets T_i do not represent any PARETO-optimal alternative. Therefore, if the user of the decision support system wants to get one or more PARETO-optimal solutions he can let compute such alternatives on this level.

There are three different procedures ordered in a sequence of increasing (computer-) time consuming. The shortest one is the so-called sequential (hierarchical) rank order optimization (Podinowski et al. (1975), Straubel et al. (1983), Wittmüß et al. (1984)) which is simulating the human decisions in a very simple (but clear) way.

Using this method it is necessary that the user can formulate a (time independent) rank order of the criteria characterizing his own subjective preference structure. That means : the user has to assign to each criterion an (increasing) number of order (inside of the interval $[1,n]$ with the decreasing commitment regarding the criteria in such a way that the rank number of the most important criterion is equal "1" and that of the last criterion is equal "n". Of course, to get insight in the influence of a chosen rank sequence on the PARETO-optimal results the user can change this order if the procedure is repeated.

Then the method starts with the optimization of the first (most important) criterion. That means symbolically in the case of maximization :

$$\underline{\text{STEP 1}} : \max_{\underline{u}} Q_1(\underline{z}(t),\underline{u}(t)) \longrightarrow \hat{Q}_1(\underline{z}^{*(1)}(t),\underline{u}^{*(1)}(t)). \quad (7)$$

(The result of this step is already stored if the utopical point was computed on the second level). Now the user has to choose a value R_1 with the restriction $R_1 \leq \hat{Q}_1$ whree R_1 is the lowest (numerical) value for the criterion Q_1 acceptable for the user in the following optimization runs regarding the criteria assigned by higher rank numbers. In general (if the criteria Q_1,.. ..,Q_n are non-cooperative) that means : a lower value of R_1 allows to attain higher values of the following maximization runs regarding the next criteria.

Then the computing process starts again with the feasible strategy (of the first step). Now there runs a maximization procedure for Q_2 with an additional restriction :

$$\underline{\text{STEP 2:}} \quad \max_{\underline{u}} Q_2(\underline{z}(t),\underline{u}(t)) \longrightarrow \widehat{Q}_2(\underline{z}^{*(2)}(t),\underline{u}^{*(2)}(t)) \quad (8)$$

$$Q_1(z(t),u(t)) \geq R_1$$

The procedure ends if the n th criterion is sequentially optimized regarding the n-1 additional restrictions. After each step i the user has to decide between two alternatives :

- he can change the chosen value R_{i-1} and repeat the optimization run regarding the criterion Q_i once more, or
- he can choose a value R_i and continue the optimization run with respect to the next criterion Q_{i+1}.

The result of the whole procedure is only one PARETO-optimal strategy in agreement with the user´s subjective preference structure expressed by the chosen rank order of the criteria and the chosen numerical values for $R_1,\dots\dots,R_{n-1}$.

In fig. 1 is shown an extremely simple example for this method. It is characterized by three nonconvex objectives $Q_1(u_1),Q_2(u_1),Q_3(u_1)$ depending on one control variable only. Here the control variable $u_1^{*(0)}$ is computed (after the step 0 of the method) with respect to the criterion Q_2 (in agreement) with the chosen rank order of the criteria). Then by choosing of the lowest level R_1 for the criterion Q_1 the feasible control region can be divided into two (nonconnected) sub-regions.

If the user cannot formulate such a particular "a priori" preference structure he can realize one "best" PARETO-optimal alternative on the basis of sequentially computed PARETO-optimal solutions influenced by him by verbally formulated preferences, if he uses now the so-called Relaxation method (Eiduk (1981)), which is implemented on this level of the decision support system,too. This procedure runs as follows :

STEP 0 : Starting with any feasible control vector there will be computed any PARETO-optimal point characterized by the values $\widehat{Q}_1^{(o)},\dots\dots,\widehat{Q}_n^{(o)}$.

STEP 1 : On the basis of the displayed result the user has to articulate his actual preference now : He has to formulate which of the objectives should be improved, which of the objectives should be fixed, and which of the objectives may be permitted to become worse.

STEP 2 : Based on these formulated restrictions a new PARETO-optimal point will be computed now (Wittmüß(1985)). If the user is not satisfied by the newly displayed result then the procedure goes to step 1 again. In the other case it ends.

By this Relaxation method the user´s subjective preference structure will be iteratively generated on the basis of the sequentially computed PARETO-optimal points.

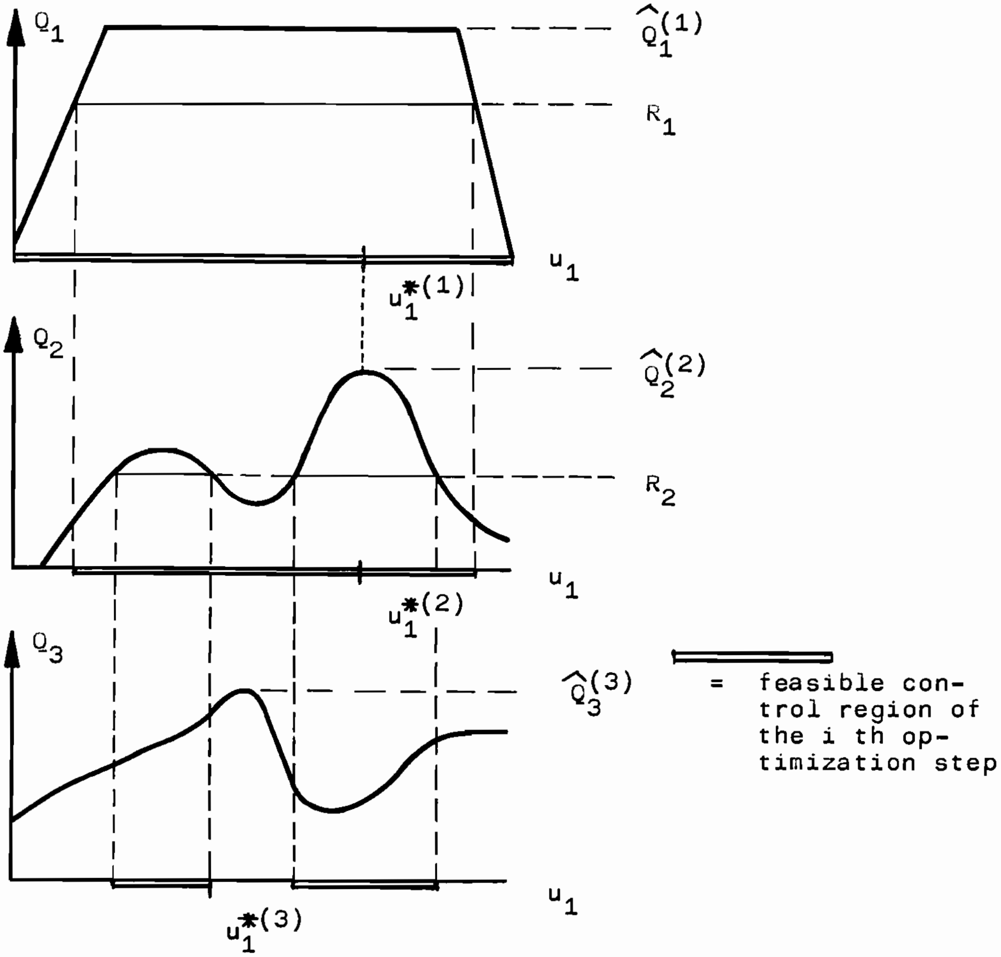


FIGURE 1 Example for the sequential rank order optimization

Finally, using the last implemented procedure on the third level of the decision support system the user can get a packed subset of the PARETO-optimal point set. Of course, this procedure is extremely time consuming. It ends if a desired number of points is computed for which a desired minimum "distance" between them is simultaneously guaranteed. In order to find quickly one computed point in the packed PARETO-subset (in agreement with the user´s subjective preference structure) there are implemented some supporting procedures giving a graphic survey of the "compromise set" and a "fuzzy" oriented evaluation (Böttner (1985)), respectively.

If the user wants to get the control vector for any PARETO point he can go to the first Target method on the second level of the decision support system where the corresponding control vector $\underline{u}(t)$ can be computed inside of a desired accuracy.

TABLE 2 Listing of implemented subroutines

Level	Implemented subroutines
0 : preparation	- choice of the language for the man-machine dialogue (English or German is available) - input of the changeable (already stored) external parameters for the process model - choice of the possibility of selection and storing PARETO-optimal points computed by any chance while the further procedures are running
1 : gaming	- I: All control variables are determined "by hand" - II: Operational game against the computer to which there is assigned one (or more) objective/s inside of the actual time horizon $\leq$ T - III: Planning game against the computer optimizing all control variables which are not fixed by the player inside of the whole time horizon T
2 : checking	- Computation of the "utopical point" - Target method considering restrictions in form of equations - Target method considering restrictions in form of unequations
3 : PARETO-optimization	- Sequential (hierarchical) rank order optimization - Relaxation method - Computation of the "compromise set"

4. THE IMPLEMENTED OPTIMIZATION PROCEDURES

On the different levels of the decision support system the implemented subroutines need different special deviations of optimization procedures. In principle we implemented two main procedures :

- a numerical gradient approximation for convex or concave objectives and connected control regions (Rosenmüller (1984))

- a stochastic search procedure based on the idea of biological evolution by mutation and selection (Rechenberg (1973), Schwefel (1977), Born et al. (1983), Born (1984)) for all other cases.

Owing to the complexity of the real process models it is impossible to test their mathematical characteristics. Therefore we use a mixed optimization method of both above mentioned procedures in a heuristic pragmatic way, too: First the gradient approximation leads to a local optimum (if it exists) and the following stochastic procedure tests it with respect to a (perhaps) better solution. If any better solution can be found the gradient approximation method is repeated.

Using the stochastic search procedure a selection test can be formulated respecting simultaneously more as one objective. For this the user has to choose a rank order of the objectives with regard to his preference. Firstly, the procedure tries to improve simultaneously all criteria. After a fixed number of calls of the process model without any success of any improvement - while the supply of control variables is changing - it will be neglected (step by step) the criterion with the lowest importance. Finally, in the case of non-cooperative criteria the"evolution test" considers the most important criterion only.

5. APPLICATIONS

Until now the described decision support system has been applied to two complex problems :

The first (continuous) problem is the computing of PARETO-optimal strategies for a national centrally planned (nonlinear) macroeconomic model with time delayed states characterized by 3 economic objectives and 12o control (rate-) variables (Wittmüß (1984)).

The second problem is the computing of a PARETO-optimal reconstruction sequence of regionally distributed emittants of air pollution. Here a region is divided into grids of equal size in which there are distributed the emittants and special neuralgic grid elements where the weather depending immission of air pollution should be reduced below a desired value inside the planning horizon. Each emittant is characterized by the emission before and after its reconstruction, by its reconstruction costs and the height of the chimney. Using the process model for each grid one can compute the time series of immission considering the average of weather conditions, the topology of the region and the actual emission distribution. Then the vector optimization procedure based on the idea of evolution supplies an (at first time-independent) sequence of control variables for each call of the process model. In such a way each emittant is represented by one control variable. In agreement with this sequence the number of different emittants which should be reconstructed in each year of the planning period depending on the given supply of total investment costs per year is computed while a model call is running. Simultaneously one gets a time series of the emission for each emittent. Finally on the bases of these computed results the values for the 4 objectives can be computed :

- (maximize) the number of neuralgic grid elements in which the reduction demands can be fulfilled,
- (maximize) the total emission reduction inside of the region,
- (minimize) the total reconstruction costs of the planning horizon,
- (minimize) the time horizon below the maximum value T.

Notice that this problem is an integer programming one which can be solved with the evolution principle, too.

6. REFERENCES

Böttner, R. (1985). A fuzzy approach for the selection from a limited number of alternatives. Systems Analysis and Simulation 1985. Akademie-Verlag Berlin, 27:259-262

Born, J., and Bellmann, K. (1983). Numerical adaption of parameters in simulation models using evolution strategies. In K. Bellmann (Ed.), Molecular genetic information systems: modelling and simulation. Akademie-Verlag, Berlin

Born, J. (1984). ASTOP Handbook. Central Institute for Cybernetics, Acad. Sci GDR. Berlin

Eiduk, J.Y. (1981). Vector-relaxational algoritms for tradeoff decision search. In: Methods and models of decision analysis, Riga

Podinowski, W.W., and Gawrilow, W.M. (1975). Optimization by sequential criteria (in Russ.). Soviet Radio, Moscow

Rechenberg,J. (1973). Evolutionsstrategie. Friedrich-Fromman-Verlag, Stuttgart

Rosenmüller, R. (1984). Ein Verfahren der Gradienten-Approximation im Rahmen der mehrkriterialen sequentiellen Rangfolge-Optimierung. 13. Jahrestagung, Grundlagen der Modellierung und Simulationstechnik, Rostock

Schwefel, H.P. (1977). Numerische Optimierung von Computer-Modellen mittels der Evolutionsstrategie. Interdisciplinary systems research, 26. Birkhäuser-Verlag, Basel und Stuttgart

Straubel, R., and Wittmüß, A. (1983). Ein interaktives Entscheidungsmodell zur optimalen Steuerung eines Systems mit mehreren Zielfunktionen. msr, 26 : 2-4

Wittmüß, A., Straubel, R., and Rosenmüller, R. (1984). Intertive Multi-Criteria Decision Procedure for Macroeconomic Planning. Syst. Anal. Model. Simul., 1 : 441-424

Wittmüß, A. (1985). Scalarizing multiobjective optimization problems and an interactive approach for multiobjective decision making. Systems Analysis and Simulation 1985. Akademie-Verlag Berlin, 27 : 255-258

TOWARD THE CONSOLIDATION OF INTERACTIVE MULTIPLE OBJECTIVE PROGRAMMING PROCEDURES

Ralph E. Steuer
College of Business Administration, University of Georgia
Athens, Georgia 30602 USA

Alan W. Whisman*
Navy Personnel Research and Development Center,
San Diego, California 92152 USA

1. INTRODUCTION

The paper discusses a consolidated approach to multiple criteria optimization that subsumes most of the most prominent interactive procedures that have been designed for solving multiple objective programming problems. The consolidated approach not only allows users to choose which interactive procedure to apply, but it also allows users to select different procedures on different iterations if so desired. Also, comments are made about multiple criteria information transfer protocols using computer graphics and voice communications capabilities at the computer/user interface.

2. THE GENERAL MULTIPLE OBJECTIVE PROGRAMMING PROBLEM

The general multiple objective programming problem is formulated as

$$\begin{array}{ll} \max \{f_1(x) = z_1(x)\} & \\ \quad \vdots & \\ \max \{f_k(x) = z_k(x)\} & \\ \text{s.t.} \quad & x \in S \end{array}$$

where the f_i need not be linear and S need not be convex. Thus, the above formulation can be used to refer to a multiple objective linear program

*The opinions expressed in this paper are those of the authors and do not reflect the official views of the Navy Department.

(MOLP), a multiple objective integer program, or a nonlinear multiple objective program.

Mathematically, a solution of a multiple objective program (regardless of whether it is linear, integer, or nonlinear) is an efficient point. A point $\bar{x} \in S$ is efficient if and only if there does not exist another $x \in S$ such that $z_i(x) \geqq z_i(\bar{x})$ for all i and $z_i(x) > z_i(\bar{x})$ for at least one i. The set of all efficient solutions is called the efficient set and is denoted E. Let $\bar{x} \in S$. Then, its image $z(\bar{x}) \in R^k$ in criterion space is nondominated if and only if $\bar{x} \in E$. The set of all nondominated criterion vectors is called the nondominated set and is denoted N.

We know from theory that a decision maker's optimal solution is efficient, and thus, the image of the optimal solution in criterion space is nondominated. Since the nondominated set N is typically very large, a number of carefully designed interactive procedures have been developed for exploring and sampling N. The purpose of these interactive procedures is to find a final solution where a final solution is defined to be either optimal, or close enough to being optimal, to terminate the decision process.

With this theory in mind, and knowledge of the fact that numerous interactive procedures have been proposed for solving multiple objective programming problems, this paper shows how a consolidated algorithm can be developed that subsumes most of the interactive procedures that are prominently mentioned in the literature. We will refer to the computer implementation of the consolidated algorithm as CONSOL.

Three CONSOL codes are envisaged: CONSOL-M (for multiple objective linear and nonlinear programs), CONSOL-I (for multiple objective integer programs), and CONSOL-N (for multiple objective network programs). In this paper we will talk primarily about CONSOL-M which will be written in FORTRAN in order to (a) be internationally transportable, (b) have large-scale capabilities, and (c) be consistent with the MINOS linear/nonlinear code [Murtagh and Saunders (1980)] which will be used as the consolidated algorithm's workhorse optimization software.

3. TWO WAYS OF CLASSIFYING INTERACTIVE PROCEDURES

Interactive multiple objective programming procedures can be classified as either

(a) feasible region reduction,

(b) weighting vector space reduction,

(c) criterion cone contraction,

(d) semi-structured probing, or

(e) line search

methods. Independent of this classification scheme, interactive procedures can be classified as either sampling the nondominated set N "from below" or "from above."

The first classification is useful for categorizing interactive procedures according to the philosophical strategies of their approaches. The second classification is useful for categorizing interactive procedures according to certain of their implementation similarities.

Let

$$\Lambda = \{\lambda \in R^k \mid \lambda_i \in (0,1),\ \sum_{i=1}^{k} \lambda_i = 1\}$$

A procedure is said to sample N "from below" if we solve the <u>weighted-sums program</u>

$$\max \{ \sum_{i=1}^{k} \lambda_i f_i(x) \mid x \in S\}$$

one or more times using different weighting vectors $\lambda \in \Lambda$ to generate the probing or sampling of criterion vectors required at each iteration.

A procedure is said to sample N "from above" if we first define a reference point $z^* \in R^k$ (typically suspended above the feasible region in criterion space) and then solve the <u>minimax program</u>

$$\begin{array}{lll} \min \{\alpha\} & & \\ \text{s.t.} & \alpha \geqq \lambda_i w_i & 1 \leqq i \leqq k \\ & f_i(x) = z_i & 1 \leqq i \leqq k \\ & z_i + w_i = z_i^* & 1 \leqq i \leqq k \\ & x \in S & \end{array}$$

(or a program similar to it) one or more times using different weighting vectors $\lambda \in \Lambda$ to generate the probing or sampling of criterion vectors required at each iteration.

4. PROMINENT INTERACTIVE PROCEDURES

A list of the most prominent interactive procedures that can be run as special cases of the consolidated approach include:

1. STEM: This algorithm was proposed by Benayoun, de Montgolfier, Tergny, and Laritchev (1971). STEM is categorized as a feasible region reduction/sampling from above method.

2. GDF: Geoffrion-Dyer-Feinberg procedure (1972). GDF is categorized as a line search/sampling from below method.
3. ICW: Interval Criterion Weights method of Steuer (1977). ICW is categorized as a criterion cone contraction/sampling from below method.
4. ECON: e-Constraint method. This is a traditional, common-sense method whose origins are not attributable to any particular author. ECON is categorized as a reduced feasible region/sampling from below method.
5. IWS: Interactive Weighted-Sums method. This is another common-sense method whose implementation has been formalized in Steuer and Schuler (1981). IWS is categorized as a weighting vector space reduction/sampling from below method.
6. TCHA: Augmented Weighted Tchebycheff procedure of Steuer and Choo (1983). TCHA is categorized as a weighting vector space reduction/sampling from above method.
7. REF: Reference Point method of Wierzbicki, Lewandowski, and Grauer (1980, 1982, and 1983). REF is categorized as a semi-structured probing/sampling from above method.
8. VIA: Visual Interactive Approach of Korhonen and Laakso (1986). VIA is categorized as a line search/sampling from above method.

At this point, it does not appear that it will be possible, because of its substantially different logic structure, to integrate the Zionts-Wallenius procedure (1976 and 1983) into the consolidated approach. Also, it is unclear at this point whether or not it will be possible to incorporate Goal Programming or the Surrogate Worth Trade-Off method of Haimes and Hall (1974) into the consolidated approach without seriously affecting the streamlined nature of the CONSOL implementation.

5. CONSOLIDATED ALGORITHM

Although the procedures that can be incorporated into the consolidated approach represent a wide range of divergent philosophical strategies, they have many significant implementation similarities. Prior to whatever optimizations are to be performed at each iteration, we must supply the procedures with λ-weighting vectors, criterion vector reference points $z^* \in R^k$, e_i-criterion value lower bounds, or criterion space search directions $d \in R^k$. This information can be developed off-line according to the idiosyncratic prescriptions of the specific interactive procedures and then placed in a standardly formatted file for access by the on-line consolidated algorithm. After accessing the appropriately developed information from the standardly formatted file, the consolidated algorithm proceeds to the optimization stage of the iteration's cycle. Then the candidate criterion vectors resulting from the optimization stage are evaluated prior to looping for another iteration. The consolidated algorithm is outlined as follows:

Step 1: Let $h = 0$ and perform necessary initializations.

Step 2: Let $h = h + 1$ and specify the interactive procedure to be performed.

Step 3: Develop appropriate weights, reference points, criterion value lower bounds, or search directions and place in standardly formatted file.

Step 4: Customize meta program (discussed in Section 6).

Step 5: Perform meta program optimizations.

Step 6: Evaluate criterion vector results.

Step 7: If it is time to stop iterating, go to Step 8. Otherwise, go to Step 2.

Step 8: Stop with final solution.

A flowchart of the consolidated algorithm is shown in Figure 1.

6. CUSTOMIZING THE META PROGRAM

At the core of the consolidated algorithm is the meta program

$$\min \; \{\alpha + \rho \sum_{i=1}^{k} (z_i^* - z_i) - \sum_{i=1}^{k} \lambda_i^{(1)} z_i\} \tag{5.1}$$

$$\text{s.t.} \quad \alpha \geqq \lambda_i^{(2)} w_i \qquad 1 \leqq i \leqq k \tag{5.2}$$

$$f_i(x) = z_i \qquad 1 \leqq i \leqq k \tag{5.3}$$

$$z_i \geqq e_i \qquad 1 \leqq i \leqq k \tag{5.4}$$

$$z_i + w_i = z_i^* + \theta d_i \qquad 1 \leqq i \leqq k \tag{5.5}$$

$$x \in S \tag{5.6}$$

$$(z,\alpha) \in R^{k+1} \text{ unrestricted} \tag{5.7}$$

The generalized nature of the meta program is that it subsumes (a) the weighted-sums types of programs inherent in the interactive procedures that sample N from below and (b) the minimax types of programs inherent in the interactive procedures that sample N from above.

To illustrate how the meta program can be adapted for use by the

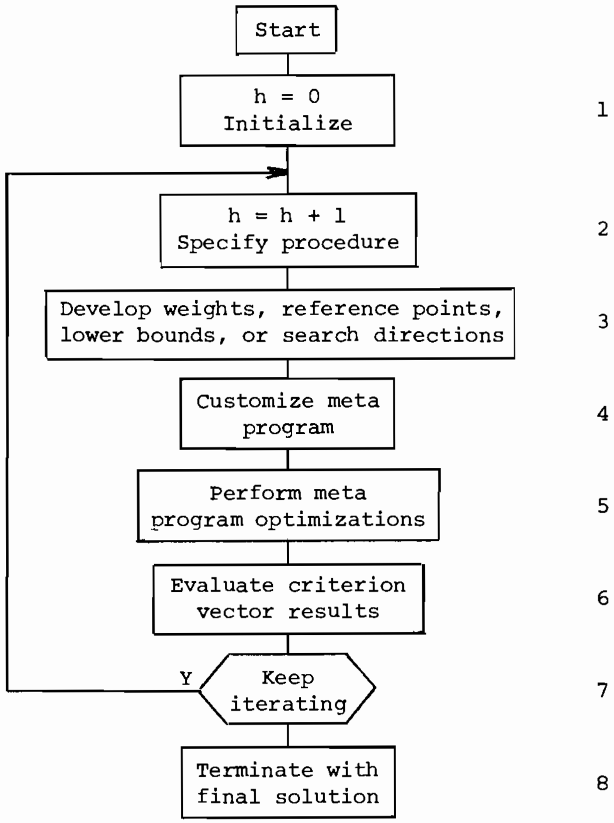


Figure 1. Flowchart of the Consolidated Algorithm

different interactive procedures, consider for instance, the minimax optimization program in STEM:

(a) $\rho = 0$ and the $\lambda_i^{(1)} = 0$ in (5.1)

(b) the $\lambda_i^{(2)}$ weights in (5.2) are calculated according to the rules specified in Benayoun, et al. (1971)

(c) some of the lower bound constraints of (5.4) are vacuous

(d) $\theta = 0$ in (5.5)

For the direction finding optimization program in the GDF procedure, we have

(a) $\rho = 0$ and $\alpha = 0$ in (5.1)

(b) the $\lambda_i^{(1)}$ weights in (5.1) are either specified directly or induced for the decision maker from responses to a pairwise comparison questioning routine

(c) constraints (5.2), (5.4), and (5.5) are vacuous

Similar customizations are performed for the other interactive procedures.

7. EMPIRICALLY-BASED COMMENTS ABOUT CONSOLIDATED APPROACH

It is commonly acknowledged that no interactive multiple objective programming procedure is universally superior to all of the others. Certain procedures appear to be better suited for certain problems, and different decision-making styles often result in users preferring different procedures all other things equal.

Also, there is the recent empirical research of Buchanan (1985) and Brockhoff (1985) that shows that many users might wish to shift among procedures when solving a problem, if such a capability were available. For example, a user may wish to start with TCHA for the first few iterations and then switch to REF for the rest.

Thus, there is a need for a common computer package such as the consolidated algorithm that not only allows one to select a specific procedure for a given problem, but also allows one to select different procedures on different iterations if so desired. With regard to switching procedures from one iteration to the next, the challenge is to ensure that enough information is saved at each iteration so that little or none of the already attained convergence achievements are lost on the interchange.

8. FURTHER COMMENTS ABOUT THE CONSOLIDATED ALGORITHM

In the initialization stage of Step 1, the consolidated algorithm computes the vector of maximal criterion values $z^{max} \in R^k$ where

$$z_i^{max} = \max \{f_i(x) \mid x \in S\}$$

This vector is required because $z^* = z^{max}$ in STEM, and the respective z^* reference points in TCHA, REF, and VIA are set in the light of z^{max}. If the problem to be solved is an MOLP, the consolidated algorithm will also compute the vector of minimal criterion values over the efficient set $z^{min} \in R^k$ where

$$z_i^{min} = \min \{f_i(x) \mid x \in E\}$$

Vector z^{min} will be computed using the algorithm described in Isermann and Steuer (1986).

Taken together, z^{max} and z^{min} will be used to portray the criterion value ranges over the efficient set. This provides an overall frame of reference for the problem being solved because the decision maker's optimal criterion vector is contained in these ranges. In the extreme, the ranges tell us how much may have to be sacrificed by a given objective to have high achievements in several of the other objectives. The range information should be particularly valuable to the ECON method because the user steers this procedure by iteratively altering e_i-criterion value lower bounds. Consistent with the Elimination by Aspects theory of Tversky (1972), Buchanan (1985) found in his experiments that many users, regardless of procedure, wanted the freedom to set e_i-lower bounds in excess of the z_i^{min}.

In general, the benefits of being able to set and reset e_i-criterion value lower bounds in multiple objective programming are two. It prevents the consolidated algorithm from developing criterion vectors that the user knows he or she does not want to see anymore, and, by reducing the efficient set, it should enable the consolidated algorithm to converge to a final solution faster.

9. COMMUNICATIONS AT THE COMPUTER/USER INTERFACE

In Step 5 of the consolidated algorithm, the meta program generates the candidate criterion vectors required by the activated interactive procedure of the current iteration. Then the generated criterion vectors are stored in a standardly formatted file so that they can be accessed in Step 6 by an off-line routine for displaying the criterion vectors. One modality for displaying and evaluating the generated criterion vectors would be to use high resolution computer graphics and voice communication capabilities.

Consider a large room with a computer workstation off to one side. On one wall is hung a large screen (6 ft. diagonal or larger). About the room are placed several chairs for viewing the screen. In the center of the room is a color overhead projection system wired to the RGB output ports of the computer terminal. Now, instead of sitting 18 inches in front of a computer screen, users can sit about the room and view computer images on the large screen. Rather than being crowded in front of a computer screen, users can more freely move about the room and discuss the contents of the displays.

In order to page through the different displays in whatever order is desired, the user must be able to communicate with the computer. This could be done through a tie clasp microphone connected to a computer workstation equipped with a voice communications adapter programmed to understand natural language commands. For example, to view a different mix of criterion vectors in a different format, the user might say "Criterion vectors 2, 3, and 5. Side-by-side bar chart."

10. THE CONSOL FAMILY OF CODES

Since CONSOL-M will use the linear-nonlinear MINOS code as its workhorse software for performing the meta program optimizations, it is ideally suited for solving multiple objective linear and nonlinear programming problems. However, users are cautioned about using any of the procedures that sample N from below (ICW, IWS, and ECON) on nonlinear problems because of the possible existence of unsupported nondominated criterion vectors.

CONSOL-I, whose purpose will be to solve multiple objective integer programs, will utilize much of the coding logic from CONSOL-M except that CONSOL-I will not support GDF and VIA because they do not lend themselves to integer applications. Also CONSOL-I would differ from CONSOL-M in that it would access an integer code such as MPSX-MIP [7] as its workhorse software for performing the meta program optimizations.

CONSOL-N is another CONSOL variation. It is directed at multiple objective network problems. In order to avoid side constraints, CONSOL-N will only support ICW and IWS so that high-speed network software such as the NETFLOW procedure in the SAS/OR [12] package can be accessed. In this way, a consolidated approach, as discussed in this paper, should be able to address the widest possible range of multiple objective programming problems with the greatest amount of flexibility and least amount of software.

REFERENCES

[1] Benayoun, R., J. de Montgolfier, J. Tergny, and O. Laritchev (1971). "Linear Programming with Multiple Objective Functions: Step Method (STEM)," Mathematical Programming, Vol. 1, No. 3, pp. 366-375.

[2] Brockhoff, K. (1985). "Experimental Test of MCDM Algorithms in a Modular Approach," European Journal of Operational Research, Vol. 22, No. 2, pp. 159-166.

[3] Buchanan, J. L. (1985). "Solution Methods for Multiple Objective Decision Models," Ph.D. Dissertation, University of Canterbury, Christchurch, New Zealand.

[4] Geoffrion, A. M., J. S. Dyer, and A. Feinberg (1972). "An Interactive Approach for Multicriterion Optimization with an Application to the Operation of an Academic Department," Management Science, Vol. 19, No. 4, pp. 357-368.

[5] Grauer, M. (1983). "A Dynamic Interactive Decision Analysis and Support System (DIDASS): User's Guide," WP-83-60, International Institute for Applied Systems Analysis, Laxenburg, Austria.

[6] Haimes, Y. Y. and W. A. Hall (1974). "Multiobjectives in Water Resources Systems Analysis: The Surrogate Worth Trade-Off Method," Water Resources Research, Vol. 10, No. 4, pp. 615-623.

[7] IBM Document No. SH19-1099-1 (1975). "IBM Mathematical Programming

System Extended/370 (MPSX/370), Mixed Integer Programming/370 (MIP/370): Program Reference Manual," IBM Corporation, Data Processing Division, White Plains, New York.

[8] Isermann, H. and R. E. Steuer (1986). "Payoff Tables and Minimum Criterion Values Over the Efficient Set," College of Business Administration, University of Georgia, Athens, Georgia.

[9] Korhonen, P. and J. Laakso (1986). "A Visual Interactive Method for Solving the Multiple Criteria Problem," European Journal of Operational Research, forthcoming.

[10] Lewandowski, A. and M. Grauer (1982). "The Reference Point Optimization Approach - Methods of Efficient Implementation," WP-82-26, International Institute for Applied Systems Analysis, Laxenburg, Austria.

[11] Murtagh, B. A. and M. A. Saunders (1980). "MINOS/AUGMENTED User's Manual," Report SOL 80-14, Department of Operations Research, Stanford University, Stanford, California.

[12] SAS Institute (1985). "SAS/OR Users Guide," Cary, North Carolina: SAS Institute, Inc.

[13] Steuer, R. E. (1977). "An Interactive Multiple Objective Linear Programming Procedure," TIMS Studies in the Management Sciences, Vol. 6, pp. 225-239.

[14] Steuer, R. E. and E. U. Choo (1983). "An Interactive Weighted Tchebycheff Procedure for Multiple Objective Programming," Mathematical Programming, Vol. 26, No. 1, pp. 326-344.

[15] Steuer, R. E. and A. T. Schuler (1981). "Interactive Multiple Objective Linear Programming Applied to Multiple Use Forestry Planning," In M. C. Vodak, W. A. Leuschner, and D. I. Navon (eds.), Symposium on Forest Management Planning: Present Practice and Future Directions, FWS-1-81, School of Forestry and Wildlife Resources, VPI, Blacksburg, Virginia, pp. 80-93.

[16] Tversky, A. (1972). "Elimination by Aspects: A Theory of Choice," Psychological Review, Vol. 79, pp. 281-289.

[17] Wierzbicki, A. P. (1980). "The Use of Reference Objectives in Multiobjective Optimization," Lecture Notes in Economics and Mathematical Systems, No. 177, Berlin: Springer-Verlag, pp. 469-486.

[18] Zionts, S. and J. Wallenius (1976). "An Interactive Method for Solving the Multiple Criteria Problem," Management Science, Vol. 22, No. 6, pp. 652-663.

[19] Zionts, S. and J. Wallenius (1983). "An Interactive Multiple Objective Linear Programming Method for a Class of Underlying Nonlinear Utility Functions," Management Science, Vol. 29, No. 5, pp. 519-529.

The interactive multiobjective package IMPROVE with application to strategic planning of carbochemical industry

Anne B. Barnikow*, Ralph Gollmer and Manfred Grauer

Academy of Sciences of the GDR, Institute of Informatics and *Institute of Chemical Technology, GDR-1199 Berlin, Rudower Chaussee 5

1. Introduction

In the field of chemical engineering and large-scale economic planning the classical approaches of single-objective thinking with an economic efficiency criterion only have shown to be inadequate for several reasons (Grauer, Lewandowski and Wierzbicki, 1983). The increasing awareness of the finiteness of natural resources and of the urgent necessity to protect the environment from hazardous emissions led to the requirement to consider in decision problems as well as economic efficiency and also resources and energy conservation and environmental protection criteria.

For that reason the Institutes of Chemical Technology and of Informatics cooperate on the generation and evaluation of efficient alternatives for the development of the carbochemical industry of the GDR. This problem is of importance due to the fact that lignite represents the only indigenous resource worth for processing in the chemical industry *and* for energy production.

Based on the experiences of one of the authors with the development and applications of the DIDASS system (Grauer, 1983, Lewandowski and Grauer, 1982, Grauer, Lewandowski and Schrattenholzer, 1982) a package for linear multiobjective programming IMPROVE (*I*nteractive *M*ultiobjective *P*rogramming using *R*eference *O*bjectives) was developed. It uses the reference point approach (Wierzbicki, 1980).

In this paper we report about some extentions of IMPROVE based on an idea of parametric programming (part 3) and its application to strategic planning (part 4). In part 2 as an introduction the achievement function approach for the linear case is briefly summarized.

2. The achievement function approach for the linear case

Let the problem be written as

$$(P) \qquad \max \{ Cx = \begin{bmatrix} C_1^T x \\ \vdots \\ C_p^T x \end{bmatrix} \mid x \in M \} \tag{1}$$

where M denotes a convex polyhedron. Let for simplicity of notation M be given in canonical form

$$M = \{ x \in R^n \mid Ax = b,\ x \geq 0 \} \quad . \tag{2}$$

Let furthermore adopt the following notations:

$$Q = \{ q \in R^p \mid q = Cx,\ x \in M \} \tag{3}$$

Q-the set of admissible points in the objective space,

$$R_+^p = \{ q \in R^p \mid q_i \geq 0\ i = 1,\ldots,p \} \tag{4}$$

R_+^P-the nonnegative orthant used as dominance cone,

$$Q^0 = \{ q \in Q \mid Q \cap (q + int\ R_+^p) = \phi \} \tag{5}$$

Q^0-the set of weakly maximal points in Q,

$$M^0 = \{ x \in M \mid Cx \in Q^0 \} \tag{6}$$

M^0-the set of weakly efficient points in M,

$$Q^* = \{ q \in Q \mid Q \cap (q + [R_+^p \setminus \{0\}]) = \phi \} \tag{7}$$

Q^*, the set of maximal points in Q,

$$M^* = \{ x \in M \mid Cx \in Q^* \} \tag{8}$$

M^*-the set of efficient (Pareto-optimal) points in M.

In order to generate efficient alternatives we use an achievement scalarizing function which is especially suitable for linear problems both, from the methodological and numerical point of view, and has shown its usefulness in the DIDASS package.

This function has the following form:

$$s(w) = \max \{ \rho \cdot \max_{i \in \{1,\ldots,p\}} (-w_i),\ -\sum_{i=1}^{p} w_i \} - \varepsilon \sum_{i=1}^{p} w_i \tag{9}$$

where the $w_i = (C_i^T x - \bar{q}_i)/\sigma_i$, denote the weighted deviations from the reference levels $\bar{q}_i$, the σ_i represents scaling factors, ρ and ε are control parameters ($\rho \geq p$, $\varepsilon \geq 0$).

We will demonstrate the structure of this function briefly by its level sets (see Figure 1):

$$s(\alpha) = \{ q \in R^p \mid s(w) \leq \alpha,\ w_i = (q_i - \bar{q}_i) / \sigma_i,\ i = 1, \cdots, p \} \quad . \tag{10}$$

The substitute problem can be rewritten as

$$min \{ s(w) \mid w_i = (C_i^T x - \bar{q}_i)/\sigma_i,\ i = 1,\ldots,p,\ x \in M \} \tag{11}$$

The advantage of this achievement function is that by introducing an additional variable and additional constraints the problem (11) can equivalently be formulated as a linear programming problem:

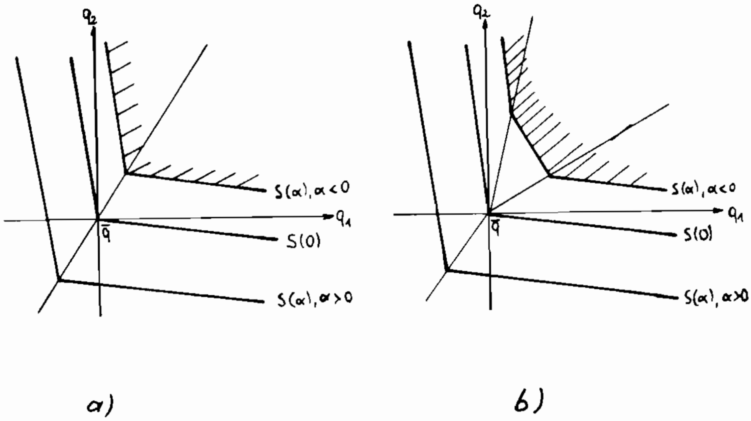


Figure 1. Level sets of the scalarizing function s a) for $\rho = p = 2$, b) for $\rho > p = 2$.

$$\left.\begin{array}{lrl} & \min \{y - \varepsilon \sum_{i=1}^{p} w_i\} & \\ \text{s.t.} & -yN_{p+1} + Dw \le 0 & (P1) \\ & Cx - diag(\sigma_i)w = \bar{q} & (P2) \\ & Ax = b & (P3) \\ & x \ge 0 & \end{array}\right\} \qquad (12)$$

where $N_K^T = (1,1,\ldots,1)$ with k elements,

$D = (-N_p^T / -\rho \cdot I^p)$ and I^p is the $(p*p)$ unit matrix .

The basic theorem characterizing the solutions of problem (12) was proved in (Kallio, Lewandowski and Orchard-Hays, 1980) for $\sigma_i = 1$, $i = 1,\ldots,p$:

Theorem: Let $(\hat{y}, \hat{w}, \hat{k})$ be an optimal solution of problem (12) for $\bar{q} \in R^p$, $\sigma > 0$, $\rho \ge p$ and $\varepsilon \ge 0$, and let $(\hat{\delta}, \hat{\mu}, \hat{\pi})$ be the Lagrange multipliers related to constraints (P1), (P2), (P3) respectively. Denote by $\hat{q} = C\hat{x}$ the corresponding objective vector and by $\hat{s} = \hat{y} - \varepsilon \sum_{i=1}^{p} \hat{w}_i$ the optimal objective value. Then $\hat{q} \in Q \cap S(\hat{s})$ and the hyperplane

$$H = \{q \in R^p \mid \tilde{\mu}^T(\hat{q} - q) = 0\} \qquad (13)$$

separates Q and $S(\hat{s})$. Furthermore it holds that $\sigma_i \mu_i \ge \varepsilon$, $i = 1,\ldots,p$ and $q = \hat{q}$ maximizes $\hat{\mu}^T q$ over Q, which yields $\hat{x} \in M^0$ if $\varepsilon = 0$ and $\hat{x} \in M^*$ if $\varepsilon > 0$.

This theorem, on the one hand assures the efficiency of the solutions of the substitute problems for every reference point $\bar{q}$ and on the other hand explains the role of the Lagrange multipliers μ as *a priori* weighting coefficients defining a supporting hyperplance to the set Q^* (see Figure 2).

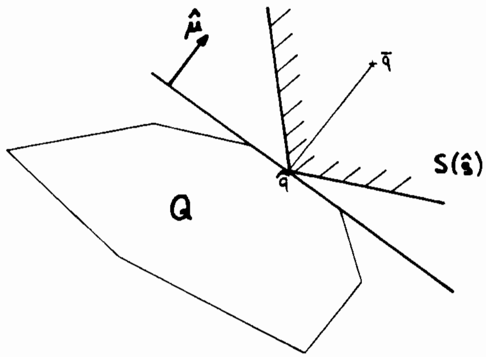


Figure 2. The separating hyperplane.

3. About some extensions

The idea of (Lewandowski and Grauer, 1982) to use the information included in μ for the estimation of efficient points corresponding to new reference points is implemented in the IMPROVE package. For this purpose the dual informations of all the solutions generated during a session are gathered in order to form an ever improving outer approximation of the set Q^*. Using this approximation the response of the model to a new reference point can be estimated by the solution of an auxiliary problem with a much smaller dimensionality than the original one - the number of variables equals to the number of objectives in the original problem.

This approximation feature may save computation time when applied to large problems because sometimes the user may conclude that the response to the current reference point would not be interesting thus rejecting this one and avoiding a longer computation with the original model. This rejection is justified because the actual solution can never be better than the estimated due to the outer approximation property.

In the current applications (120 * 150 matrix, see part 4) the construction and solution of the estimated problem takes about 40% compared with the time for the solution of the original model for a new reference point. But in the case of trajectory optimization (dynamic linear programming) the decrease could be expected to be more substantive.

Another extension of the IMPROVE package is the inclusion of an "improvement facility", i.e. the user can work in a way known from other approaches (e.g. the e-constraint-method). Starting from an efficient point he can specify a subset of the objectives to be improved which will cause a deterioration of the others. We included this feature into the framework of the reference point approach by proving the following Lemma.

Lemma: Let $q^* \in Q^*$ and $k \subset \{1,\ldots,p\}$, $\rho \geq p$, $\sigma > 0$ be fixed.

$$\text{Choose } \bar{q}_i = \begin{cases} q_i^* & i \notin k \\ q_i^* + t \cdot \sigma_i & i \in k \end{cases}$$

where t denotes a positive real number. Then for sufficiently small $\varepsilon > 0$ it holds that if $(\hat{y}, \hat{w}, \hat{x})$ is a solution of problem (12) then $C_i^T \hat{x} \geq q_i^*$ for all $i \in k$.

The proof of this Lemma relies on the properties of the chosen achievement function, and on results from linear parametric programming. Besides these two extensions a more flexible user interface is implemented. It is based on a specific designed command language and includes extended information facilities to assist the evaluation of the generated solutions. This includes the possibility to recall history information and selected parts of the solution as well as processing simple operations (+, -, *, /) on indicated variables of the solution file for the purpose to give aggregated output information.

4. The strategic planning problem

The IMPROVE package is used to analyse the decision problem of strategic planning of the carbochemical industry of the GDR. This industry in a nationwide view represents a complex system of several technological processes and material flows with a great number of interdependencies and various possibilities of combining them. Furthermore there exist several connections to the rest of the economy like for instance to the petrochemical industry and the energy system. The modelling approach used gives also the possibility to take into account new technological processes. The current version of the carbochemical industry model is of static linear input-output type.

The connections of the carbo-chemical industry to the rest of the economy are reflected by bounds on feedstock materials, energy consumptions and by the demand for target products. The single technological process is represented in Figure 3.

Inputs are the feedstock materials to be processed, costs of several types and resources like energy, water and labor. Outputs are the products of the given process.

A scoring approach is used to evaluate the environmental effects of the system. Using experts opinion, a score from the range 0 (no hazards) to 1 (heavy environmental hazards) is assigned to each of the processes. These scores are treated for the whole system like costs or material streams.

In order to obtain a linear model of the whole system all input-output relations for each process are assumed to be linear. For that purpose all data is normalized to a certain unit size for each process. The transportation costs are also included in a way that they are fixed for the unit size of the specific process.

The model on the one hand reflects the existing structure of the carbochemical industry and on the other hand the impact of investments for capacity expansions. The investments behavior is modelled by a piecewise linear function. This means an existing process is bounded by its capacity and causes no investment but "new" processes have no bounds on capacity but require investments.

The single processes are compelled by their interactions and form so the whole system. Figure 4 shows the structure of the whole carbochemical system for processing of two types of lignite under the discussed simplifications. The

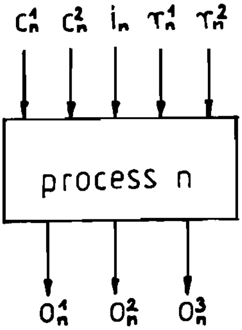


Figure 3. Modelling of a single process (i_n-input materials, C_n^k-costs, r_n^k-resources, O_n^k- products).

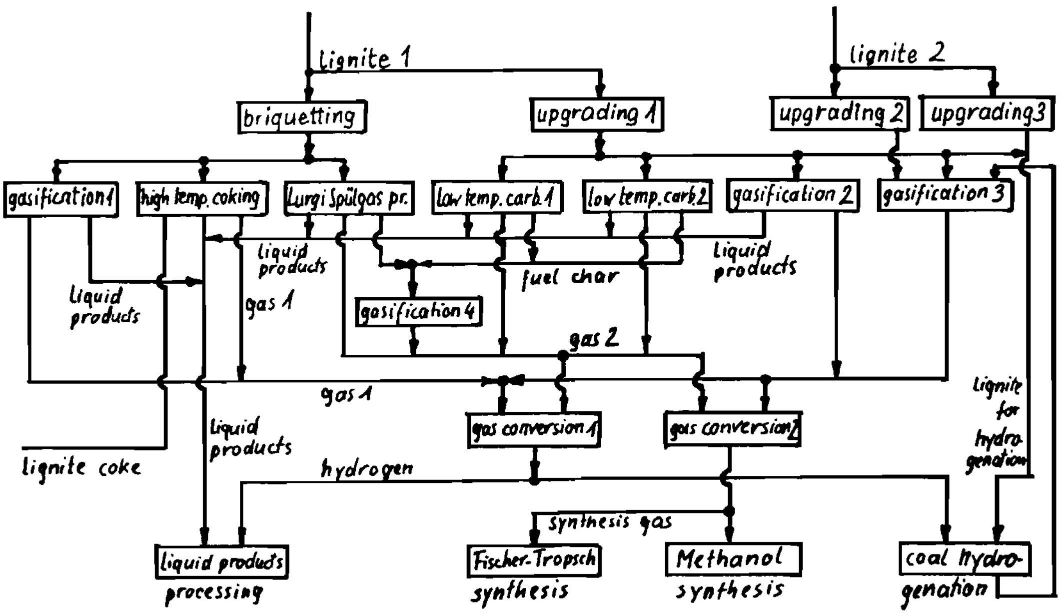


Figure 4. Simplified scheme of lignite processing.

following variables appear in the system: the load rates of the existing capacities, capacities for new plants, amounts of products (like final, intermediate products and auxiliary material), different types of costs, use of feedstock materials, further resources, and profit.

The strategic planning problem than consists in an evaluation procedure of efficient alternatives of utilization of the existing capacities, the possibilities of extension and the installation of new capacities to meet the projected demand of the inner market and the export. The criteria are of different economic nature and environmental aspects.

5. Results and conclusions

The use of the IMPROVE package for the analysis of the above formulated decision problems has shown that the approach and the current implementation is rather flexible. We will demonstrate this statement by three cases of analyzed alternatives. Figure 5 sows the result of a typical optimization where only the profit is maximized. This gives a very one-sided structure of the carbochemical industry which leads for all other criteria to unsatisfactory values and even more is inefficient from an engineering point of view.

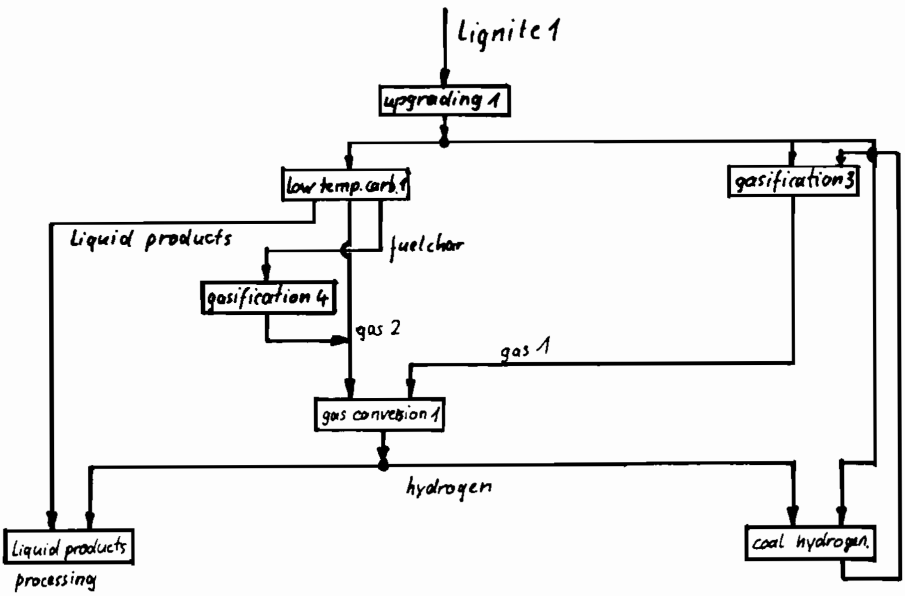


Figure 5. Solution dominated by profit maximization.

The Figures 6 and 7 represent efficient alternatives which are from their structure already broader than the "profit"-solution in Figure 5. The strategy which stands behind the solution of Figure 6 could be called a "conservative" one because in it traditional well-tested technological processes are prefered. In opposite to this, Figure 7 represents a solution where new technological processes have a higher preference.

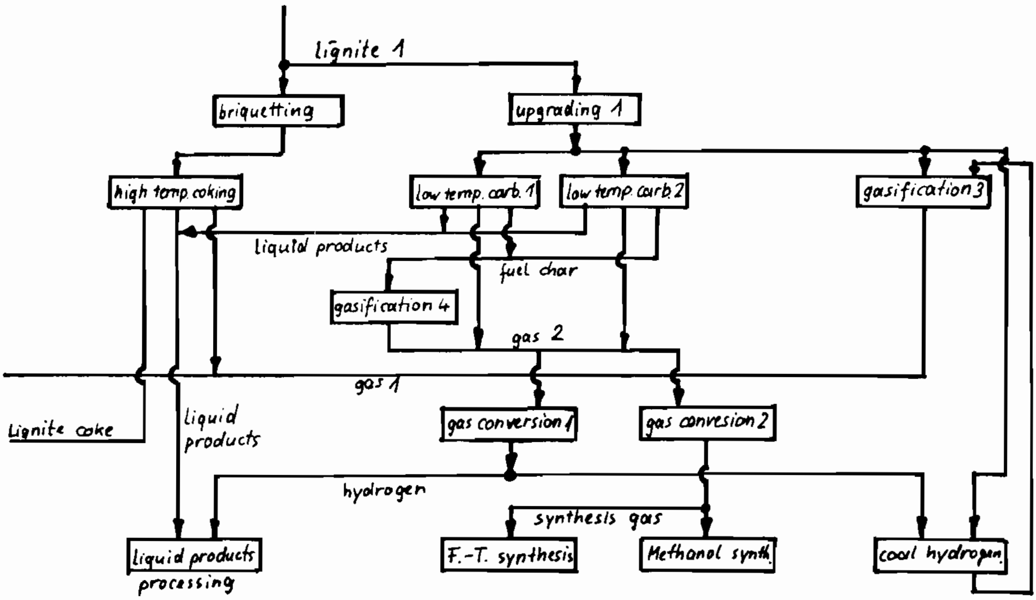


Figure 6. IMPROVE-solution 1 ("conservative" view)

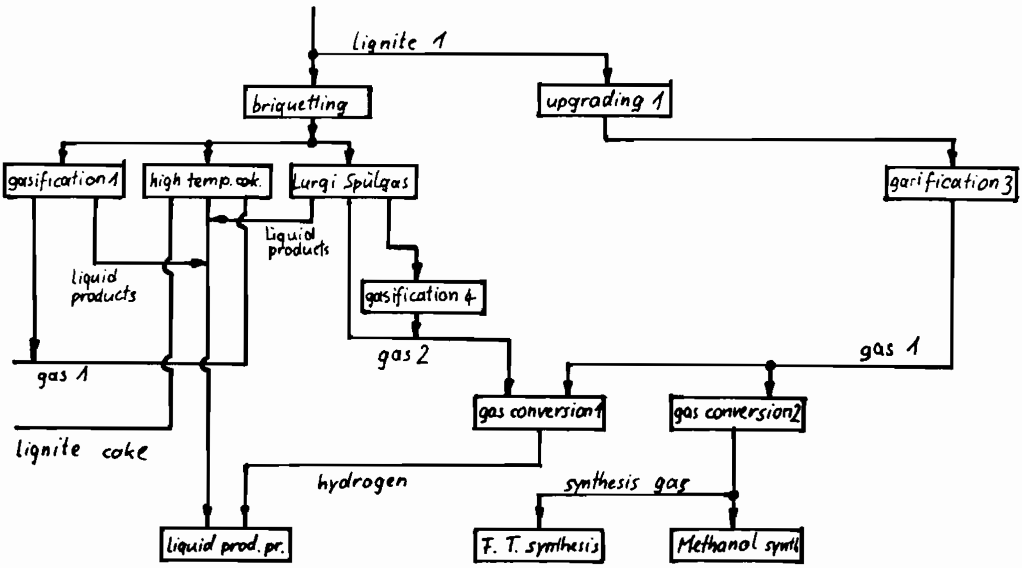


Figure 7. IMPROVE-solution 2 ("progressive view")

The further developments of the IMPROVE-package and the carbochemical industry model are planned in two directions. For strategic planning purposes a "dynamization" and nonlinear descriptions of the process behavior have to be introduced.

To fit the interface better to the user the postprocessor has to be further to be qualified to support more numerical operations on the elements the solution file and to integrate graphical presentations.

References

[1] Grauer, M., Lewandowski, A. and Schrattenholzer, L. (1982). Use of the reference level approach for the generation of efficient energy supply strategies. WP-82-19. International Institute for Applied Systems Analysis, Laxenburg, Austria.

[2] Grauer, M. (1983). A dynamic interactive decision analysis and support system (DIDASS) - user's guide. WP-83-60. International Institute for Applied Systems Analysis, Laxenburg, Austria.

[3] Grauer, M., Lewandowski, A. and Wierzbicki, A (1983). Multiobjective decision analysis applied to chemical engineering, *Applied Systemanalysis* 4 (1), pp. 32-40.

[4] Kallio, M., Lewandowski, A. and Orchard-Hays, W. (1985). An implementation of the reference point approach for multiobjective optimization, WP-80-35, International Institute for Applied Systems Analysis, Laxenburg, Austria.

[5] Lewandowski, A. and Grauer, M. (1982). The reference point optimization approach-methods of efficient implementation, WP-82-26, International Institute for Applied Systems Analysis, Laxenburg, Austria.

[6] Wierzbicki, A.P. (1980). A mathematical method for satisficing decision making. WP-80-30. International Institute for Applied Systems Analysis, Laxenburg, Austria.

EXPERIENCES WITH INTERACTIVE MULTI-OBJECTIVE LINEAR PROGRAMMING METHODS IN ENERGY PLANNING

Matthijs Kok
Department of Mathematics and Informatics
Delft University of Technology
P.O. Box 356, 2600 AJ Delft, The Netherlands

1. INTRODUCTION

Many models have been developed in the last decade to support decision making in the energy sector (e.g. Manne et al. (1979); Lev (1983)). The contribution of all these models to practical decision problems, however, seems to be limited, since today's decision making is a very complicated activity: it involves multiple objectives, various actors and the uncertainties are big. Moreover, the contribution of the models is also often theoretical: the influence of the decision makers in the final results of the models is often not explicitly visible. This paper adresses this last deficiency: we explore the use of interactive methods in a multiple objective model of the energy system in the Netherlands, using preferences of several decision makers, each belonging to different groups and organizations. The decision makers will be called *actors* in the remainder of this paper, since decision making in the government's energy policy is formally the responsibility of the Ministers and the Parliament.

The organization of the paper is as follows. First, we discuss the energy policy issues and the model which has been developed by the Energy Study Centre in order to address these problems. Special attention will be given to the definition of the feasible region in this study. Next, we present results of the energy model with multiple objectives. Interactive multiple objective programming methods are used in experiments to assess the preferences of various actors involved in the energy policy process. Finally, we present an evaluation of these methods and the treatment of uncertainties by the actors.

2. ENERGY POLICY PROBLEMS

Energy policy in the Netherlands is subject of many political debates. Three years ago, a special Steering Group (using funds of the government) organized a Public Debate, concerning among others:

a) The use of nuclear power plants in the future. In this debate, every organization and individual was allowed to express his opinion on the most desirable future society. The conclusion of this debate, now almost two years ago, was not to build new nuclear reactors. This did not end the policy debate, since many organizations feared economical losses (especially in the international product-markets) if nuclear power was ruled out. There are more conflicting views and objectives, however:

b) The utilization of the big indigenous natural gas resources. Here the conflict concerns low depletion rates (a conservation policy to save gas for the generations in the next century) and high depletion rates (to reduce the deficits on the budget of the government).

c) The reduction of acid-rain pollutants SO_2 and NO_x and fossil fuels. The environmentalists argue that one should take care in polluting the air, water and soil: if it is not reduced now, there will be very high future costs. Another point of view is that the anti-pollution measures are very costly and one should therefore not reduce the pollution too much. Moreover, there exist many uncertainties with respect to the effects of these pollutants and the anti-pollution measures, so it can be better to follow the international standards.

It is clear that we have a long-term problem here with multiple objectives and many uncertain elements. Because of uncertainties, some of the costs can not easily be included in the cost-function (e.g. the future costs of nuclear waste and the acid-rain pollutants, and the economical value of future natural gas reserves are unknown). Hence, it is quite natural to formulate this problem with multiple objectives.

3. THE ENERGY MODEL

In order to investigate the policy-problems sketched above, a static Linear Programming model (SELPE) has been developed by the Energy Study Centre in Petten (Boonekamp (1982)). It is an energy-supply model: the final demand of the sectors of the economy is given exogeneously. Thus, it is a partial equilibrium model - there is at this moment no feedback of the solution (e.g. consumption of natural gas; imports of coal and oil; investments) to a model of the remainder of the economy. Similar models have been developed at Brookhaven, USA (Sailor and Rath-Nagel (1980)) and at IIASA, Laxenburg (Schrattenholzer (1981)).

The SELPE-model can be characterized as a network: the nodes represent processes (e.g. refineries, power plants) and the arcs represent the energy flows (often transportation of energy). This flow is calculated in fysical units, and the conversion processes are described with technological parameters. The equations of this model (e.g. balance-and process-equations in the nodes, bounds on the flows in arcs) define the feasible set. In the last three years, the model has been used in several applications with one objective function: minimization of total costs.

In this study we will adopt the following approach. *First*, the feasible set is defined carefully: no potentially acceptable option is excluded, but mainly because of practical considerations (e.g. computer costs), options which are theoretically sound but highly improbable in practice are excluded as much as possible. No option which may be acceptable for an actor or (interest) group is a-priori excluded! (See also Spronk and Veeneklaas (1983)). Much time was spent to formulate the feasible region of the model in this study. *Second*, we defined various objective functions in order to compare different energy policies. The energy flow and the capacities of the equipment can now be optimized with respect to these functions.

The definition of the feasible region, the projections of the exogeneous variables and the set of objective functions have been discussed with the actors involved in this project. The feasible region which we defined originally was acceptable for all of them; some minor changes, however, in order to enlarge the feasible region were proposed and implemented. Although there are big uncertainties in the projections of the exogeneous variables (e.g. final demand and price of oil, coal, etc.) one scenario for these variables was acceptable. An exception was the price of nuclear energy: several scenarios have been defined for this variable. In this paper we will present results of one scenario; for more details see Kok (1985b). Finally, note that, compared with single-objective models, the use of the optimization-algorithm is strengthened: the feasible region can be explored now. Hence, if the feasible region of the model is very small (e.g. for a case-study Keepin (1984)), multiple objective programming does not make much sense.

4. RESULTS WITH MULTIPLE OBJECTIVE METHODS

In this section we will present the results of some interactive multiple objective methods and the multiple objective energy model. The assumptions and aims of the interactive methods will not be discussed here; see e.g. Chankong and Haimes (1983); Kok (1985a). We present first definitions of some key-concepts in multiple objective linear programming.
The problem of maximizing a linear vector function subject to a number of linear constraints is in general written as

$$\left.\begin{array}{ll} \text{maximize} & Cx \\ \text{subject to} & Ax \leq b, \\ & x \geq 0, \end{array}\right| \qquad (1)$$

where C denotes a (pxn)-matrix and A a (mxn)-matrix. We assume that the feasible set S is non-empty and bounded. Let $\bar{x}^i$ denote a feasible solution which maximizes the i-th objective function regardless of the remaining ones. The ideal vector z^* is the (unique) p-vector with components

$$z_i^* = (c^i)^T \bar{x}^i ,$$

where the n-vector c^i stands for the i-th row of C. the nadir vector n^* is defined as the p-vector with components

$$n_i^* = \min_{j=1,\ldots,p} (c^i)^T \bar{x}^j .$$

Note that the nadir vector is not necessarily unique, and is not necessarily equal to the minimum objective value over the entire efficient set (Isermann and Steuer (1985); Kok and Spronk (1985)). Solutions of the energy model have been obtained by using the mathematical programming package APEX for the CYBER 855 at the Energy Study Centre. We have used the "user option" of this package and we have written a FORTRAN program to control the input-output of the model userfriendly. All methods are carried out in *batch processing* since the Linear Programming model (roughly 600 variables and 500 constraints) is too big to use in the interactive mode. The turnaround-time was always acceptable (less than one quarter of an hour).

4.1. PAY-OFF MATRICES

For a given scenario in the year 2000 (i.e. projections of the exogeneous variables) we calculated the Pay-off matrix P with elements $p_{ij} = (c^i)^T \bar{x}^j$, with $\bar{x}^j$ the solution which optimizes the j-th objective. Table 1 shows us the pay-off matrix in one scenario with relatively cheap nuclear energy (8.7 cts/kwh). We found that some objectives have alternative optima: the pay-off matrix and the nadir vector are therefore not unique. We think that it is not satisfactory that the solutions in a pay-off matrix are randomly calculated from the set of alternative solutions. In order to obtain a uniquely defined pay-off matrix (and also to ensure efficiency of these solutions) we applied *lexicographic* optimization.

In this study, we choose the objective "minimization of total costs" as second in the lexicographic ordering, if objective j, j = 1, ..., p is optimized. In this way we obtained unique solutions in the pay-off matrix. Note that the optimal solutions of some objectives are equal; it does not necessarily mean that some of these objectives are redundant. Figure 1 illustrates the pay-off matrix of table 1: each vertical bar represents an objective function, with the ideal value at the bottom and the nadir value at the top. The j-th level of bar i stands for the value $(c^i)^T \bar{x}^j$, and the j-th broken line connects the respective objective values at $\bar{x}^j$.

4.2. MODEL WITH NINE OBJECTIVE FUNCTIONS

We first applied the Pairwise-Comparison method (for details see Kok and Lootsma (1985)), using the model and the objective functions as described in the preceeding section. The actors are asked to estimate the ratios which are acceptable for deviations from the ideal vector (in the direction of the nadir vector). These ratios are estimated via a Pairwise-Comparison method; accordingly a weighting vector w is calculated using the following ingredients: the qualitative answers of the actors, a conversion of the category scale into a real magnitude scale and logarithmic regression (since the answers are often not entirely consistent, and to calculate the group weights). Thus we minimize in order to estimate the weights $w^{(k)}$ of actor k:

$$\sum_{i<j} (\ln r_{ijk} - \ln w_i^{(k)} + \ln w_j^{(k)})^2$$

where r_{ijk} is the estimate of ratio w_i/w_j by actor k. The group weights are

Objective	value at optimum of objective								
	1	2	3	4	5	6	7	8	
1. min. total costs (Dfl. 10^9)	82.1	84.0	87.7	82.2	83.0	89.9	83.9	82.1	86.0
2. min. SO_2-emissions (1000 ton)	522.7	93.7	113.3	709.7	482.9	186.3	542.1	522.7	192.8
3. min. NO_x-emissions (1000 ton)	495.4	303.4	216.2	572.4	474.2	398.3	526.6	495.4	435.7
4. min. nuclear energy (MW)	5000.0	5000.0	500010	0.0	5000.0	0.0	5000.0	5000.0	5000.0
5. max. renewable energy (PJ)	60.4	60.4	191.2	60.4	191.2	1.3	129.2	60.4	191.2
6. max. profits on gas (Dfl. 10^9)	8.1	13.9	8.0	8.1	7.1	19.7	6.1	8.1	7.6
7. min. natural gas (Dfl. 10^9)	18.1	35.8	18.2	18.1	15.6	53.0	12.4	18.1	17.1
8. min. net oil-imports (10^6 TOE)	17.1	17.1	27.6	17.1	17.5	18.1	18.1	17.1	26.9
9. min. fossil fuels (PJ)	2256.8	2234.5	2114.2	2532.1	2118.8	2608.5	2178.4	2256.8	2051.2

Table 1. Pay-off matrix for the Dutch economy in the year 2000 (price of nuclear energy 8.7 cts/kWh).

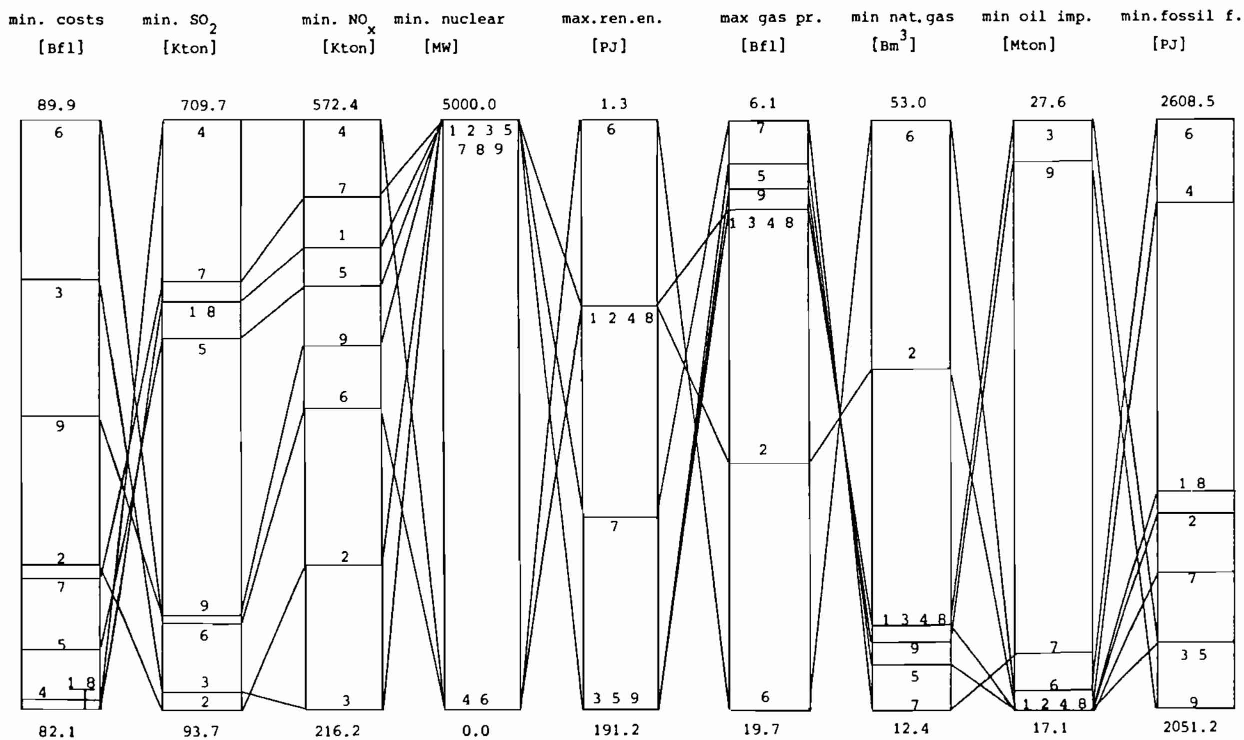


Figure 1. Graphical representation of the Pay-off matrix in table 1.

obtained by minimizing

$$\sum_{k \in D_{ij}} \sum_{i<j} (\ln r_{ijk} - \ln w_i + \ln w_j)^2$$

where D_{ij} stands for the set of actors who have judged objective i versus objective j. Note that it is not necessary to force the actors to judge every pair of objectives.

Next, we use these weights in order to find an efficient solution $\bar{x}$; we solve the following single-objective programming problem:

$$\begin{aligned} &\text{minimize} \quad \{ \max_{i=1,..,p} \lambda_i \, (z_i^* - (c^i)^T x)\} + \sum_{j=1}^{p} \varepsilon_j \, (z_j^* - (c^i)^T x) \\ &\text{subject to } Ax \le b, \\ &\qquad\qquad\quad x \ge 0, \end{aligned}$$

with $\lambda_i = w_i/(z_i^*-n_i^*)$. Table 2 shows us the results in one scenario. Here, preference information is given by five representatives of five groups and organizations: the Federation of Netherlands Industries, the Netherlands Trade Union Federation, The Centre for Energy Conservation, the Ministry of Housing, Physical Planning and Environment, and the Ministry of Economic Affairs. We see that the ratios ρ_i:

$$\rho_i = w_i \frac{z_i^* - (c^i)^T \bar{x}}{z_i^* - n_i^*} \quad ,$$

are not necessarily equal for all i = 1, ..., p. Figure 2 illustrates the six efficient solutions of the energy model in table 2. The j-th broken line connects the solution of actor j.

The results show clearly the conflicts between the various actors: especially the objective "minimization of total costs" and "minimization of nuclear energy" are highly conflicting. The consensus with respect to the acid-rain pollutants (SO_2, NO_x) and the low priority of each actor with respect to the strategic objectives ("minimization of net oil imports") are remarkable. The objectives "minimization of fossil fuels" and "maximization of profits on Dutch natural gas" are also given low priority by each actor. Hence, we decided to carry out a second experiment: because the objectives 6-9 have obtained low priority (see table 2) we dropped these objectives from consideration.

			1			2			3			4			5			group		
	ideal vector	nadir vector	weight	solution	ρ_i	weight	solution	ρ_i	weight	solution	ρ_i	weight	solution	ρ_i	weight	solution	ρ_i	weight	solution	ρ_i
1. Min. Costs (Dfl 10^9)	82.1	89.9	2.3	90.2	2.3	47.7	82.9	4.6	5.8	85.4	2.4	3.7	86.1	1.9	62.2	82.6	4.1	15.8	84.2	4.2
2. Min. SO_2 (Kton)	93.7	709.7	13.5	300.0	4.5	12.6	319.8	4.6	12.6	211.0	2.4	34.6	127.4	1.9	10.5	333.4	4.1	22.1	210.5	4.2
3. Min. NO_x (Kton)	216.2	572.4	10.2	374.1	4.5	11.2	363.3	4.6	12.6	284.0	2.4	14.2	263.7	1.9	10.5	354.8	4.1	17.1	303.5	4.2
4. Min. Nuclear (MW)	0.0	5000.0	34.7	651.5	4.5	0.9	3940.7	0.7	53.4	224.6	2.4	9.1	1039.1	1.9	1.8	5000.0	1.8	11.2	1870.5	4.2
5. Max. Ren. En. (PJ)	191.2	1.3	16.9	191.2	0.0	5.2	40.4	4.1	5.8	165.4	0.7	19.8	191.2	0.0	4.8	60.4	3.3	12.6	128.1	4.2
6. Max. Gas. Pr. (Dfl 10^9)	19.7	6.1	3.0	9.5	2.2	9.0	12.7	4.6	1.9	14.8	0.7	2.2	14.4	0.9	1.4	10.9	0.9	4.0	13.7	1.7
7. Min. Nat. gas (Bm^3)	12.4	53.0	5.3	22.7	1.4	2.4	32.1	1.2	1.4	39.0	0.9	1.9	37.8	1.2	1.3	26.7	0.4	3.1	35.5	1.7
8. Min. Oil Im. (M.ton)	17.1	27.6	3.2	24.6	2.3	4.6	17.2	0.1	2.4	17.1	0.1	10.2	17.3	0.2	4.3	17.2	0.1	6.4	17.2	0.1
9. Min. Fossil F. (PJ)	2051.2	2608.5	10.8	2284.5	4.5	6.4	2322.1	3.1	4.1	2377.2	2.4	4.2	2302.1	1.9	3.1	2274.6	1.2	7.6	2344.8	4.0

Table 2. Individual and group weights, and efficient solutions of the energy model with nine objective functions in the year 2000 (scenario A: price of nuclear energy 8.7 cts/kWh).

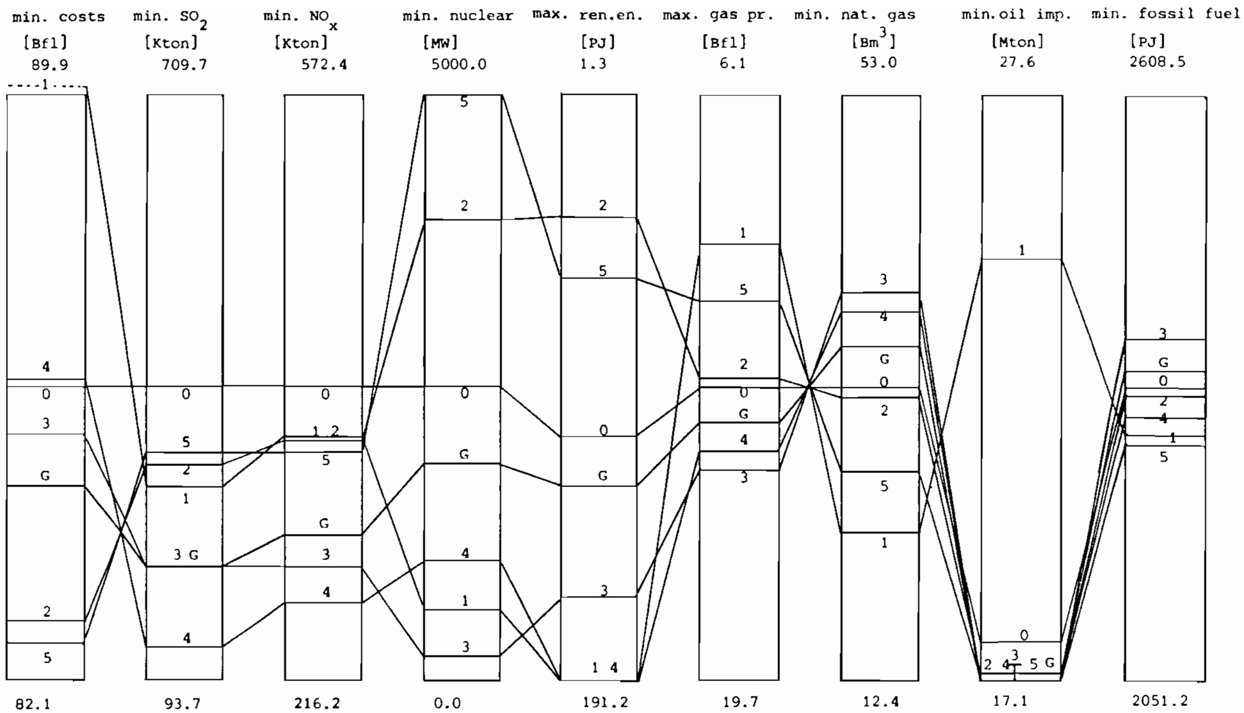


Figure 2. Graphical representation of the solutions of the energy model in table 2. Each broken line represents the solution of actor i; the 0-broken line represents the trial solution with equal weights and the G-broken line is the group-compromise solution.

4.3. MODEL WITH FIVE OBJECTIVE FUNCTIONS

In this second experiment we included only five objectives, and we constrained the acid-rain pollutants to upperbound levels. The following objectives are included:

1. minimization of total costs (Dfl, 10^9),
2. minimization of SO_2 emissions (Kton); upperbound: 350 Kton,
3. minimization of NO_x emissions (Kton); upperbound: 400 Kton,
4. minimization of Nuclear Energy (MW),
5. maximization of Renewable Energy (PJ).

Again we applied the Pairwise Comparisons method. Table 3 shows us the results. The judgments (weights) of the actors in this experiment are, more or less, identical with the judgments in the first experiment: we did not, however, succeed in reducing the divergence of opinions: the same conflicts, but now more distinct, can be discovered. Because the issue of nuclear energy is very polarized in the Netherlands, the group-compromis solution may therefore not be acceptable for all actors.

In the experiment we have used the multiple objective programming method as a one-shot procedure. In order to investigate the influence of the method and the fuzziness of the preferences we applied two other methods to the model with five objectives. Of course, the conclusions which can be drawn from these experiments will be limited, since actors will be biased after application of one method, but, because a one-shot procedure has almost no learning effect, application of other methods is meaningful. Table 4 shows the final results of the Interactive Multiple Goal Programming method (see Spronk (1981)) and the Satisficing Trade-off method (Nakayama and Sawaragi (1984)). The number of iterations in the experiments was limited (maximum of 4), and especially the Satisficing Trade-off method has a good performance in this respect.

5. EVALUATION

With respect to the use of the interactive methods in the experiments with the energy model we remark:

1. There are no big differences in the final solutions of the three interactive methods (table 4).
2. The uncertainties which are inherently involved in the projections of the

			1			2			3			4			5			group		
	ideal vector	nadir vector	weight	solution	ρ_i	weight	solution	ρ_i	weight	solution	ρ_i	weight	solution	ρ_i	weight	solution	ρ_i	weight	solution	ρ_i
1. Min. Costs (Dfl 10^9)	82.4	85.5	2.5	85.8	2.8	44.9	83.1	10.2	7.3	85.1	6.3	3.0	85.4	2.9	78.9	82.6	6.7	18.9	84.0	9.5
2. Min. SO_2 emission (Kton)	93.7	350.0	13.5	146.0	2.8	30.1	180.2	10.2	16.1	194.3	6.3	60.5	106.1	2.9	8.7	290.7	6.7	33.8	165.8	9.5
3. Min. NO_x emission (Kton)	227.5	400.0	12.2	266.5	2.8	16.5	333.7	10.2	19.7	282.8	6.3	18.2	255.2	2.9	8.7	360.1	6.7	24.0	295.8	9.5
4. Min. Nuclear Energy (MW)	0.0	5000.0	60.7	227.0	2.8	1.0	5000.0	1.0	53.6	589.4	6.3	10.0	1460.7	2.9	1.4	5000.0	1.4	14.3	3323.6	9.5
5. Max. Renewable Energy(PJ)	191.2	60.4	11.1	169.9	1.8	7.4	77.8	6.4	3.3	151.0	1.0	8.2	144.6	2.9	2.2	60.4	2.2	9.0	121.2	4.8

Table 3. Individual and group weights, and solutions of the energy model with five objective functions in the year 2000.

	1			2			3			4			5		
	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
1. Min. Costs (Dfl 10^9)	85.8	86.0	85.9	83.1	83.0	83.0	85.1	86.4	85.9	85.4	84.9	84.0	82.6	83.1	83.4
2. Min. SO_2 emissions (Kton)	146.0	120.0	156.5	180.2	300.0	313.1	194.3	125.0	161.5	106.1	100.0	151.2	290.7	220.0	212.1
3. Min. NO_x emissions (Kton)	266.5	250.0	252.9	333.7	350.0	357.8	282.8	250.0	276.2	255.5	300.0	352.6	360.1	350.0	323.6
4. Min. Nuclear Energy (MW)	227.0	500.0	557.3	5000.0	3812.0	3190.3	589.8	0.0	0.0	1460.7	1000.0	1021.2	5000.0	4000.0	3750.1
5. Max. Renewable Energy (PJ)	169.9	162.8	158.9	77.8	100.0	94.2	151.0	184.0	190.9	144.6	150.0	149.1	60.4	77.8	121.2

Table 4. Comparison of solutions of energy model with five objective functions, obtained via three different methods (A = Pairwise Comparison, B = Interactive Multiple Goal Programming, C = Satisficing Trade-off).

exogeneous variables in this application, result in neglecting the outcomes of the model. For example, if the model shows that the nuclear power plants reduce the costs, this is not taken seriously by those actors who expect high nuclear energy prices in the future. In order to investigate this, we organized another experiment: we asked the actors to assess their subjective probabilities with respect to the expected production costs of nuclear energy. Table 5 shows the results. It is clear that the expected nuclear costs

production costs	actor 1	2	3	4	5
6.0 - 9.0	0.10	0.40	0.05	0.25	0.70
9.0 - 12.0	0.35	0.35	0.15	0.50	0.25
12.0 - 15.0	0.35	0.20	0.30	0.20	0.04
15.0 - 18.0	0.20	0.05	0.50	0.05	0.01

Table 5 Subjective Probabilities of five actors with respect to production-costs of a 1000 MW - nuclear power plant (cts/kWh, 1982) in the year 2000.

are related to the judgements concerning the use of nuclear energy. In figure 3 we have plotted the mean of the expected nuclear energy costs versus the mean of installed nuclear power plants in the three interactive methods. *It is not clear which variable explains the other*. Careful treatment of uncertainties in the uncontrollable variables is therefore essential in long-term planning (Leemhuis (1985); Dijk and Kok (1985)).

3. It is surprising that there is much consensus concerning many policy-issues: conservation of the indigenous natural gas stocks is assessed as less relevant, and constraining the acid-rain pollutants is given high priority by the actors.

4. The main conflict in energy policy with respect to the use of energy technologies is the contribution of nuclear energy.
The priorities of acid-rain pollutants are partly responsible for the group compromise solution of 2500-3000 MW nuclear energy, where the individual compromise solutions are in the range 0-5000 MW.
It is surprising that two actors prefer a constrained use of nuclear energy of, say, 500 -1000 MW (table 4) because in the national debate only the question "nuclear energy: yes or no" seems to be relevant. These

Nuclear energy: Costs vs. Capacity

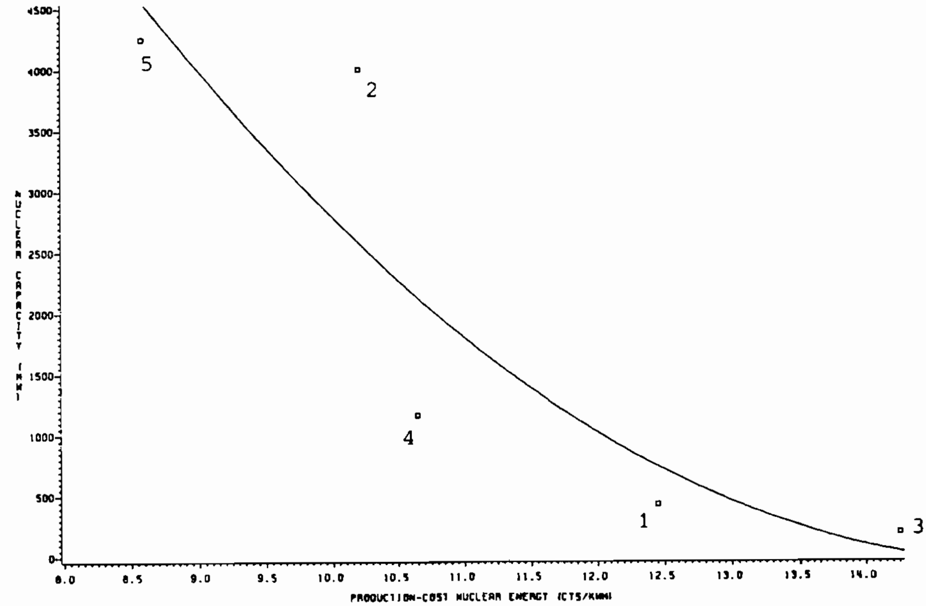


Table 3. Expected production costs of nuclear energy [cts/kWh] versus preferred nuclear capacity [MW] in the year 2000 assessed by the five actors. The regression line is also included.

actors argue that a not restricted use of nuclear energy will damage the energy conservation objective. Finally, the group-compromise solution is rather close to a recent government proposal.

The final conclusion concerning the use of interactive multiple objective methods in long-term (strategic) planning is that they can be succesfully applied in order to obtain insight in the latitude of the policy problem, and to generate explicitly preferences and value-judgements of the various actors involved in the planning-policy process. There is empirical evidence that the influence of the chosen interactive method is acceptably small.

ACKNOWLEDGEMENT

It is a great pleasure to acknowledge the valuable support of professor F.A. Lootsma (Delft University of Technology) and F. van Oostvoorn (Energy Study Centre, Petten). This research was supported by the Ministry of Economic Affairs, The Hague.

REFERENCES

Boonekamp, P.G.M. (1982). Beschrijving van SELPE, een model van de Nederlandse energievoorziening, ESC-17, Energy Study Centre, Petten.

Chankong, V. and Haimes, Y.Y. (1983). Multiobjective Decision Making: Theory and Methodology, North Holland, Amsterdam.

Dijk, D. and Kok, M. (1985). Strategic planning and energy models. Report 85-21. Reports of the Department of Mathematics and Informatics, Delft University of Technology, Delft, Netherlands.

Isermann, H. and Steuer, R.E. (1985). Pay-off tables and minimum criterion values over the efficient set, Report 85-178, College of Business Administration, University of Georgia, Athens, Georgia, USA.

Keepin, B. (1984). A technical appraisal of the IIASA energy scenarios, Policy Sciences 17, 199-276.

Kok, M. (1985a). Interface with Decision Makers and some experimental results in interactive Multiple Objective Programming methods, Report 85-36. Reports of the Department of Mathematics and Informatics, Delft University of Technology, Delft, Netherlands. (to appear in European Journal of Operational Research).

Kok, M. (1985b). Multiple Objective Energy modeling: Experimental results with interactive methods in the Netherlands. Report 85-49. Reports of the Department of Mathematics and Informatics, Delft University of Technology, Delft.

Kok, M. and Lootsma, F.A. (1985). Pairwise-comparison methods in multi-objective programming, with applications in a long-term energy-planning model. European Journal of Operational Research, 22, 44-55.

Kok, M. and Spronk, J. (1985). A note on the pay-off matrix in multiple objective programming. Omega (to be published).

Leemhuis, J. (1985). Using Scenarios to develop strategies. Long Range Planning 18, 30-37.

Lev, B. (ed.) (1983). Energy models and Studies, North Holland Publishing Company, Amsterdam.

Manne, A.S., Richets, R.G., and Weynant, J.P. (1979). Energy Policy Modelling: a survey. Operations Research 27, 1-36.

Nakayama, H. and Sawaragi, Y. (1984). Satisficing Trade-off method for multi-objective programming and its applications, 9th IFAC World Congress, 2-6 July 1984, Budapest, Hungary.

Sailor, V.L. and Rath-Nagel, St. (1980). MARKAL - A computer model designed for multi-national energy systems analysis. Seminar on Modelling Studies, Economic Commission for Europe, Washington.

Schrattenholzer, L. (1981). The energy supply model MESSAGE, RR-81-31, IIASA, Laxenburg, Austria.

Spronk, J. (1981). Interactive Multiple Goal Programming: Applications to Financial Planning, M. Nijhof, Boston.

Spronk, J. and Veeneklaas, F. (1983). A feasibility study of economic and environmental scenarios by means of Interactive Multiple Goal Programming, Regional Science and Urban Economics 13, 141-160.

DECISION-SUPPORT MODEL SYSTEM FOR LONG-TERM WATER MANAGEMENT IN OPEN-PIT LIGNITE MINING AREAS

Stefan Kaden
IIASA, A-2361 Laxenburg, Austria

1. Introduction

Open-pit *lignite* mining creates significant impacts on the environment as well as conflicts within the socio-economic system in such areas. Generally any open-pit mining activity affects both groundwater and surface water resources. These problems concern especially countries in Central and Eastern Europe, in particular, the GDR, FRG, CSSR, and Poland, but also in Asia and the USA.

In the German Democratic Republic more than 70% of the total output of primary energy is based on lignite extracted exclusively by open-pit mining. The annual output of lignite amounts to more than 300 million tons/annum. To satisfy the geomechanical stability of the slopes and the bottom of the open-pit mines it is necessary to pump out 1.8 billion m^3/annum water for dewatering of the open-pit mines. Compared with the stable runoff of the GDR of about $9.0 \cdot 10^9$ m^3/annum the amount of mine water exceeds 20% of the stable runoff of the whole country. Hence, in the mining area significant environmental and resource use conflicts between different water users are created. The most important interest groups are mining, municipalities, industry, in many cases located downstream, and agriculture. Recreation and environmental protection are conflicting interests too. A detailed description of the major impacts of open-pit mining on the environment is given in Kaden et al. (1985a,b).

The activities of each of the interest groups modify more or less the water resources system and at the same time the conditions for resources use by other groups. It is also important that these activities might lead to a deterioration of the natural environment.

There is an apparent need for the development of methods and models supporting the analysis and implementation of long-term regional water policies, to reconcile the conflicting interests in open-pit lignite mining areas, to achieve a proper balance between economic welfare and the state of the environment. Focusing on these goals within the Regional Water Policies Project at the International Institute for Applied Systems Analysis in collaboration with research institutes in the GDR and in Poland, the prototype decision support model system DSS MINE has been elaborated combining interactive software with computer graphics. It is implemented for a GDR test area that includes typical water-related elements of mining regions, significant conflicts and interest groups.

2. Methodological Approach

2.1. General Structure of the DSS MINE

The analysis of regional water policies in mining regions is a problem of dynamic multi-criteria choice. It is embedded in a complicated policy making process. An advanced system of decision aids is needed which allows:

- to consider the controversy among different water users and interest groups,
- to include multiple criteria some of which can not be evaluated quantitatively,
- to take into the account the uncertainty and the stochastic character of the system inputs as well as the limited possibilities to analyze all the decisive natural and socio-economic processes and impacts,
- to offer a set of decision alternatives, demonstrating the necessary trade-offs between different water users and interest groups.

At present, no mathematical method is available or practical applicable considering all these problems in one single model. Only hierarchical model systems can satisfy all requirements.

In general, dynamic problems of regional water management are approached by time-discrete dynamic systems models. The step-size and the available mathematical methods are the structural factors of the necessary model hierarchy. Frequently already a two-level model hierarchy satisfies most requirements. For the DSS MINE such a two-level system has been realized, combining a first-level *Planning Model* with a second-level *Management Model*, see for details Kaden et al. (1985a,b) and Kaden (1986b).

The first-level *Planning Model* realizes a dynamic multi-criteria analysis for a relatively small number of *planning periods* as the time step for principal management/technological decisions. The time step depends on the variability in time of relevant processes, on the required criteria and their reliability, and on the frequency of decisions. As a compromise between accuracy and both, data preparation and computational effort, for the DSS MINE variable time steps are used, starting with one year and increasing with time up to 15 years. This has been done taking into account the uncertainties in predicting model inputs and the required accuracy of model results, decreasing with time.

The planning model serves for the estimation of rational strategies of long-term systems development. These strategies are selected by multi-criteria analysis considering a number of *criteria*. The criteria have to be chosen from a given set of *indicators*, e.g. cost of water supply, cost of mine drainage, satisfaction of water demand and environmental requirements. These indicators are assumed to be integral values over the whole planning horizon. With the purpose of a unified model being independent on the chosen criteria all indicators are bounded and treated as constraints.

The reference point approach is used for multi-criteria analysis, Wierzbicki (1983). Starting from aspiration levels (reference point) of the decision makers for the indicators of systems development efficient responses are generated (non-dominated solutions "closest" to the reference point). The best suited solution can be obtained by correcting the aspiration levels in an interactive procedure. A description of the problem solver for multi-criteria analysis is given in Kaden, Kreglewski (1986).

The planning model is based on a series of more or less strong simplifications in order to obtain a manageable system being suitable for multi-criteria analysis. The major simplifications are:

- the discretization of the planning horizon into a small number of planning periods; all model data, e.g. decisions, state variables, are assumed to be constant within the planning period,
- the neglect of uncertainties in model inputs,
- the application of simplified environmental submodels based on comprehensive models,
- the neglect of relevant environmental subprocesses as the interaction between groundwater and surface water depending on the surface water table.

That is why a second-level *Management Model* for the simulation of systems behavior for a larger number of smaller *management periods* (monthly and yearly time steps) is applied. It is used to analyze managerial decisions by the help of stochastic simulation and to verify results obtained with the planning model.

The planning model considers principal management/technological decisions for estimated input values (expectation values). The feasibility of the estimated decisions is checked only in the mean for planning periods. Problems arise if the principal decisions are superimposed by managerial decisions for shorter time intervals, depending on the actual partly random systems development. This is especially typical for water demand/supply. Both, the models for the actual water demand, and for the available water resources have to consider autocorrelated and random components. The water demand has to be satisfied according to its variations between and within years.

By the help of stochastic simulation based on the Monte Carlo method the feasibility of strategies estimated in the planning model is verified and the strategies are statistically evaluated. Details on the management model can be found in Kaden et al. (1986).

2.2 Design of the Model System

In the past years rapid development in electronic data processing has opened up completely new possibilities for model applications in practical decision making for regional water management and planning. In developing such model system one should not try to replace real-world policy analysis and decision making. The systems should be designed to support it effectively. To be accepted and used any Decision Support Model System must fit in the decision making reality, it has to be user-friendly, reliable, robust and credible.

The development of the DSS MINE for the analysis of regional water policies in open-pit lignite mining areas was directed towards those goals. With the above described methodological approach the policy making reality is well reflected. The model system focuses on the necessary decisions and common criteria for regional water management. The underlying time discretization corresponds to the common planning and management practice. The reference point approach is well suited for an interactive multi-criteria analysis and it is close to engineering-like policy making.

Especial attention has been directed towards the interactive and user-friendly model handling and data management. The following aspects have been considered:

- hierarchical data base for input as well as output data with a robust screen oriented data display and editing system,
- menu-driven model control,
- style and language of model use according to the planning and decision making reality,
- use of computer colour graphics for visual display of computational results.

Below the practical realization of these aspects is illustrated for the implementation of the DSS MINE for a GDR test area, see also Kaden (1986a).

The DSS MINE has been developed in FORTRAN 77 for the computer VAX 11/780 with an AED-512 colour graphics system. It is modular structured (210 modules). Adaptations to different practical problems are possible with low effort as long as the dimension and the character (mathematically) of the problems are adequate. A detailed description of the DSS MINE is given in Kaden (1986b).

3. Application of the DSS MINE for a GDR Test Area

The DSS MINE has been implemented for a test region in the German Democratic Republic. It is an about 500 km² large area in the Lusatian Lignite District. A detailed description is given in Kaden et al. (1985a).

In Figure 1 an overview of the major steps of model application is depicted. The figure illustrates the possible choices in the course of interactive model use and the related terminal output. The DSS MINE is mostly self-explaining. All alternatives are verbally displayed and menu-controlled.

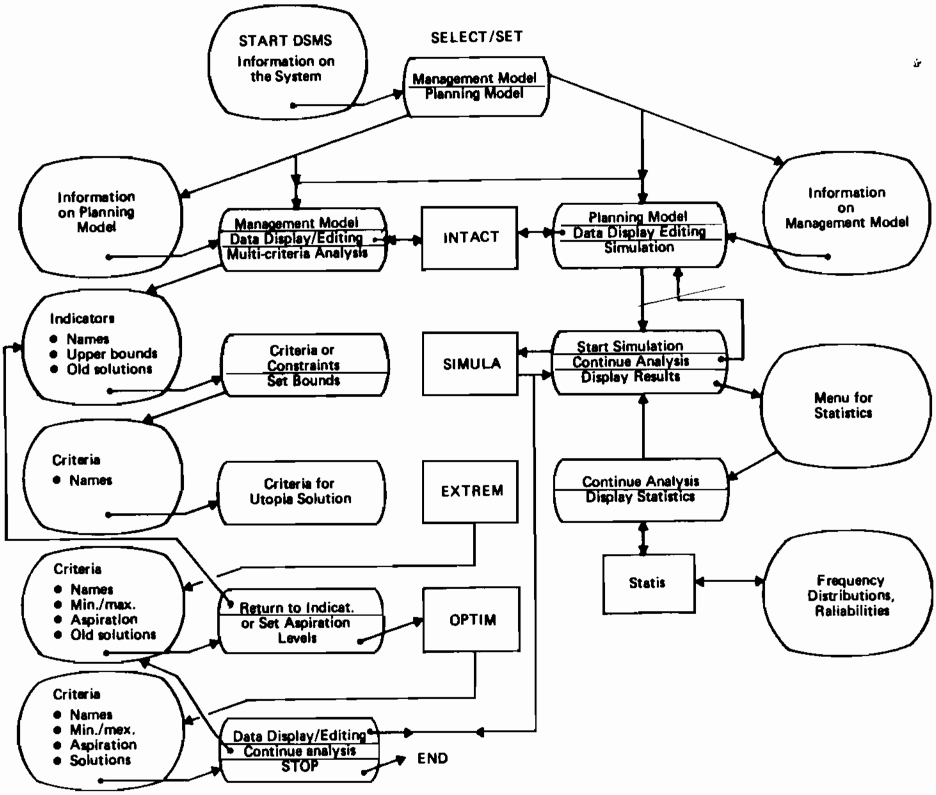


Figure 1: Interactive use of the DSS MINE

A detailed description of all features of the system is impossible in this paper. In the following some major aspects will be illustrated. Thereby outputs of both, an alpha-numeric terminal, and a colour graphic terminal are depicted, unfortunately mono-chrome only. Due to the "page"-oriented display (spread-sheet type) in stead of the common serial output the demonstration of terminal outputs in serial form in a paper is difficult. Originally the "pages" are often not changed, only corrected, for instance, by a newly typed number. In the following for better understanding the full "pages" are listed again after any change. In order to save space, the outputs are condensed. A new "page" at the terminal is signed by ***********. For some outputs revers mode is used; this is signed here by underlining.

The analysis can be started with a prepared data set without any additional data input, but all data can be displayed and edited interactively as it will be demonstrated below.

After starting the system the following output is displayed:

```
******************************************************************************
            *** MINWAT ***  INTERACTIVE DECISION SUPPORT SYSTEM
         DESIGN OF REGIONAL WATER POLICIES IN LIGNITE MINING AREAS

  INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS, Laxenburg  Austria
                 --  Regional Water Policies Project  --

 Implementation:      GDR MINEWATER STUDY, stage 1 for the test area.


 To GET  Full  graphics                          type  F      ;
 To GET  Short graphics                          type  S   .  ;
 To OMIT color graphics                          press RETURN ;
 ...........................  ..................  type  ANSWER  :F
******************************************************************************
```

The user has to decide on the amount of colour graphic output. Only significant answers are accepted, i.e. 'F', 'S' or 'RETURN'. Any wrong answer, e.g. wrong characters, numbers or any garbage, is neglected. This dialogue principle holds true for the entire system. If numbers have to be typed a series of data checks (mostly minimum-maximum tests) is realized. Consequently even unexperienced users can apply the system without unexpected crashes.

After typing 'F' on the colour terminal a series of colour graphics is displayed. First a simplified map of the whole test region is shown, see Figure 2. In this map all elements of the regional water system are depicted. But not all of these elements are significant for policy-making. Some of the elements are either not significant or due to some reason not to be controlled. (In another figure the considered elements of the regions can be displayed).

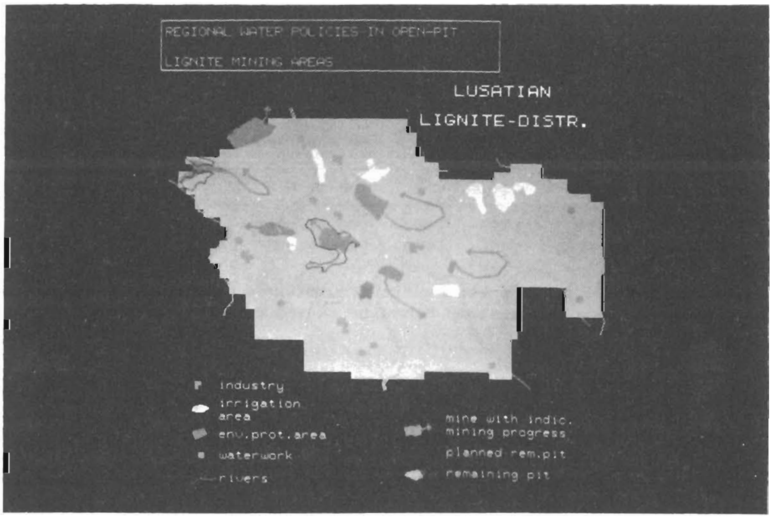


Figure 2: Schematic map of the GDR test area

The next picture shows a scheme of the test area, Figure 3. The thickness of lines indicates the water quantity, i.e. the flux through the stream and its tributaries and through the allocation pipe lines.

The colour indicates the water quality coordinated to given ranges of the pH-value and the Fe^{2+}-concentration, varying between 'excellent' and 'very bad'. The same scheme is used in a reduced size to display model results, see Figure 4.

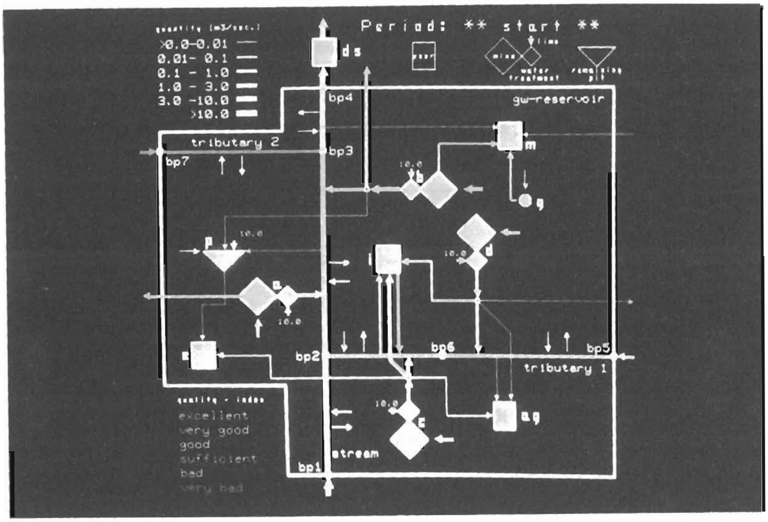


Figure 3: Scheme of the test area

After this graphical information the user has to select the submodel he wants to run:

```
*****************************************************************************
          *** MINWAT ***  INTERACTIVE DECISION SUPPORT SYSTEM
       DESIGN OF REGIONAL WATER POLICIES IN LIGNITE MINING AREAS

  INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS, Laxenburg  Austria
                 -- Regional Water Policies Project --

 Implementation:       GDR MINEWATER STUDY, stage 1 for the test area.
 Components of the Model System:

 (1) PLANNING MODEL for the analysis of the planning horizon 1981 - 2030
     with the planning periods:
       1      2      3      4      5      6      7      8      9      10
     ----------------------------------------------------------------------
     1981   1982   1983-  1985-  1987-  1989-  1991-  1998-  2006-  2016-
                   1984   1986   1988   1990   1997   2005   2015   2030

 (2) MANAGEMENT MODEL for the stochastic simulation and statistic
     evaluation of monthly systems behavior in the planning periods.

  To SELECT a model                                 type  NUMBER  ;
  ... ....................... ...................  type  VALUE   :1
*****************************************************************************
```

Having started the planning model by typing '1' the user selects between:

```
*****************************************************************************
  CURRENT MODEL SELECTION:                       LONG-TERM PLANNING MODEL


                          MENU FOR PLANNING MODEL


 To DISPLAY/EDIT model data                       type  D        ;
 To GET Informations on the planning model        type  I        ;
 To RETURN to MANAGEMENT MODEL                    type  R        ;
 To CONTINUE with analysis                        press RETURN   ;
 . ................. . .. ... ..... ...........  type  ANSWER   :D
*****************************************************************************
```

In the data display/editing mode 'D' the user gets access to the hierarchical data base for input as well as output data. The topics or data are selected by their number. Data are changed by defining the number of row and if required of column, and by typing the new value. The handling is depicted below:

```
********************************************************************************
 CURRENT MODEL SELECTION:                               LONG-TERM PLANNING MODEL

                TOPICS

 ( 1)           INDICATORS       of systems behavior
 (-2)           DECISIONS        on systems development
 ( 3)           DESCRIPTORS      of systems development
 ( 4)           SOCIO-ECONOMICS for systems development

 To SELECT a topic for DISPLAY/EDITING              type  NUMBER  ;
 To RETURN to GENERAL INFORMATIONS                  type  R       ;
 To CONTINUE with analysis                          press RETURN  ;
 ............................................... type  ANSWER  :1
********************************************************************************
 TOPIC SELECTED:                                                  INDICATORS

        SUB-TOPICS

 ( 1)   WATER DEMAND-SUPPLY      deviation
 ( 2)   ENVIRONMENTAL CRITERIA  (water quality index)
 ( 3)   ECONOMICS MINE DRAINAGE
 ( 4)   ECONOMICS WATER SUPPLY
 ( 5)   ECONOMICS ENV. PROTECT.

 To SELECT a sub-topic for DISPLAY/EDITING          type  NUMBER  ;
 To RETURN to GENERAL TOPICS                        type  R       ;
 To CONTINUE with analysis                          press RETURN  ;
 ............................................... type  ANSWER  :2
********************************************************************************
 TOPIC SELECTED:                                        ECONOMICS WATER SUPPLY

          [Mill.Mark]
                                          upper bound    sol-1     sol-2     sol-3

 ( 1) cost-m  : MUNICIPALITY                 900.00 |     28.12     28.12     28.12
 ( 2) cost-i  : INDUSTRY                    2500.00 |   1304.37   1304.17   1157.31
 ( 3) cost-ag : AGRICULTURE                   40.00 |      7.40      7.44      7.46

 To EDIT upper bound of a topic                     type  NUMBER  ;
 To DISPLAY Time series                             type  T       ;
 To RETURN to          INDICATORS                   type  R       ;
 To CONTINUE with analysis                          press RETURN  ;
 ............................................... type  ANSWER  :2
********************************************************************************
 TOPIC SELECTED:                                        ECONOMICS WATER SUPPLY

          [Mill.Mark]
                                          upper bound    sol-1     sol-2     sol-3

 ( 1) cost-m  : MUNICIPALITY                 900.00 |     28.12     28.12     28.12
 ( 2) cost-i  : INDUSTRY                    2500.00 |   1304.37   1304.17   1157.31
                                            -------
 ( 3) cost-ag : AGRICULTURE                   40.00 |      7.40      7.44      7.46

 To EDIT upper bound of a topic                     type  NUMBER  ;
 To DISPLAY Time series                             type  T       ;
 To RETURN to          INDICATORS                   type  R       ;
 To CONTINUE with analysis                          press RETURN  ;
 ............................................... type  VALUE   :2000
```

```
*****************************************************************************
 TOPIC SELECTED:                                        ECONOMICS WATER SUPPLY


          [Mill.Mark]
                                          upper bound      sol-1     sol-2     sol-3

 ( 1) cost-m  : MUNICIPALITY                900.00 |      28.12     28.12     28.12
 ( 2) cost-i  : INDUSTRY                   2000.00 |    1304.37   1304.17   1157.31
                                           -------
 ( 3) cost-ag : AGRICULTURE                  40.00 |       7.40      7.44      7.46

 To EDIT upper bound of a topic                        type  NUMBER  ;
 To DISPLAY Time series                                type  T       ;
 To RETURN to          INDICATORS                      type  R       ;
 To CONTINUE with analysis                             press RETURN  ;
 .............................................         type  ANSWER  :'RETURN'
*****************************************************************************
```

In the illustrated way all input data can be specified by the user, e.g. the bounds for decisions or indicators.

After starting the multi-criteria analysis a list of all indicators considered is displayed. The user has to define the criteria for multi-criteria analysis, the remaining indicators are treated as constraints with an upper bound. The upper bound can be changed interactively. In the given test run the deviation between municipal as well as industrial water demand and supply, the total cost for mine drainage and the cost for water supply have been selected as criteria (signed by '*').

```
*****************************************************************************
   SPECIFICATION of  INDICATORS  as  CRITERIA (*)   or   CONSTRAINT  ( )

                                             upper bound     sol-1    sol-2    sol-3
 ( 1)    dev-m    : MUNICIPALITY          *     10.00  |     16.20    16.21    16.20
 ( 2)    dev-i    : INDUSTRY              *    100.00  |      3.54     3.68    17.52
 ( 3)    dev-ag   : AGRICULTURE                 50.00  |     22.40    21.50    19.97
 ( 4)    dev-e    : ENV. PROTECT.               50.00  |     27.98    24.93    25.24
 ( 5)    dev-ds   : DOWN-STREAM                 0.     |     -8.42    -8.40    -8.41
 ( 6)    c-ds     : Quality DOWN-STREAM          3.00  |      0.80     0.80     0.80
 ( 7)    c-e      : Quality ENV. PROT. A.        3.00  |     -0.96    -0.96    -0.96
 ( 9)    cost-a   : MINE A                    1900.00  |    216.96   216.96   259.19
 (10)    cost-b   : MINE B                    2500.00  |   1496.95  1493.58  1614.40
 (11)    cost-c   : MINE C                    1900.00  |    653.54   654.24   776.38
 (12)    cost-d   : MINE D                    1900.00  |    712.33   712.31   751.53
 (13)    cost-mi  : TOTAL DRAINAGE COST   *   6000.00  |   3079.78  3077.10  3401.48
 (14)    cost-m   : MUNICIPALITY          *    900.00  |     28.12    28.12    28.12
 (15)    cost-i   : INDUSTRY              *   2000.00  |   1304.37  1304.17  1157.31
 (16)    cost-ag  : AGRICULTURE                 40.00  |      7.40     7.44     7.46
 (17)    cost-e   : ENV. PROTECT. AREA          40.00  |      3.30     3.36     3.35

 To SELECT an indicator as CRITERIA                    type  NUMBER  ;
 To SELECT an indicator as CONSTRAINT                  type  -NUMBER ;
 To CONTINUE with multi-criteria analysis              press RETURN  ;
 .............................................         type  ANSWER  :
*****************************************************************************
```

In the next step the so-called utopia solution can be estimated for selected criteria (each criteria is optimized separately without considering of the others). Next, the multi-criteria analysis is prepared finally with following display:

```
******************************************************************************
        SPECICFICATION of  ASPIRATION LEVEL for selected INDICATOR

   indicator    minimum  maximum aspiration           sol-1    sol-2    sol-3
 1 dev-m           0.05   239.80      0.05            16.20    16.21    16.20
 2 dev-i           0.09   113.18      0.10             3.54     3.68    17.52
13 cost-mi      2598.45  3371.30   2598.45          3079.78  3077.10  3401.48
14 cost-m         20.93   935.47     20.93            28.12    28.12    28.12
15 cost-i       1205.81  1381.67   1205.81          1304.37  1304.17  1157.31

To CHANGE the ASPIRATION LEVEL of an indicator   type  NUMBER  ;
To RETURN to SELECT CRITERIA/CONSTR.             type  R       ;
To CONTINUE with analysis                        press RETURN  ;
................................................ type  ANSWER  :2
******************************************************************************
```

The user gets informations on the possible minimum and maximum value, as well as the last three solutions for each criteria. He has to define for each of the criteria his aspiration level (the reference point). After typing the number of row the new value has to be entered.

```
******************************************************************************
        SPECICFICATION of  ASPIRATION LEVEL for selected INDICATOR

   indicator    minimum  maximum aspiration           sol-1    sol-2    sol-3
 1 dev-m           0.05   239.80      0.05            16.20    16.21    16.20
 2 dev-i           0.09   113.18      0.10             3.54     3.68    17.52
                                      ----
13 cost-mi      2598.45  3371.30   2598.45          3079.78  3077.10  3401.48
14 cost-m         20.93   935.47     20.93            28.12    28.12    28.12
15 cost-i       1205.81  1381.67   1205.81          1304.37  1304.17  1157.31

To CHANGE the ASPIRATION LEVEL of an indicator   type  NUMBER  ;
To RETURN to SELECT CRITERIA/CONSTR.             type  R       ;
To CONTINUE with analysis                        press RETURN  ;
................................................ type  ANSWER  :50.
******************************************************************************
        SPECICFICATION of  ASPIRATION LEVEL for selected INDICATOR

   indicator    minimum  maximum aspiration           sol-1    sol-2    sol-3
 1 dev-m           0.05   239.80      0.05            16.20    16.21    16.20
 2 dev-i           0.09   113.18     50.00             3.54     3.68    17.52
                                     -----
13 cost-mi      2598.45  3371.30   2598.45          3079.78  3077.10  3401.48
14 cost-m         20.93   935.47     20.93            28.12    28.12    28.12
15 cost-i       1205.81  1381.67   1205.81          1304.37  1304.17  1157.31

To CHANGE the ASPIRATION LEVEL of an indicator   type  NUMBER  ;
To RETURN to SELECT CRITERIA/CONSTR.             type  R       ;
To CONTINUE with analysis                        press RETURN  ;
................................................ type  ANSWER  :'RETURN'
******************************************************************************
```

If the aspiration levels for all criteria are defined the multi-criteria analysis can be started. During the solution intermediate results are graphically displayed. The results are listed in a similar table as that above, but with the new solution. Now the user can display the results in detail in the hierarchical data base as it was illustrated above for the input data. Furthermore results can be displayed graphically. For instance, the results for different planning periods can be compared, see Figure 4, left part.
The values of selected indicators can be displayed in a bar chart, comparing the last three solutions, Figure 4, right part. Time series of selected data can be displayed in the form of step functions.

After the display of model results the multi-criteria analysis can be continued at any of the previous steps, e.g. input data editing, definition of criteria or aspiration levels etc.

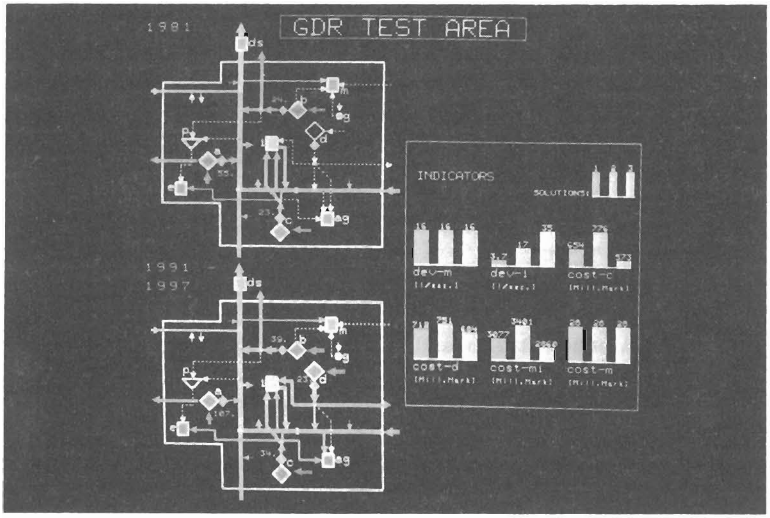


Figure 4: Comparison of results for different planning periods

If the results of the planning model are satisfactory the management model can be started, see Kaden (1986a).

With the DSS MINE a highly interactive model system for the analysis of regional water policies in open-pit lignite mining areas has been developed and implemented.

Although the system has been designed with special regard to lignite mining ares the basic structure and concepts of the system are practicable for other water management problems too.

References

Kaden, S., Hummel, J., Luckner, L., Peukert, D., Tiemer, K. (1985a). Water Policies: regions with open-pit lignite mining (Introduction to the IIASA study). WP-85-4. International Institute for Applied Systems Analysis, Laxenburg, Austria.

Kaden, S., Luckner, L., Peukert, D., Tiemer, K. (1985b). Decision support model system for regional water policies in open-pit lignite mining areas. International Journal of Mine Water, 4(1):1-16.

Kaden, S. and Kreglewski, T. (1986). Decision Support System MINE - Problem Solver for Nonlinear Multi-Criteria Analysis. CP-86-5. International Institute for Applied Systems Analysis, Laxenburg, Austria.

Kaden, S., Michels, I., Tiemer, K. (1986). Decision Support System MINE; the Management Model. CP-86-9. International Institute for Applied Systems Analysis, Laxenburg, Austria.

Kaden, S. (1986a). Interactive Software and Computer Graphics for the Analysis of Regional Water Policies in Open-Pit Lignite Mining Areas. UNESCO/IHP-III Symposium: Decision Support Systems and Related Methods in Water Resources Planning, Oslo, May 1986, Proceedings, in press.

Kaden, S. (1986b). Decision Support System MINE; Description of the Model System, WP-86, forthcoming. International Institute for Applied Systems Analysis, Laxenburg, Austria.

Wierzbicki, A.P. (1983). A Mathematical Basis for Satisficing Decision Making. Mathematical Modeling USA 3:391-405 (Report IIASA RR-83-7).

An Approach for Integrated Energy-Economy Decision Analysis

The Case of Austria

Manfred Strubegger

International Institute for Applied Systems Analysis
Laxenburg, Austria

1. Introduction

Many modeling efforts have been undertaken to gain better insight in the interactions between the energy sector and the rest of the economy. One of the first attempts was made with ETA-MACRO [1], consisting of a macro-economic model linked with an energy model, that had its main focus on electricity generation. One of the latest developments in that field is the ZENCAP model [2]. It is based on a multi-sector economy model consisting of an input-output table and a large number of econometrically estimated equations, e.g. for investments, energy demand and employment.

This paper describes an instrument that falls into the same category of models. However, by reducing the detailed representation of the economic relations and focusing more on an exact description of the energy system, it is easier to handle than the ZENCAP model. This model is designed for the investigation of different energy strategies considered in energy importing countries. The economic submodel serves as a tool for an overall consistency check of the energy scenarios investigated. In general the setup of the model allows to exclude an optional economic sector from the internal logic of the model and investigate it separately in a more detailed model. This separate model is driven by the demands calculated from the economy model. Its results are then used as input to the economy model. For objectives related to the energy system the model system described here supports multi-objective decision analysis [3].

The following section gives an overview of the current energy situation. Thereafter the mathematical formulation of the model is outlined, followed by some preliminary results in the last chapter.

2. The Current Energy Situation

After more than a decade of tight energy markets we are now facing a buyer's market where energy exporters fight for their market position by lowering export prices. This is certainly a situation where research in energy related matters seems to be of no great importance. However, periods like this allow one to prepare for the future. Nobody can predict the development of energy prices over the next few decades or even over the next few months with certainty. Thus various strategies should be tested against conceivable energy price and availability scenarios. Their effectiveness in avoiding future shocks should be proved.

One of the most important factors influencing the development of energy prices are the decisions taken in the oil importing countries. If consumption begins to increase due to the low energy prices and investments in energy conservation are reduced, the current buyer's market may soon turn into a seller's market again and thus allow for price increases. If, on the other hand, the efforts for fuel saving and diversification are continued an energy price increase becomes more unlikely. An ongoing effort in fuel saving can only be imposed by the governments and be enforced by keeping the domestic energy price stable at near the current high levels.

Since the model system was initially developed for Austria, some specific problems of this country will be addressed in the following. Besides the effects mentioned above, where small countries like Austria can only react, a number of internal problems require a careful investigation of the future development of energy demand and supply options. The problem is two-sided. Given the assumption that domestic energy prices will be kept at a high level, then, at least for the next ten to twenty years the growth in energy consumption will be limited and even a decrease in energy consumption due to savings is conceivable. Such a situation requires extreme care when investment decisions are to be taken. In periods with high growth a wrong decision can be paid from future income relatively easily. In low or zero growth periods a wrong investment decision can cause serious financial problem. On the other hand, if consumption increases due to falling energy prices, new power plants have to be taken into operation. But recent developments in public perception of large scale investment projects has hindered the construction or operation of three large power plants. The nearly completed nuclear power station in Zwentendorf was, due to a referendum, not taken into operation. The construction of the base load hydro-power plant near Hainburg and the peak-load hydro-power station in the Dorfer Tal, was--at least--delayed due to strong public protest and the desire to create natural resorts at these locations.

However, the problems are not limited to the public acceptance of new plants. The question arises, wether these projects are at all needed presently to secure Austria's energy supply. Therefore a comprehensive investigation of the future demand situation is required. Related problems, e.g., the environmental damage caused by energy utilization, have to be included in the analysis.

3. The Model

The model developed for analyzing energy-economy interactions on the national level consists of three main modules--an *energy model*, an *economy model* and a *model for private consumption*--that exchange information about physical and money flows. The energy model is a linear programming model. The economy model is a vintage type model consisting of a production and a price module. The model for private consumption is an econometric model calculating private consumption as a function of prices and disposable income. In the following the term economy model stands usually for both the economy model as such and the private consumption model.

3.1. The Energy Model

The energy model is the core module of the model set. It gives a detailed picture of the energy sector and allows the analysis of various energy supply strategies and for the energy demand paths generated by the economy model. Additionally it calculates final energy prices for the energy forms consumed by the end-users. These prices represent the costs of the energy carriers used and the annualized costs of the energy conversion and transportation systems employed.

The energy model is a linear programming model that encompasses all energy flows from energy imports and domestic resource extraction to the final utilization of the different energy carriers. Each energy form i,.e., primary, secondary, final and useful energy, is represented as a node in a physical flow model. Resource balancing and demand constraints ensure that the consumption of an energy carrier never exceeds the amount available.

The links between these nodes, i.e. the energy conversion and transportation technologies are represented as capacities that allow a maximum throughput with a given efficiency. The capital stock can be operated during the entire technical lifetime of a technology and can be increased by new investments.

$$x_{i,t} - \sum_{k=t-\nu_i}^{t} \pi_i y_{i,k} \le hc_{i,t}$$

$x_{i,t}$ production of energy by technology i in year t

$y_{i,k}$ new capacity built in year k

ν_i technical plant life

π_i availability factor

$hc_{i,t}$ historic capacity still available in year t

Various additional constraints are used to make the model behave more realistically, i.e. to avoid 'flip-flop' solutions. Such constraints are, e.g., resource depletion constraints that limit the amount of resources extracted to a given share of the remaining resources.

$$r_t \le \delta_{r,t}(R - \sum_{\tau=1}^{t-1} r_\tau)$$

R total amount of resource at the beginning of the time horizon

r_t extraction of the resource in year t

$\delta_{r,t}$ maximum fraction of the resource, that can be extracted in period t

Dynamic constraints of the form

$$a_{i,t} \begin{matrix} \le \\ = \\ \ge \end{matrix} \gamma a_{i,t-1} \pm g$$

a_t activity, $x_{i,t}$, or new capacity, $y_{i,t}$, installed in year t

γ growth/decline factor

g initial/terminal value

are used to control the annual growth of energy conversion for a technology or annual new investments.

In addition to the energy flows the emission of pollutants is also considered in that model.

$$\sum_{i=1}^{n} e_{i,k} x_{i,t} \le \bar{E}_{k,t}$$

$e_{i,k}$ emission of pollutant k per unit of energy produced in plant i

$\bar{E}_{k,t}$ maximum level of emissions allowed for pollutant k in year t

In the energy model the emissions for SO_2, NO_x, CO, C_xH_y and particulates are considered. Scrubbers, catalysts or filters can be used to reduce the emissions to a prescribed level.

For use in the model set the energy model calculates the investment, intermediate goods and labour requirements, as well as the development of final energy prices and the energy trade balance.

A number of important decision variables are contained in the energy model:

- minimize the costs of supplying adequate amounts of energy to the end-user;
- control the amounts of imported energy and diversify its sources and
- control environmental pollution from energy conversion.

Even if the model is presently only used in cost minimization mode the general set-up, and the implementation, allow for the utilization of multi-objective decision methods using, e.g., the three objectives outlined above. The approach proposed is based on the reference point optimization as developed for the DIDASS system at IIASA [4].

3.2. The Economy Model

3.2.1. Philosophy

The model encompasses the complete economy of a country. For this special application the energy sector was excluded from the internal dynamics of the economy model. Energy prices are then calculated in the energy model and are thus exogenous to the economy model. Investment and intermediate demands of the energy sector are included in the demand vector. The import/export balance as well as the work force employed are taken into consideration.

The model consists of five submodules

- a module for updating the input-output coefficients, based on vintage production theory [5],
- a consumer demand part of the economy module,
- a physical flow module, based on an Input-Output framework,
- a price module, and

- a foreign trade module.

These sub-modules are operated iteratively until a stable solution is found.

3.2.2. The Module for Updating Input-Output coefficients

Based on the vintage production theory each economic sector is depicted by capital stocks, i.e. vintages, that represent plants installed in a particular time period. In each period the existing capital stock is depreciated according to the profitability of each vintage, i.e. the higher the ratio between wages paid and value added produced in that vintage, the more of the capital stock is scrapped.

For the new capital stock an 'optimal' input mix is determined on the basis of labour, capital and intermediate goods prices in that specific period. The actual amount of new investments is then determined in the physical flow module. The current implementation gives two options for selecting new plants

- a Cobb-Douglas function for each economic sector can be used to calculate the input coefficients or
- vintages can be selected from a pre-defined set with given input structures.

The Cobb-Douglas function could easily be substituted by other production functions estimated from past economic behaviour. In order to anticipate trends, that are not deducible from historic trends, the coefficients of such functions can be adjusted over time. The predefined vintages allow for a direct comparison of different processes for producing a specific good, e.g., alternative steel production technologies.

The chosen properties (input factors) of a specific plant cannot be changed at a later point in time when economic realities differ and the plant is no longer optimal. Thus the representation of plants follows the putty-clay hypothesis. The plants can only, as described above, be removed during later periods. Given this representation, the Input-Output table is updated at the end of each period according to the capital stock removed and the capital stock added.

3.2.3. The Physical Flow and the Private Consumption Modules

The physical flow module reflects an Input-Output framework with endogenous formation of capital stock. As described above each sector is represented by a number of vintages reflecting the technology built in a specific period. Unfortunately, the available data did not allow adherence to that procedure for the capital stock existing in the base year (i.e.; vintages built before 1976). Therefore the average capital stock existing in 1976 represents the old capital stock. As time passes the age structure is approximated by parameters that improve the quality (i.e.; labour productivity, capital output ratio, consumption of intermediate goods per unit output) of the historic capital stock over time. This procedure is used to approximate the nonhomogeneity of the existing capital stock.

The mathematical formulation of the physical flow model is:

$$\sum_{j=1}^{ns}(1-a_{i,j}^{o})z_j^{o}+\sum_{j=1}^{ns}(1-a_{i,j}^{n})z_j^{n}-\sum_{j=1}^{ns}k_{i,j}\bar{z}_j-d_i^{p}y = d_i^{g}+e_i+x_i-m_i \qquad (i=1,nc)$$

$$\chi_j^{o}z_j^{o} \le s_j^{o} \qquad (j=1,ns)$$

$$\chi_j^{n}z_j^{n}-\bar{z}_j^{n} \le 0 \qquad (j=1,ns)$$

$$\sum_{j=1}^{ns}(l_j^{o}z_j^{o}+l_j^{n}z_j^{n}) = \bar{L}$$

ns,nc	number of sectors and commodities considered
a_{ij}^{o}	average input coefficients of existing vintages
a_{ij}^{n}	input coefficients of new vintages
z_j^{o},z_j^{n}	production of old and new vintages, respectively
$k_{i,j}$	capital matrix, $\sum_i k_{i,j}=1$
$\bar{z}_j$	investment of sector j
χ_j^{o},χ_j^{n}	capital requirements per unit of output in sector j
s_j^{o}	existing capital stock of sector j
y	disposable income
d_i^{p}	private consumption of commodity i, given as share of the disposable income y
d_i^{g}	governmental consumption of commodity i
e_i	requirements for commodity i of the energy sector
m_i	imports of commodity i
x_i	exports of commodity i
$l_j^{o,n}$	specific labour requirements in sector j for old and new vintages, respectively
$\bar{L}$	total labour force

This system of 3nc+1 equations with 3nc+1 variables is solved as a linear programming problem. It determines the new investments, the disposable income, and the activities of the economic sectors. The formulation as a linear programming problem requires private consumption to be linear in disposable income. Thus a Linear Expenditure System (LES) is used to calculate the demand for consumer goods. A bridge matrix transforms these demands to demands for the commodities produced by the sectors.

$$d_i^c = \gamma_i + \frac{\beta_i}{p_i}\left[y - \sum_{l=1}^{nc} \gamma_l p_l\right]$$

d_i^c	demand for consumer good i
p_i	price for consumer good i, including value added tax
y	disposable income
β_i, γ_i	estimated parameters

If other systems of demand equations are to be used, the solution procedure has to be changed, but apart from this there is nothing preventing consideration of other representations of consumer behaviour.

3.2.4. The Price Module

After solving this system of equations and averaging over the old and new vintages in terms of input coefficients the resulting sector prices can be calculated from

$$p_j^s = \sum_{i=1}^{nc} p_i^c a_{ij} + \Pi_j + \omega_j \bar{l}_j$$

with

$$\Pi_j = \left[\frac{1}{\alpha_j}\right]\sum_{i=1}^{nc} p_i k_{i,j} \bar{z}_j .$$

p_j^s	sector price for sector j
p_i^c	price for commodity i
a_{ij}	average input factor for commodity i in sector j
Π_j	specific profits gained in sector j
ω_j	specific wages paid in sector j
$\bar{l}_j$	average specific labour requirements in sector j
α_j	propensity to invest

The last equation makes use of the simple investment function $I = \alpha\Pi$ to determine the profits that must have been made in order to invest in the capital needed.

3.2.5. The Foreign Trade Module

After the production system is solved the import/export balance resulting from the given prices is compared with the exogenously given trade balance T.

$$T = \sum_{i=1}^{nc} p_i^x x_i + p_i^{c(1.+\varepsilon)} f_i - p_i^m m_i$$

The imports and exports are determined by the estimated functions

$$m_i = m_0 \left[\frac{p_i^c}{p_i^m}\right]^{\varepsilon_i^m} z_i^{\varepsilon_i^z} e^{\gamma_i^m t}$$

$$x_i = x_0 \left[\frac{p_i^c}{p_i^x}\right]^{\varepsilon_i^x} M_w^{\gamma_i^x t}$$

p_i^x	export price for commodity i
p_i^m	import price for commodity i
p_i^c	domestic price for commodity i
f_i	amount of commodity i demanded by tourists
ε	price elasticity of tourists
ε_i^m	price elasticity for imports
ε_i^x	price elasticity for exports
ε_i^z	income elasticity for imports
M_w	forecasted import demand for Austrias trading partners
γ_i^x, γ_i^m	time trend factors
m_0, x_0	used for the adjustment to base year values

Since these equations depict the competition of Austria on the world market, they are very essential for deriving reasonable solutions. An important problem is the different development of the quality of goods imported and exported which cannot be distinguished in the model due to the necessary aggregation of goods. An attempt to solve that problem is made by differentiating between import- and export prices.

4. Linking the Energy and the Economy Modules

The general problem of linking energy and economy models is to find the optimal trade off between an exact and well defined link and a representation that utilizes the information available on the energy sector. A well defined link would be at the level of final energy. All changes taking place between final and useful energy are actually already represented by the Cobb-Douglas functions and the

Linear Expenditure System. The information available is, however, not sufficient to study the development of the energy system. The linking procedure chosen here is of the second type; i.e.; the energy system is represented as completely as possible, down to the level of useful energy. This, of course, no longer represents an exact link, as the final energy demand calculated in the economy model has to be translated to useful energy. This procedure cannot fully reflect the results of the economic equations as this information cannot be extracted from them. As generally the investment directly related to energy equipment is not very large compared to other investments the error is assumed to be rather small.

In the actual linking procedure the economy module calculates the final energy demand, as well as the costs for labour, capital and intermediate inputs needed in the energy sector. The final energy demand is then 'translated' into useful energy using the end-use efficiencies calculated from the previous iteration. The energy module in turn calculates final energy prices and the demand for investment goods, labour and intermediate goods. This information is exchanged between the two models until convergence is achieved. The stopping criteria is, when none of the figures exchanged between the two modules changes by more than half a percent. Convergence is usually achieved after four to six iterations.

5. Data Requirements and Availability

5.1. Data for the Energy Model

The data required for the energy model encompass

- *technical data* regarding the energy technologies represented, like efficiency, maximum availability, flexibility in the input- and output structure, emissions per unit of output, maximum build-up rates, technical plant life,
- *economic data* describing the investment, fixed and variable operation and maintainance costs,
- *structural data* describing the existing capital stock of energy technologies and its age structure, potentials for grid dependent energy carriers and
- *scenario variables* reflecting future development options, like e.g.; acceptance of nuclear power, maximum emissions or import dependence.

The basis for establishing the energy model is a consistent energy balance sheet for the base year of the calculations. According to that information the model is built up as described above and calibrated for that base year. In this special application it is also taken care that labour requirement of the energy sector, as well as the emission pattern in the base year is reflected accurately.

Additionally the demand for investments, split into buildings and equipment, and for intermediate goods, represented as one homogeneous good except energy carriers, are calculated.

The data needed to run the energy model are generally available. Problems are usually related to obtaining cost figures for the different energy conversion systems. Specially investment figures for technologies converting final to useful energy cannot be separated clearly from investments related to the installation of new equipment.

5.2. Data for the Economy Model

The data required for the economy model are qualitatively different from those for the energy model. The availability of data is rather limited in certain aspects, especially as the different data sources use different aggregation levels. The set of the most important data required for each sector considered contains time series for

- capital stocks,
- investments,
- gross production,
- the number of employees,
- wages
- demand for and prices of consumer goods, and
- amounts and prices of imports and exports.

Most of these time series, besides the ones concerning imports and exports, are readily available from official statistics and publications. The time series for imports and exports had to be reaggregated from time series ordered according to the SITC-revised code to ones fitting the aggregation used in the model. Only after that step the necessary import and export functions could be estimated.

The main problem was, that no capital matrices and only two Input-Output matrices are available for Austria (1964 and 1976). Obviously these two snapshots of the economic structure, especially when reported for rather different economic situations, are not sufficient to estimate meaningful factors describing structural change for the different economic sectors. To avoid that problem, a Cobb-Douglas function, for which the coefficients can be calculated from one observation, is used to generate the input coefficients for new technologies. The capital matrix was generated from data on total and sectoral investments in buildings, machinery and transport equipment. The expansion to the sectoral representation used in the

Input-Output framework was done using the structure as reported in the 1976 Input-Output table that distinguishes between these three types of investment goods. No attempt was made so far to dynamize that matrix.

6. Preliminary Results

Preliminary runs show the reaction of the model set to different scenarios for energy import prices. In the first case the energy prices are assumed to remain constant; in the second case the energy import prices increase by some 2% per year above the price of other imported goods. The results derived for primary energy consumption are shown in Table 1.

	1984	1988	1992	1996	2000	2004	1984	1988	1992	1996	2000	2004
Coal	5.91	8.56	9.55	10.41	10.75	11.00	5.91	9.73	10.42	10.58	10.53	10.47
Crude	12.69	12.31	12.77	13.39	14.05	14.78	12.69	11.99	12.15	12.09	12.38	12.33
Gas	6.09	6.94	9.16	10.69	12.56	14.63	6.09	6.14	6.87	7.53	8.27	9.22
Hydro	4.19	4.54	4.58	4.61	4.65	4.68	4.19	4.54	4.89	5.24	5.43	5.47
Others	1.99	1.87	1.70	1.49	1.18	0.85	1.99	1.99	2.64	3.25	3.27	3.28
Total	30.87	34.22	37.76	40.59	43.19	45.94	30.87	34.39	36.97	38.69	39.88	40.77

Table 1: Primary Energy Consumption in Austria, 1984 to 2004, in GWyr/yr. (left: constant energy prices, right: growing energy prices)

The similar development of coal utilization can be explained by the constraint imposed on emissions, that are not allowed to grow above the 1984 level. The high share of hydro power in electricity production limits the potential for emission reduction in this sector. As it is very costly to reduce emissions at the level of the end user, gas is used to substitute for high sulphur fuels in the end use heat markets. By the end of the planning period the high energy price case shows a reduction of final energy demand by some 11%, when compared to the constant energy price scenario. Expensive imports are substituted by growing supplies from hydro power and 'other fuels', i.e. industrial and municipal waste and wood.

A comparison of the results from the economy model (see Table 2) show a slight increase of real GDP in the high energy price scenario. However, a much larger share of produced goods has to be exported and tourism to be increased in order to keep the trade balance at the required level in both cases.

The two main constraints imposed on the economy model, the trade balance and full employment, result in a reduction of wages by some 13% relative to the constant price scenario. At this point a decision maker could test his hypotheses about how to solve the problem of reduced growth of private income while still avoiding unemployment.

	1984	1988	1992	1996	2000	2004	1984	1988	1992	1996	2000	2004
Private	413.5	467.5	507.4	544.7	589.1	642.1	413.5	462.0	492.6	517.0	547.4	578.3
Tourism	61.6	66.0	74.5	83.3	94.3	106.2	61.6	66.9	76.9	87.8	101.3	117.4
Public	172.4	183.6	201.1	216.5	224.4	236.3	172.4	183.6	200.8	216.0	223.6	235.3
Investm	180.6	194.7	206.5	218.1	238.1	254.6	180.9	195.4	208.1	220.9	242.8	262.1
Exports	277.8	313.7	357.5	406.1	469.0	540.3	278.2	318.0	369.6	429.7	507.4	602.7
Imports	-287.3	-338.3	-391.8	-450.7	-524.6	-609.9	-287.7	-338.0	-391.2	-450.4	-526.3	-615.7
Total	818.60	887.20	955.20	1018.00	1090.30	1169.60	818.90	887.90	956.80	1021.00	1096.20	1180.10

Table 2: Gross Domestic Product by Use in Austria, 1984 to 2004, in 10^9 AS. (left: constant energy prices, right: growing energy prices)

7. Final Remarks

The model presented in this article is a first attempt towards an integrated tool for decision support in the field of energy-economy analysis. It can be used for assessing various policies regarding the development of the Austrian energy supply system. Currently the sensitivity of the model is analyzed. Different assumptions about the development of the factors influencing energy demand and supply patterns as well as ones relating to other policy issues are tested for their implications.

Important questions to be tackled include the question of the construction of new hydro power plants. Currently this is a main issue in the Austrian energy debate. The model can also help in the analysis of another, even more urgent problem. This is the question about the preferable path towards a reduction of the emissions stemming from energy utilization.

References

[1] A.S. Manne. A Model for Energy-Economy Interactions. Paper presented at the Operations Research Society of America/The Institute of Management Sciences Meeting, San Francisco, California, May 9-11,1977.

[2] R. Coldoni, B. Fritsch, (Eds.). Capital Requirements of Alternative Energy Strategies, A Techno-Economic Assessment, Center for Economic Research, Swiss Federal Institute of Technology, Zürich, September 1980.

[3] M. Grauer, S. Messner, M, Strubegger. An Integrated Programming Package for Multi-Criteria Decision Analysis. In Plural Rationality and Interactive Decision Processes. Proceedings, Sopron, Hungary,1984. Springer-Verlag. Berlin, 1985.

[4] M. Grauer, A. Lewandowski, A. Wierzbicki. DIDASS--Theory, Implementation and Experiences. In M. Grauer and A. Wierzbicki (Eds). Interactive Decision Analysis. Springer-Verlag, Berlin, 1984.

[5] L. Johansen. Production Functions. North Holland. Amsterdam, 1972.

Natural Gas Trade in Europe and Interactive Decision Analysis

Sabine Messner

International Institute for Applied Systems Analysis
Laxenburg, Austria

Abstract

The IIASA International Gas Study focuses on the development of the European energy system with special emphasis on the evolution of the natural gas market. The competitiveness of natural gas in the various energy end-use markets, resource availability, long term gas supply options and implications on international trade are investigated. Studying natural gas trade requires the analysis of complex decision problems with several interest groups involved.

In the next chapters an outline of the International Gas Study performed at IIASA is given, followed by a description of the overall methodological approach used to model the development of the energy system. In particular the setup of the Gas Trade Model (GATE) and some findings are presented. This Gas Trade Model is based on the decision support system DIDASS [1], which is applied to simulate the decisions of the partners involved in international gas trade in Europe. It turned out to be very instrumental in the investigation of different strategies regarding the future development of gas trade.

Introduction

Over the last decades the volume of international gas trade in Europe has been growing significantly. Table 1 shows that, while gas trade was negligible up to the late 1960ies, it reached a volume of nearly 150 billion m^3* by 1980. In that year the non-CMEA countries--hereafter called Western Europe--imported roughly 50% of the natural gas they consumed. Already in the beginning of the 60ies natural gas had started to increase its market share in primary energy consumption, stimulated by the development of the large and inexpensive Groningen field in the Netherlands.

* 1 billion m^3 = 1 bcm = 10^9 m^3 = 37.26 10^{15}J

This growing importance of natural gas on the European energy markets and in international trade lead to the initialization of the IIASA International Gas Study. It is set up to investigate the potential future development of gas consumption and international gas trade and resource potentials.

Table 1: Development of the European natural gas market, 1950-1980 (in billion m^3).

Year	Exports				Imports	
	The Netherlands	Norway	Soviet Union	Other Exporters	non CMEA Europe	CMEA
1950	-	-	.1	-	-	.1
1960	-	-	.2	.2	-	.4
1970	11.3	-	3.3	5.3	13.8	6.1
1980	50.8	24.6	56.5	13.3	111.6	33.9

(Source: [2])

Scope of the Study

Previous analyses performed at IIASA [3] indicate that it takes some 50 years until a new energy form can increase its market share from 5 to 50%. As one of the goals of this study is to investigate the possible penetration of natural gas into the energy system a time frame of 45 years (1985 to 2030) was adopted. The potential consumers of natural gas include all European Countries. They were aggregated into five regions (Central Europe, East Europe, North Europe, Southeast Europe and Southwest Europe) in accordance with political system, degree of economic development and geographic aspects. The potential sources of natural gas, i.e. areas exporting to the consuming regions, are The Netherlands, Norway, the Soviet Union, North Africa and Other Exporters, mainly the Middle East.

The focal issues investigated in the analysis of natural gas potentials in the European market are

- resource potentials, costs and constraints of extraction,

- demand development, options of energy utilization and conversion, burner tip competitiveness of natural gas and rates of market penetration; and
- trade options and patterns, related infrastructure requirements and economic, policy and geopolitical considerations.

The data needed for this analysis were compiled and reviewed together with experts from various gas/energy businesses and energy research institutions from East and West. In many cases these basic research areas were amended by additional investigations. So the work on energy conversion and utilization also includes the environmental component. SO_2 and NO_x emissions, desulfurization and denoxing units, and changes in environmental standards were included in the analyses.

The data related to domestic resources and energy conversion and utilization were synthesized using a dynamic optimization model, MESSAGE II [4]. It depicts the whole energy conversion chain from resource extraction via conversion, transport and distribution to the utilization of final energy in households, the industrial sectors and for transportation. In order to get a comprehensive picture about the possible development of the regional energy systems such a model was built for each of the five consumer regions. Scenario variables like the international oil price, the development of environmental standards or the development of energy demand in the future are analyzed for their impact on the modeled energy systems. A more comprehensive description of the setup of the study can be found in [5].

Figure 1 shows the—historical and projected—primary energy mix for the period 1950 to 2030 for Western Europe. The scenario shown is based on the assumption that economic growth will accelerate again after 1990 and that conservation measures have gained so much impetus that they inhibit a growth in primary energy consumption up to the turn of the century. Environmental standards are assumed to remain unchanged for the next decades. The relations between the world market prices of oil, coal and gas are changed only marginally, resulting in a growing market share of coal and gas. Oil is loosing its leading position because it can no longer be burnt economically in power generation units and large industrial applications. These markets are conquered by coal. Natural gas gains mainly in the residential and commercial markets. The growth of gas end use is counteracted by increases in efficiency, conservation measures and the implementation of new technologies like heat pumps, resulting in a moderate growth of primary gas consumption.

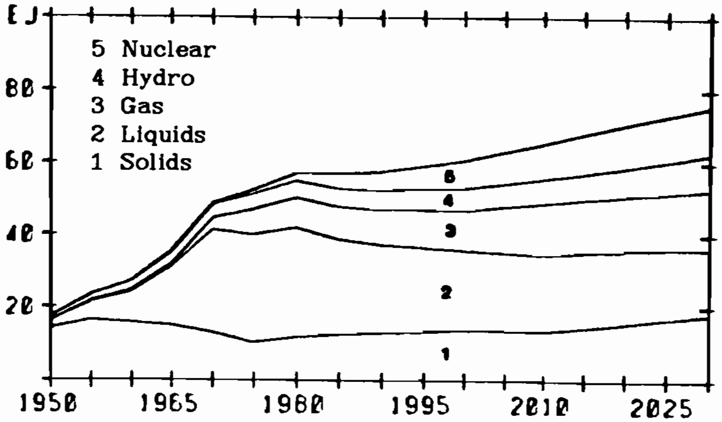


Figure 1: Primary Energy Consumption in Western Europe (1950 to 2030 in EJ*)

The Trade Model

Since Europe is a relatively resource poor area it is highly dependent on energy imports. The availability of natural gas for the consumers depends to a high degree on imports, turning the analysis of gas trade flows into a central issue of this gas study. The typical features required in a gas trade model that links the technology oriented cost minimization models of the European consumer regions are

- The capability to capture *technical and economic realities*. Resource potentials and extraction technologies in the exporting, but also in the importing regions fall into this category, the capital intensive transport of natural gas via pipeline or LNG tankers and regasification units have to be considered.
- The ability to represent the *different objectives* of the parties involved in international gas trade. They include cost minimization for consuming regions and profit maximization for producers.

Conventional modeling approaches like simulation or optimization are not adequate to this problem. They neglect the different objectives and realities as they are seen by the parties involved.

* $1 EJ = 10^{18} J$

The IIASA International Gas Study team adopted a new approach. A multiobjective optimization method, that was originally designed to assist decision makers interactively by investigating the problem in question, was used to simulate the trade options in Europe. This method, the Reference Point Optimization method developed at IIASA [1], supports the use of time-dynamic linear programming models. It requires the specification of reference trajectories for all objectives opposed to expressing preferences in terms of relative weights. This feature facilitates the combination of objectives expressed in units which are not directly comparable, like East or West European currencies. Thus, cost minimization can be applied independently for each demand node without having to specify exchange rates. Different types of objectives like energy quantities and energy costs can also be included simultaneously.

Although this trade model cannot predict the future, it can help to investigate possible optimal consensus strategies under various assumptions regarding the aspirations of the trading partners, resource extraction limitations, demand and price developments, etc. The software package used, IMM, and some potential applications are described in [6].

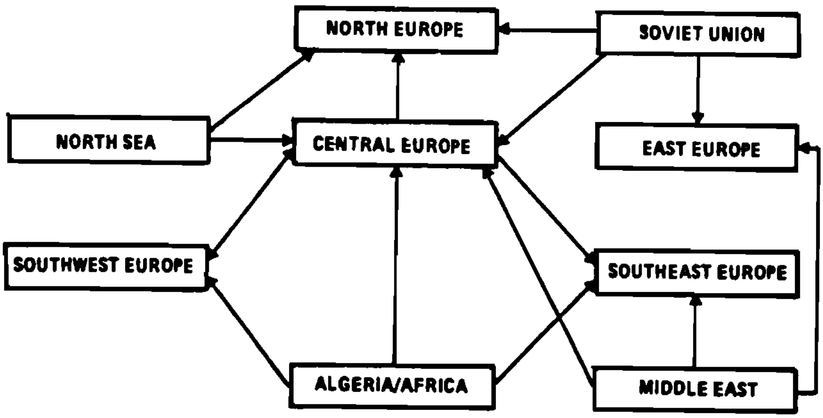


Figure 2: Gas Trade Links

Model Description

The potential natural gas transportation links depicted in figure 2 serve as a basis for investigating the effect of different trade strategies on the gas flows between exporters and importers. Each of the consuming regions has the options to extract gas, import it, or conserve energy. The reference level of natural gas consumption is derived from the disaggregated regional models that were outlined above.

The objectives considered are cost minimization for the consumers, maximization of profits for the European exporters, maximization of income for the Soviet Union and the Other Exporters and minimization of gas imports of Europe from outside (see figure 3). The gas trade model includes 37 extraction technologies, 43 transport and link technologies, 16 contracts and 15 variables to represent gas consumption and conservation. It covers 6 periods, resulting in a size of 2170 rows and 1410 columns with a density of .332%.

Objective	Region
Minimize Costs of Natural Gas Supply	North Europe Central Europe Southwest Europe Southeast Europe East Europe
Maximize Income minus Gas Extraction Costs	United Kingdom The Netherlands Norway
Maximize Income	Soviet Union Other Gas Exporters
Minimize Imports of Europe from	Soviet Union Other Gas Exporters

Figure 3: Objectives considered in the gas trade model

Based on *utopia trajectories*, that are compiled from the optimal solutions for each single objective, and the *reference trajectories* set for these objectives, IMM determines a pareto-optimal solution. The procedure of setting reference trajectories and solving the problem can be iterated interactively until a satisfactory result is generated.

The utopia trajectories of the gas trade model (the combination of optimal solutions from the single objective optimization of each all trajectory, see table 2) show some of the characteristics of the modeled regions. East Europe will have to import increasing amounts of natural gas, partly due to falling oil supplies from the Soviet Union, partly due to the domestic resource situation, and will, even in the best case, have to pay rapidly increasing gas bills. The Netherlands are confronted with a lifetime of their gas export potential of roughly 30 years, thereafter the income will be essentially zero. The Soviet Union and the Other Exporters could supply growing amounts of gas, especially when presently binding contracts with other exporters expire. On the other hand the importers could, by enforcing their preference for low imports of gas from outside Europe, limit these imports to virtually zero after 2010.

Table 2: Utopia trajectories for the objectives (1985 = 1.)

Objective	1985	1990	2000	2010	2020	2030
Cost North	1.00	1.38	1.78	3.44	3.88	3.69
Cost Central	1.00	1.13	1.24	1.77	2.03	2.21
Cost Southeast	1.00	1.23	1.27	1.60	2.58	3.44
Cost Southwest	1.00	0.66	0.94	2.42	2.75	3.12
Cost East	1.00	1.04	1.10	1.66	2.34	3.21
Profits Norway	1.00	0.96	0.86	0.79	0.76	0.31
Profits Netherlands	1.00	0.69	0.57	0.25	0.24	0.16
Profits United Kingdom	1.00	0.68	0.40	0.35	0.15	0.10
Income Other Exporters	1.00	3.62	4.96	7.99	11.28	14.62
Amount Other Exporters	1.00	0.83	0.72	0.27	0.00	0.00
Income Soviet Union	1.00	1.35	1.57	1.95	2.35	2.56
Amount Soviet Union	1.00	0.88	0.83	0.25	0.01	0.01

Findings

In the following a sequence of runs will be described that were made to check the consistency of the approach as well as its sensitivity.

In a reference run the trade model was formulated as single objective problem. The overall cost of supplying natural gas to Europe was minimized. The solution shows the cost optimal structure of the supply picture if profits of the European exporters are neglected and the Soviet Union and Other Exporters sell at given prices. Figure 4 shows this cost optimal supply pattern: European production can be increased, specially by using the large and expensive Norwegian Troll field. After 2010, when consumption of gas grows at a rate of 2%/yr, domestic production starts to decline due to diminishing natural gas resources in the consuming regions and The Netherlands. By 2030, 50% of all gas is supplied by the Soviet Union. This is due to the relatively lower price charged by the Soviet Union compared to the Other Exporters. Having in mind the extraction potentials of the Soviet Union, that are not included in this model, these exports are certainly too high. In order to avoid such problems, later versions of the model will also contain gas extraction and consumption technologies for the Soviet Union.

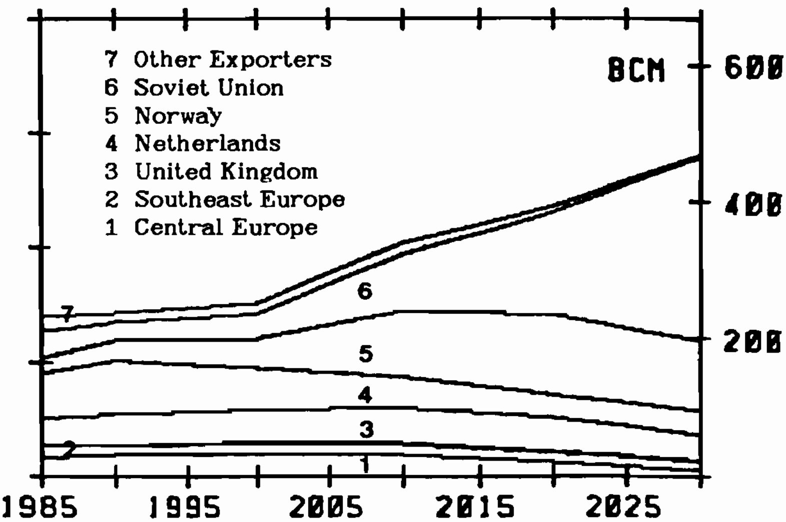


Figure 4: Natural Gas Supply in Western Europe, Cost Minimization Run (1985 to 2030 in billion m^3).

For a first test case the reference levels for all cost objectives were set to the same level of deviation from the utopia trajectories (the deviation was set to

5%), while the import constraining objectives for the Soviet Union and Other Exporters are set further away form the utopia trajectories (the deviation was set to 10%*). The result depicted in figure 5 shows that extraction inside Europe is kept at a low level. Specially Norway and the United Kingdom do not exploit their expensive resources in the northern North Sea, because they cannot be produced economically if profits are accounted for. The Soviet Union and the Other Exporters share the export market equally.

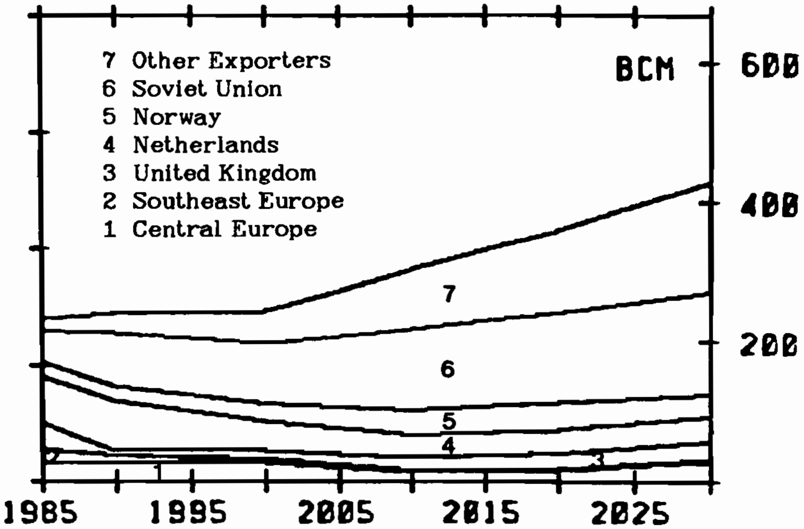


Figure 5: Natural Gas Supply in Western Europe, High Dependence Run (1985 to 2030 in billion m^3).

However, if the European countries would decide to reduce the dependence from outside (see figure 6), this could be accomplished quite fast by reducing LNG and piped gas imports from North Africa. After 2000 the consuming regions extract expensive gas that was even left untouched in the cost minimization run. Also, Norwegian and United Kingdom resources are stretched to the limit. The imports from outside Europe are constant up to 2010, thereafter they increase very moderately.

These sensitivity runs show that the solutions can be reasonably explained by the assumptions made regarding the reference levels for the various objectives and the model structure as a whole. By expressing the willingness to pay for the

* This somewhat arbitrary approach was taken in order to test the sensitivity of the model.

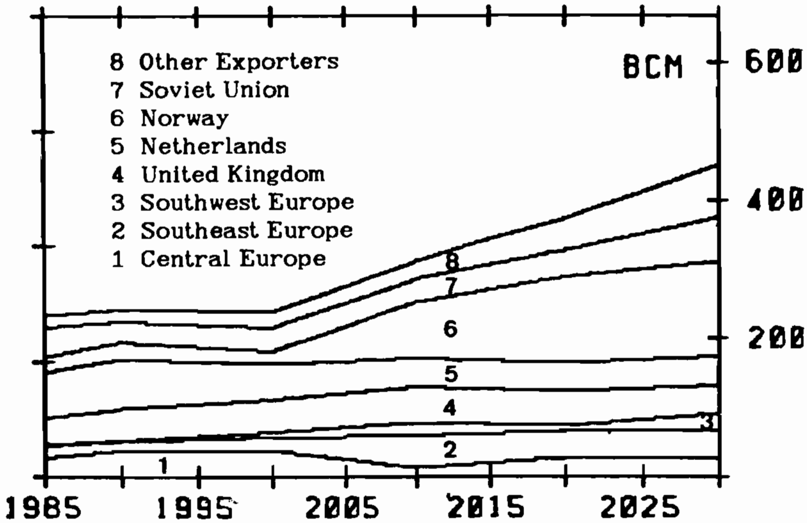


Figure 6: Natural Gas Supply in Western Europe, Low Dependence Run (1985 to 2030 in billion m^3).

supply of natural gas for the importers, the desired profits/income for the exporters and the political consideration of the amount of gas allowed to come from outside Europe, very descriptive and logical scenarios of gas extraction and trade can be generated.

Whereas it is difficult to get a good picture of the shape and size of the feasible region in conventional optimization methods, the Reference Point Optimization Method opens a way to achieve that information. Thus it is possible to give ranges which allow a more intricate interpretation of the results than simple point solutions.

The synthesis of the described runs can be used to illustrate this. Table 3 shows gas supply by major producer for all three scenarios in the year 2030. For the Netherlands the solutions differ only marginally: by 2030 their extraction just matches domestic demand. For the other three large suppliers the range covered by the scenarios is very wide and shows that the feasible region, i.e. the range of achievable solutions, is quite large. For 2000 the according table would show much less variations. The flows are still restricted by contracts and the low growth in the demanded volume does not allow for much flexibility.

Table 3: Gas supply by major producer (2030 in billion m^3)

2030	Netherlands	Norway	Soviet Union	Other Exporters
Cost Minimization	41.	123.	320.	5.
High Dependence	41.	39.	175.	188.
Low Dependence	41.	170.	86.	89.

Final Remarks

This analysis of the potential development of international gas trade in Europe proved that interactive decision support systems like DIDASS can, besides their original field of applications, also be used to simulate decision processes in the development of scenarios. It can show the impact of different approaches to a problem and thus help to investigate a range of possible outcomes in complex multi-actor problems.

In the case of the IIASA Gas Study the application of a decision support system opened the opportunity to investigate the structure of optimal consensus strategies under different assumptions concerning the objectives of the importers and exporters of natural gas. For future work the most intriguing problem would be to investigate the results in situations where no basic consensus can be achieved between the actors. However, presently no operational formal approach to deal with such problems is known. Thus, parallel to the improvement and dissemination of methods that deal with the problem of consensus finding, methods have to be developed that tackle the problem of searching for solutions in cases when no consensus can be achieved between negotiating partners.

References

[1] M. Grauer, A. Lewandowski and A.P. Wierzbicki. DIDASS—theory, implementation and experiences. In M. Grauer and A.P. Wierzbicki (Eds.). Interactive Decision Analysis. Springer-Verlag. Berlin, 1984.

[2] M. Valais, P. Boisserpe and J.L. Gadon. The World Gas Industry. Editions Technip. Paris, 1982.

[3] C. Marchetti and N. Nakicenovic. The Dynamics of Energy Systems and the Logistic Substitution Model. Research Report RR-79-13, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1979.

[4] S. Messner. User's Guide for the Matrix Generator of MESSAGE II. Working Paper WP-84-71. International Institute for Applied Systems Analysis, Laxenburg, Austria, 1984.

[5] H-H. Rogner, S. Messner and M. Strubegger. European Gas Trade: A Quantitative Approach. Working Paper WP-84-44. International Institute for Applied Systems Analysis, Laxenburg, Austria, 1984.

[6] M. Grauer, S. Messner and M. Strubegger. An Integrated Programming Package for Multiple-Criteria Decision Analysis. In: Plural Rationality and Interactive Decision Processes. Springer-Verlag. Berlin, 1985.

AN APPLICATION OF SATISFICING TRADE-OFF METHOD TO A BLENDING PROBLEM OF INDUSTRIAL MATERIALS

H. Nakayama*, J. Nomura**, K. Sawada** and R. Nakajima**

* Department of Applied Mathematics, Konan University
8-9-1 Okamoto, Higasinada, Kobe 658, JAPAN

** Research and Development Laboratory, Matsushita Electric Works, Ltd.
1048 Kadoma, Kadoma-city, Osaka 571, JAPAN

1. INTRODUCTION

In recent years, there have been remarkable developments in interactive multiobjective programming methods. It is very important to feedback from applications in order to sophisticate methodology. Although some trials of application have reported, it seems to us that it is just now the time to apply our methods to the real world. From this viewpoint, the authors have been trying to apply several methods to real problems. In this paper, one of the results in a blending problem of some industrial material will be reported.

In the recent industrial world, many kinds of industrial materials are used and they are mostly made from some kinds of raw materials. By ratios of the adopted raw materials, production cost and various characteristics (mechanical characteristic, chemical characteristic, electrical characteristic, etc.) of a resulting material are changed. Hence, the production cost and these various characteristics should be considered as the criteria for blending problems of industrial materials, and are generally conflicting with each other. This paper treats a blending problem of an industrial material as an optimization problem with six objectives. The satisficing trade-off method is proved to be effectively used for this problem.

2. FORMULATION OF A BLENDING PROBLEM

The industrial material considered here is a plastic material made from six raw materials, and it is often used in many kinds of household appliances, for example, bathtubs, surface-materials of wall and etc. These six raw materials are oxides (alminium oxide, silica oxide, calcium oxide, etc.). The characteristic values of the material are represented by nonlinear functions of the raw material blending-ratios x_i $(i=1,2,\ldots,6)$ as follows:

(a) Cost (yen)

$$f_1 = (0.231\times10^1x_1+0.583x_2+0.813\times10^1x_3+0.671x_4+0.240x_5 +0.10x_6)\times10^3 \;/\; (0.696x_1+0.620x_2+0.299x_3\times10^1x_4 +0.60x_5+0.561x_6)$$

(b) Melting point (°C)

$$f_2 = 0.140x_2^2+0.203x_3^2+0.114x_1x_3-0.105x_1x_4+0.768\times10^{-1}x_1x_6 -0.148\times10^1x_1-0.112\times10^2x_2-0.167 10^2x_3+0.307\times10^1x_4 -0.444\times10^1x_6+0.540\times10^3$$

(c) Coefficient of thermal expansion (10^{-5}/°C)

$$f_3 = 0.153\times10^1x_2+0.106\times10^1x_3-0.227x_4+0.351x_6+0.206\times10^2$$

(d) Acid-resisting characteristic (%)

$$f_4 = 0.117\times10^1x_6-0.16x_1x_4-0.881x_2x_3-0.830\times10^{-1}x_2x_4-0.990x_2x_4 -0.218x_4x_5+0.336x_1+0.117\times10^2x_2+0.115\times10^2x_3+0.986\times10^1x_4 +0.505\times10^2x_5-0.174\times10^2x_6-0.189\times10^3$$

(e) Alkaline-resisting characteristic (%)

$$f_5 = 0.230\times10^{-1}x_1^2-0.735\times10^{-1}x_4^2+0.445\times10^{-1}x_1x_3-0.915\times10^{-1}x_2x_5 -0.124x_2x_4+0.255\times10^{-1}x_3x_4-0.10x_4x_5-0.164\times10^{-1}x_1 +0.366\times10^1x_2-0.146\times10^1x_3+0.543\times10^1x_4+0.259\times10^4x_5 +0.114\times10^4x_6-0.728\times10^2$$

(f) Aqua-resisting characeristic (%)

$$f_6= 0.148x_6^2+0.258\times10^{-1}x_1x_2-0.765\times10^{-1}x_5x_6+0.189\times10^1x_1 -0.70\times10^{-1}x_1x_5-0.830x_2-0.75x_3+0.714x_5+0.314\times10^1x_6 +0.839\times10^2$$

The functions f_1 through f_6 are obtained by using the multiple regression analysis. These are considered to be the objectve functions of the problem.

The constraints on x_i are given by

(g) Blending constraint

$$\sum_{i=1}^{6} x_i = 100$$

$$0 \leq x_i \leq 100 \qquad i=1,2,\ldots,6$$

Hence, our blending problem is formulated into the following optimization problem with six objectives:

$$\text{(P)} \qquad \text{Min } \{f_1(x), f_2(x), f_3(x), f_4(x), f_5(x), f_6(x)\}$$

subject to

$$\sum_{i=1}^{6} x_i = 100$$

$$0 \leq x_i \leq 100 \qquad i=1,2,\ldots,6$$

Here, it should be noted that the desirable levels of the six criteria vary depending on the final products. Therefore, we need methods which can be used in a flexible manner according to the final products.

3. APPLICATION OF SATISFICING TRADE-OFF METHOD

Some of authors have already tried to apply the Surrogate Worth Trade-off Method (Haimes-Hall-Freedmann [2]) and Multi-attribute Utility (Value) Analysis to the problem described in the previous section (Nishikawa-Nomura [7], Nishikawa-Nomura-Sawada [8]). They have observed that the Surrogate Worth Trade-off Method is very difficult to use as it is, because it requires the decision maker (designer) to answer his surrogate worth many times. If all surrogate worths are required to be consistent, the method would not be realistic. Therefore, some modifications were tried. A major feature of them is to show the trade-off relation among criteria using computer graphics. This made the decision maker to judge his value trade-off very easily. Even so, a lot of calculation was needed for auxiliary optimization.

On the other hand, in the multi-attribute utility analysis, we could learn much about the preference of the decision maker. However, it took terribly much time to get the decision maker accustomed with the method. Moreover, since it usually takes much time to obtain an overall utility function, it is lacking in flexibility for variable desirability depending on the final products.

For such a circumstance, one of the authors have developed an effective interactive programming method called the satisficing trade-off method (Nakayama [3], [4]), which is similar to DIDASS (Grauer-Lewandowski-Wierzbicki [1]), and applied to some experimental problems (Nakayama-Sawaragi [5], Nakayama-Furukawa [6]). In this paper, we shall report a result using the satisficing trade-off method in our blending problem, in which the method will be verified to be very effective for obtaining the solution in a simple, easy and fast way.

3.1 SATISFICING TRADE-OFF METHOD

We shall review the satisficing trade-off method briefly. As long as we take an approach of optimization, we can not help requiring decision makers such difficult judgment as margianl rate of substitution or ordering vectors. In order to decrease the burden of decision makers, interactive programming methods taking the approach of satisficing as a decision rule have been developed. The goal programming may be formulated as one of them, in which the solution is just a satisfactory one if the aspiration level is feasible, or it is as close as possible to the aspiration levels if it is infeasible.

Note here that the satisficing was originally suggested due to the observation of the limit of human cognitive ability and available information through the decision making. In our problem formulation, however, we are supposed to be able to evaluate the values of criteria functions by some means, although we also observe the limit of human cognitive ability. For these problems, the mere satisficing is not necesarily satisfactory to the decision maker as a decision rule. Therefore, we try to show DM a solution $\tilde{x}$ which is satisfactory and in addition guarantees that there is no other feasible solution which is superior to $\tilde{x}$ in terms of all criteria. We call a solution which is satisfactory and (weakly) Pareto-efficient a satisfactory (resp. weak) Pareto solution. In other words, when $\bar{f}$ is an aspiration level of the decision maker, an alternative $\tilde{x}$ is said to be a satisfactory Pareto solution if

$$f(x) \not\leq f(\tilde{x}) \qquad \text{for all } x \in X$$

and

$$f(\tilde{x}) \leqq \bar{f}.$$

A satisfactory weak Pareto solution can be defined analogously by replacing $\not\leq$ with $\not<$ in the first inequality above.

In general, satisfactory (weak) Pareto solutions constitute a subset of X. However, we can narrow down the set of satisfactory Pareto solution set and obtain a solution close to the preferentially optimal one by tightening the aspiration level, if necessary. On the other hand, even if there does not exist any satisfactory Pareto solution for the initial aspiration level, we can attain one by relaxing the aspiration level. Originally, the aspiration level is fuzzy and flexible. It seems that decision makers change their aspiration levels adaptively according to situations in practice.

From this point of view, the interactive satisficing methods provide a tool supporting decision makers to find an appropriate satisfactory (weak) Pareto solution by some adjustment of the aspiration level: usually the aspiration level is changed in such a way that

$$\bar{f}^{k+1} = T \circ P\,(\bar{f}^{k})$$

where $\bar{f}^k$ represents the aspiration level at the k-th iteration. The operator P selects the Pareto solution nearest in some sense to the given aspiration level $\bar{f}^k$. The operator T is the trade-off operator which changes the k-th aspiration level $\bar{f}^k$ if the decision maker does not compromise with the shown solution $P(\bar{f}^k)$. Of course, since $P(\bar{f}^k)$ is a Pareto solution, there exists no feasible solution which makes all criteria better than $P(\bar{f}^k)$, and the decision maker has to trade-off among criteria if he wants to improve some of criteria. Based on this trade-off, a new aspiration level is decided as $ToP(\bar{f}^k)$. Similar process is continued until the decision maker obtain an agreeable solution.

The algorithm of the satisficing trade-off method is summarized as follows:

Step 1 (setting the ideal point): The ideal point $f^*=(f_1^*,\dots,f_r^*)$ is set, where f_i^* is small enough, for example, $f_i^*=\mathrm{Min}\{f_i(x) \mid x\varepsilon X\}-\varepsilon$ ($\varepsilon>0$). This value is fixed throughout the following process.

Step 2 (setting the aspiration level): The level $\bar{f}_i^k$ of each objective function f_i at the k-th iteration is asked to the decision maker. Here $\bar{f}_i^k$ should be set in such a way that $\bar{f}_i^k > f_i^*$ $(i=1,\dots,p)$. Set k=1.

Step 3 (weighting and finding a Pareto solution by the Min-Max method): Set

$$w_i^k = \frac{1}{\bar{f}_i^k - f_i^*}, \tag{3.1}$$

and solve the Min-Max problem

$$(P_\infty^k) \qquad \underset{x\varepsilon X}{\mathrm{Min}}\ \underset{1\le i\le r}{\mathrm{Max}}\ w_i^k |f_i(x) - f_i^*|$$

or equivalently

$$(P_\infty'^k) \qquad \underset{x,\xi}{\mathrm{Min}}\ \xi$$

$$\text{subject to} \quad w_i^k(f_i(x) - f_i^*) \leq \xi, \quad i=1,\dots,r$$

$$x\varepsilon X$$

Let x^k be a solution to these problems.

Step 4 (trade-off): Based on the value of $f(x^k)$, the decision maker classifies the criteria into three groups, namely,

(i) the class of criteria which he wants to improve more,

(ii) the class of criteria which he may agree to relax,

(iii) the class of criteria which he accept as they are.

The index set of each class is represented by I_I^k, I_R^k, I_A^k, respectively. If

$I_I^k = \emptyset$, then stop the procedure. Otherwise, the decision maker is asked his new acceptable level of criteria, $\tilde{f}_i^k$, for the class of I_I^k and I_R^k. For $i \varepsilon I_A^k$, set $\tilde{f}_i^k = f_i(x^k)$.

Step 5 (feasibility check): Let λ_i $(i=1,\dots,r)$ be the optimal Lagrange multipliers to the problem $(P_\infty'^k)$. If for a small nonnegative ε

$$\sum_{i=1}^{r} \lambda_i w_i^k (\tilde{f}_i^k - f_i(x^k)) \geq -\varepsilon, \tag{3.2}$$

then set the new aspiration level $\bar{f}_i^{k+1}$ as $\tilde{f}_i^k$ and return to the step 3. Otherwise, $\tilde{f}_i^k$ might be infeasible in the sense of linear approximation (Nakayama [3]). Then, by taking the degree of difficulty for solving the Min-Max problem into account, we choose either to trade-off again or to return to the step 3 by setting $\bar{f}_i^{k+1} = \tilde{f}_i^k$. In case of trading off again, the acceptable level of criteria for I_I^k and/or I_R^k should be reset lower than before, and go back to the beginning of the step 5.

Now, we shall list outstanding features of the satisficing trade-off method in the following:

1) We do not need to pay much attention to setting the ideal point f^*. It suffices to set f^* sufficiently small enough to cover all or almost of all Pareto solutions as candidates for a decision solution in the following process. In case of Min $\{f_i(x) \mid x \varepsilon X\}$ being finite, for example, set $f_i^* = \mathrm{Min}\{f_i(x) \mid x \varepsilon X\} - \varepsilon$ $(\varepsilon > 0)$. Otherwise, set f_i^* to be sufficiently small.

2) The weights w_i $(i=1,\dots,r)$ are automatically set by the ideal point f^* and the aspiration level $\bar{f}$. For the weight with (3.1), the value of $w_i(f_i(x) - f_i^*)$ can be considered to represent the normalized degree of non-attainability of $f_i(x)$ to the ideal point f_i^*. This enables us to need not to pay an extra attention to the difference among the dimension and the numerical order of criteria.

3) By solving the Min-Max problem with the above we can get a satisfactory weak Pareto solution in case of the aspiration level $\bar{f}$ being feasible, and just a weak Pareto solution even in case of $\bar{f}$ being infeasible. Interpreting this intuitively, in case of $\bar{f}$ being feasible the obtained satisfactory weak Pareto solution is the one which improves equally in some sense each criterion as much as possible, and in case of $\bar{f}$ being infeasible we get a weak Pareto solution nearest to the ideal point which share an equal amount of normalized sacrifice for each criterion. This practical meaning encourages the decision maker to accept easily the obtained solution.

4) At the stage of trade-off, Tthe new aspiration level is set in such a way that

$$\bar{f}_i^{k+1} < f_i(x^k) \qquad \text{for any} \quad i \varepsilon I_I^k,$$

$$\bar{f}_i^{k+1} > f_i(x^k) \qquad \text{for any} \quad 1 \varepsilon I_R^k,$$

In this event, if $\bar{f}^{k+1}$ is feasible, then we have a satisfactory weak Pareto solution by solving the Min-Max problem (Nakayama [3]). In setting the new aspiration level, therefore, it is desirable to pay attention in such a way that it becomes feasible.

5) In order that the new aspiration level $\bar{f}^{k+1}$ may be feasible, the criteria f_i $(i \varepsilon I_R^k)$ should be relaxed much enough to compensate for the improvement of $f_i(x^k)$ $(i \varepsilon I_I^k)$. To make this trade-off successful without solving a new Min-Max problem, we had better make use of sensitivity analysis on the basis of informatin of Lagrange multipliers (Nakayama [3]). In the recent version of the staisficing trade-off method, we revised to assign the sacrifice for f_j $(j \varepsilon I_R)$ automatically set in the equal proportion to $\tilde{\lambda}_i w_i$, namely, by

$$\Delta f_j = \frac{-1}{N\tilde{\lambda}_j w_j} \sum_{i \in I_I} \tilde{\lambda}_i w_i \Delta f_i$$

where N is the number of elements of the set I_R. By doing this, in cases where there are a large number of criteria, the burden of the decision maker can be decreased so much. Of course, if the deicion maker does not agree with this quota Δf_j laid down automatically, he can modify them in a manual way.

Remark 3.1

In some cases, decision makers want to know the global feature of the feasible set in trading-off. To this end, it is convenient to show the so-called pay-off matrix by minimizing each objective function independently, which was introduced in STEM. Of course, since the pay-off matrix does not change during the interactive process, system analysists had better prepare it in advance to the decision process.

3.2 Experimental Results

The decision makers are two chemical engineers who are engaged in this problem for a long period and have an experience using SWT method and multiattribute utility analysis before. In the following, we shall report one of the results in our experiment: The ideal point was set as

$$f_1^*=0.0,\ f_2^*=0.0,\ f_3^*=10.5,\ f_4^*=0.0,\ f_5^*=0.0,\ f_6^*=0.0$$

For the first aspiration level $\bar{f}$ = (300, 520, 10.9, 2.9, 2.9, 0.9), we obtained the follwong (weak) Pareto solution:

	Pareto solution	aspiration level	sensitivity
f_1	243.9	300.0	0.167
f_2	422.8	520.0	0.815
f_3	10.8	10.9	0.014
f_4	1.47	2.9	0.0
f_5	0.98	2.9	0.0
f_6	0.73	0.9	0.0034

Although every criteria is satisfactory at the obtained Pareto solution, the first criterion (cost) was tried to improve a little. The next aspiration level was set $\bar{f}_1$=200.0 with the other criteria fixed at the level of the obtained Pareto solution.

	Pareto solution	aspiration level	sensitivity
f_1	207.4	200.0	0.192
f_2	438.5	422.8	0.788
f_3	10.8	10.8	0.012
f_4	0.29	1.47	0.0
f_5	1.02	0.98	0.0
f_6	0.76	0.73	0.0065

At this stage, the decision maker wanted to improve f_1 to 180.0, while relaxing f_2 to 490.0 and f_4 to 0.9.

	Pareto solution	aspiration level	sensitivity
f_1	169.5	180.0	0.193
f_2	462.1	490.0	0.792
f_3	10.82	10.8	0.010
f_4	0.68	0.90	0.0
f_5	0.88	1.02	0.0
f_6	0.71	0.76	0.0051

Since the obtained solution is sufficiently satisfactory, the iteration was stopped here.

Remark 3.2

It is seen that there are many sensitivities with zero value at each iteration. Therefore, since the obtained Pareto solutions might be just weak Pareto solutions, we adopted the following auxiliary problem in order to obtain strong Pareto solutions.

(AP) Maximize $\sum_{i=1}^{r} \varepsilon_i^2$

subject to

$$f_i(x) + \varepsilon_i^2 = f_i(\hat{x}) \qquad (i=1,\ldots,r)$$

$$x \in X$$

where $\hat{x}$ is a (weak) Pareto solution. If all ε_i for the solution to (AP) are zero, then $\hat{x}$ itself is a strong Pareto solution. If there are some $\varepsilon_i \neq 0$, then the solution $\tilde{x}$ to the problem (AP) is a strong Pareto solution.

By performing the auxiliary problem (AP) in our problem, we had all values of ε_i $(i=1,\ldots,r)$ to be zero. Therefore, we can see that the obtained solutions at each iteration are all strong Pareto solutions.

Remark 3.3

By our computer (UNIVAC 1100/63E, 2.52 mips), it took us much more time (about 6 minutes CPU time) to solve the auxiliary problem (AP) than the original min-max problem (about 2 minutes CPU time). Therefore, we next tried to use an augmented norm:

$$S_a = \max_{1\leq i\leq r} w_i|f_i(x)-f_i^*| + (1/\alpha)\sum_{i=1}^{r} w_i|f_i(x)-f_i^*|. \tag{3.3}$$

As is well known, the solution minimizing the above augmented norm is a strong Pareto solution. However, unfortunately, it is not necessarily satisfactory, even if the given aspiration level is feasible (see, for example, Nakayama [4] and Sawaragi-Nakayama-Tanino [9]). However, in many problems, we can get a satisfactory Pareto solution by making the value of α sufficciently large. We shall show below the result using the above augmented norm. Here, instead of minimizing S_a directly, we solved the following equivalent problem with α=100.0:

$$(S_a') \qquad \underset{x,\ \xi}{\text{Minimize}} \quad \xi + (1/\alpha)\sum_{i=1}^{r} w_i(f_i(x) - f_i^*)$$

subject to

$$w_i(f_i(x) - f_i^*) \leq \xi \qquad i=1,\ldots,r$$

$$x\varepsilon X$$

	Pareto solution	aspiration level	sensitivity
f_1	244.1	300.0	0.189
f_2	423.1	520.0	0.808
f_3	10.83	10.9	0.033
f_4	0.07	2.90	0.0
f_5	1.20	2.90	0.0
f_6	0.49	0.90	0.0

We can see that the values of f_1, f_2 and f_3 have almost no difference from those resulted from the usual min-max problem: f_4 and f_6 are improved: f_5 is worse than the result using the usual min-max problem. We proceeded in the same way as the result using the usual min-max problem:

	Pareto solution	aspiration level	sensitivity
f_1	209.7	200.0	0.183
f_2	443.4	423.1	0.813
f_3	10.84	10.8	0.004
f_4	0.05	0.07	0.0
f_5	0.17	1.20	0.0
f_6	0.0009	0.49	0.0

	Pareto solution	aspiration level	sensitivity
f_1	173.8	180.0	0.822
f_2	464.3	490.0	0.000
f_3	10.83	10.8	0.063
f_4	0.87	0.90	0.075
f_5	0.16	0.17	0.040
f_6	0.0	0.0009	0.0

The CPU time for solving (S_a') was almost the same time as the one for usual min-max problem.

4. CONCLUDING REMARKS

Through our experiment, the following has been observed:

(i) In SWT method, we had to ask the decision maker his preference on many sample points. As a result, the decision makers got tired as the interaction with computer proceeded. However, in the satisficing trade-off method the satisfactory solution was obtained in only a few interactions. Moreover, the number of calculation by computer in the satisficing trade-off method was overwhelmingly fewer than that in SWT method.

(ii) In the multiattribute utility analysis, it took terribly much time for the decision makers to get accustomed with the method. We were obliged to give lectures many times to the decision makers in order that they can master the meaning of the method. In particular, lotteries consisted of the best level and the worst one was very difficult to understand, because the decision makers could not so much as imagine the situation under which the best level occur with the probability of 0.5 and the worst one with the probability of 0.5. In many engineering problems, the decision makers are usually engineers who have much experience in the problems. Therefore, although they are familiar with the real situation, they hardly think about hypothetical situations. The fact that the aspiration levels are very easy for the decision makers to think about makes the satisficing trade-off method very easy to carry out without any previous knowledge.

(iii) The above observation does not imply that the satisficing trade-off method is best for any kind of problems. However, in particular, it can be used most effectively in such problems as the one treated in this paper in which the mathematical formulae of criteria are given and their desirable levels vary depending on the situation.

(iv) In the satisficing trade-off method, we used three types of auxiliary optimization problems in order to get Pareto solution closest to the given aspiration level: the usual min-max problem, the min-max problem with an additional optimization problem and the augmented norm problem. In the min-max problem with an additional optimization problem, although the obtained solution is guranteed to be a strong Pareto solution, it took much more computing time than the other methods. In the augmented norm method, although the obtained solution is not necessarily satisfactory even if the given aspiration level is feasible, the computing time is almost the same as the one for the usual min-max problem and moreover the obtained solution is always guaranteed to be a strong Pareto solution. In practice, we have devised our program so that we may select one of them appropriately case by case.

REFERENCES

[1] Grauer, M., A. Lewandowski and A.P. Wierzbicki, DIDASS- Theory, Implementation and Experiences, in M. Grauer and A.P. Wierzbicki (eds.) Interactive Decision Analysis, Proceeding of an International Workshop on Interactive Decision Analysis and Interpretative Computer Intelligence, Springer, 22-30 (1984)

[2] Haimes, Y.Y., W.A. Hall and H.B. Freedman, Multiobjective Optimization in Water Resources Systems, The Surrogate Worth Trade-off Method, Elsevier Scientific, New York (1975)

[3] Nakayama, H., 'Proposal of Satisficing Trade-off Methd for Multiobjective Programming', Proc. of Society of Instrument and Control Engineering, Vol. 20, NO. 1, 29-35 (1984) (in Japanese)

[4] Nakayama, H., 'On the Components in Interactive Programming Methods', in M. Grauer, M. Thompson and A. Wierzbicki (eds.) Plural Rationality and Interactive Decision Analysis, Springer, 234-247 (1985)

[5] Nakayama, H. and Y. Sawaragi, 'Satisficing Trade-off Method for Interactive Multiobjective Programming Methods', in M. Grauer and A.P. Wierzbicki (eds.) Interactive Decision Analysis, Proceeding of an International Workshop on Interactive Decision Analysis and Interpretative Computer Intelligence, Springer, 113-122 (1984)

[6] Nakayama, H. and K. Furukawa, 'Satisficing Trade-off Method with an Application to Multiobjective Structural Design, Large Scale System, Vol. 8, 47-57 (1985)

[7] Nishikawa, Y. and J. Nomura, 'Multiobjective Optimization of Ratios for Blendiing Industrial Materials. Proc. 2nd Mathematical Programming Symposium, Japan, 159-183 (1981)

[8] Nishikawa, Y., J. Nomura and K Sawada, 'Applications of Multiobjective Optimization in a Firm', Communications of the Operations Research Society of Japan, Vol. 27, 325-332 (1982)

[9] Sawaragi, Y., H. Nakayama and T. Tanino, Theory of Multiobjective Optimization, Academic Press, 1985

Multiobjective systems analysis for impact evaluation of the Hokuriku Shinkansen (Super-express railroad)

Fumiko Seo[1], Masatoshi Sakawa[2]
Keizo Yamane[3] and Takao Fukuchi[4]

[1] Kyoto Institute of Economic Research, Kyoto University, Kyoto, 606
[2] Department of Systems Engineering, Kobe University, Kobe, 657
[3] Department of Management and Information, Setsunan University, Osaka, 572
[4] Institute for Social-Economic Planning, University of Tsukuba, Ibaraki, 305

1. Introduction

This paper presents a large-scale systems analysis for multiobjective impact evaluation of the Hokuriku Shinkansen (Super-express railway) Construction Plan in Japan. As assessment of the Shinkansen network construction, IIASA has published research results in 1977 and 1982. Although these works are pioneer in this field and include some interesting results such as commodity flow analysis based on spatial econometric model by Sakashita (1980), many researches are generally wide-ranging and methodological discussions have not enough been presented. This paper intends to present multiple impact assessment with some methodological discussions from the point of view of multiobjective systems analysis.

Our methodology is primarily based on multiattribute utility analysis for the system's evaluation but it is combined with the results of macro econometric modeling and simulation for future prediction. It is shown that intrinsic characteristics of both the methods can be effectively complemented one another in their combined utilization. The interactive computer program ICOPSS (Sakawa and Seo 1980, 1982) is used effectively for assessing the large-scale regional systems and for performing sensitivity analysis for construction of complementary policies.

2. Methodological Discussion

Large-scale systems analysis is usually confronted with incompatible requirements: preciseness in problem structuring and comprehensiveness in problem setting. On the one hand, problems included in a system should be structured in more precise forms quantitatively. For meeting this request, the problem setting for the large-scale systems analysis should be confined to a rather narrow particular range. On the other hand, the systems analysis demands as its intrinsic property more wide-ranging recognition of problems. Thus systems analysis should provide a combined utilization of different methods which possess their own characteristics for treating the complex systems. This paper intends to meet the gap between the precise problem structuring and the wide-ranging problem setting. In addition, we are particularly concerned with a hierarchical configuration for treating the multiple objectives systems. Our approach is composed of the macro-econometric modeling and the multiattribute utility analysis. The macro-econometric method provides the effective forecasting with quantitative structuring of economic aspects under a given parameters with statistical estimations, and the multiattribute utility analysis provides a comprehensive evaluation of a problem complex with multiple objectives depending on the subjective judgement. Thus these methods have their own limitations. In this paper, both of these methods are used for future prediction in a combination and, after a coherence check for the results obtained by both the methods, sensitivity analysis based on the multiattribute utility analysis is used for presentation of complementary policies (Figure 1).

First, the prediction is performed at the prefectural level for three regions (Toyama, Ishikawa, Fukui) using the macro-econometric model, and then the assessment is performed for Ishikawa prefecture at the local level for more detailed and more wide-ranging multiobjective evaluation using the multiattribute utility analysis.

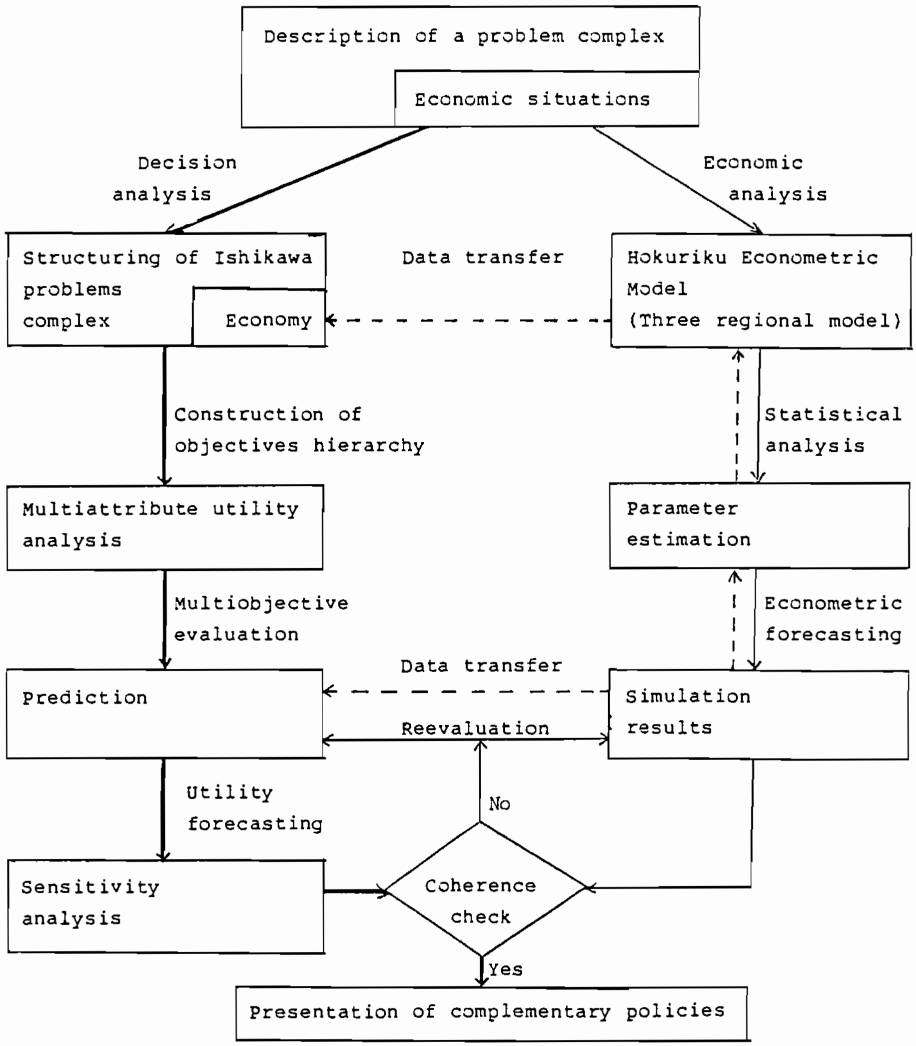


FIGURE 1 Large-scale systems analysis for Hokuriku Shinkansen construction.

3. Macro-Econometric Model (Three Regions)

For the macro-econometric model, the total region in Japan is partitioned into 11 regions of which 9 regions except Toyama, Ishikawa, Fukui prefectures are treated as the exogenous sector. For constructing the macro-econometric model, 35 equations in total including 35 endogenous variables and 17 exogenous variables are used and 27 equations are estimated with the least square method.

The particular macro-econometric model in this analysis provides a regional gravity model based on income potential ΠY_{ij}, which concerns interregional economic effects defined as a sum of prefectural income Y_j divided by "general distance" d_{ij} between regions i and j (major cities). The general distance is defined as a sum of the railroad passenger's fare C_{ij} and the time distance T_{ij} multiplied by a wage rate W_{Ri}. The income potential is interpreted as representing a regional demand pull, and makes an impact on prefectural income or labor productivity with a time lag. The public capital stocks G_{Ri} and G_{Ii} for transportation and industry-bases are introduced as exogenous variables and effects of them on the private capital stock K are considered. The public capital stock for life-base G_w is estimated as an endogenous variable. The Impact assessment of the Shinkansen construction in three prefectural levels is mainly performed on reduction of time-distances between 11 regions, and on its inducement effects for increase of the regional population flow, private investments and production. A flow chart showing structural interrelationship among variables is depicted in Figure 2 and a list of variables is shown in Table 1.

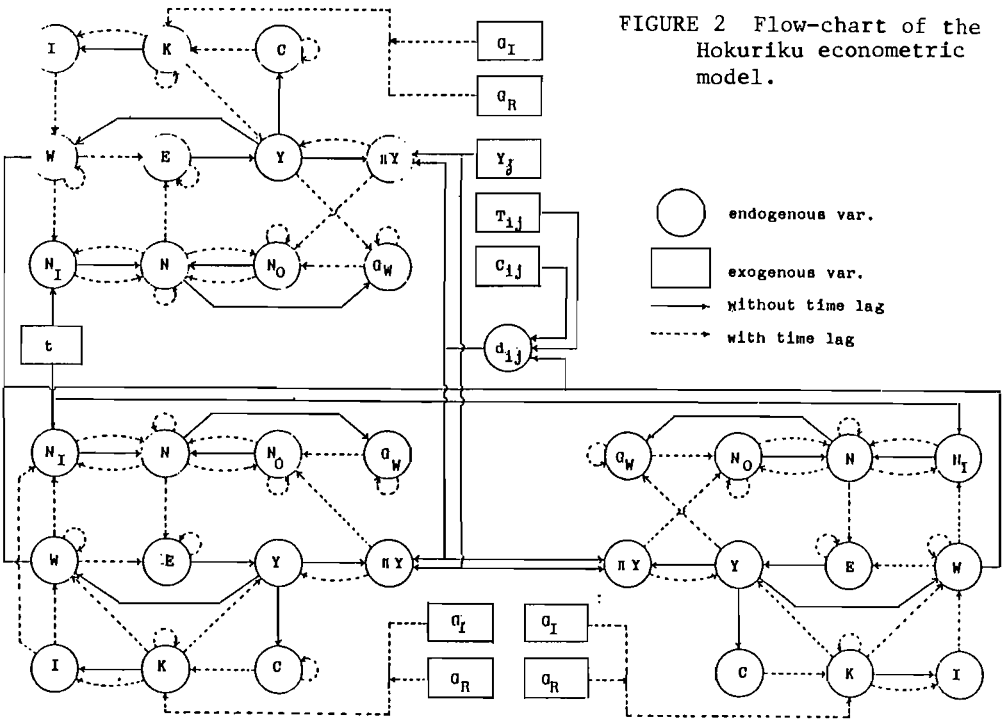


FIGURE 2 Flow-chart of the Hokuriku econometric model.

Simulation with the Hokuriku Econometric model are performed under the alternative assumptions for the total construction expenditure G_R (2250 billion or 2925 billion yens) and for the construction period (7 or 10 years). Exogenous policy variables are (i) interregional time distance (T_{ij}), (ii) cost distance (C_{ij}) and (iii) transportation-related public capital stock (G_R).

In short, the following predictions are obtained.

(i) After opening of the Hokuriku Shinkansen, reduction of the interregional time distance (T_{ij}) and increase of the interregional cost distance (C_{ij}) have considerable positive effects on the income potential ΠY (Toyama 23.35 ~ 25.58%, Ishikawa 17.35 ~ 19.05%, Fukui 5.18 ~ 6.33% increase each). These effects are particularly reflected on increase of net inflow of population and increase of income in Toyama.

(ii) In Ishikawa, there are the largest effects of the time distance (T_{ij}) reduction on inducement of the private investment (I), increase of capital intensive ratio (K/E) and labour productivity (Y/E). Thus in Ishikawa, activation of regional economy, in particular, qualitative changes of the local industrial structure are predicted.

Although the Hokuriku Econometric model provides effective forecasting of economic impacts on the region at the prefecture level, the large-scale systems analysis requires more comprehensive problem setting beyond the

TABLE 1 List of variables of the Hokuriku econometric model.

notation		name of variables	unit
endogenous variables	N	Residential population	1000
	N_1	Population inflow	1000
	N_0	Population outflow	1000
	E	Employment	1000
	Y	Income	million yen
	C	Consumption	million yen
	W	Wage rates	yen/hour
	I	Private investment for equipment	million yen
	K	Private capital stock	million yen
	G_w	Social capital stock for life-base	million yen
	ΠY	Income potential	index
	d_{ij}	General distance between i & j	index
exogenous variables	G_I	Social capital stock for industry-base	million yen
	G_R	Social capital stock for transportation-base	million yen
	Y_j	Prefectural income (j = 4,...,11)	million yen
	T_{ij}	Time distance (JNR)	hour
	C_{ij}	Cost distance (JNR)	yen
	t	Time trend	1965=1.0

economic modeling and forecasting. The multiobjective investigation treats the regional problems in more partitioned levels with a hierarchical configuration for Ishikawa prefecture.

4. Multiobjective Utility Analysis

4.1. Structuring of a Problems complex

The multiobjective impact assessment of the Hokuriku Shinkansen construction is performed based on structuring of the complex problems which the local communities face. The whole system is decomposed in regional, industrial and functional subsystems; which are structured in nine levels of a hierarchical system. The first level concerns a comprehensive evaluation of the Ishikawa prefecture. The second level is partitioned into "region" which can be separately evaluated for each subregion and "prefectural administration" which can not be separately evaluated. In the third level, the "region" is partitioned into "Kanazawa", "Hakusan", "Kaga" and "Noto" districts and evaluated according to each peculiarity for each district. "Kanazawa" is an urban area which include the capital city of Ishikawa prefecture, Kanazawa city, where the administrative functions of the prefecture are agglomerated and traditional small industries such as the Kutani potteries, Lacquer ware, and Japanese cloth and cakes are located. "Hakusan" is a natural environment conservation area which is mainly composed of the Hakusan National Park, but also partially includes industrial areas such as Mikawa, Tsurugi and Nonoich towns where some modern industries are located. "Kaga" is a famous sightseeing and health resort area which is composed of the Kaga Hotsprings country (Yamanaka, Yamashiro, Katayamazu, and Awazu Hotsprings, etc.) and the Kaga Beach, but also includes major industrial areas such as Komatsu city, where the Komatsu Manufacturing Company and the Komatsu airport are located, and Yamanaka town which is famous for the Yamanaka Lacquer ware. "Noto" is well known for the Noto Peninsula National Park and includes the Wajima city, which is famous for Wajima Lacquer ware, and electric and atomic power plants sites in Nanao city and in the Hagui district. Noto is also important forestry and fishery areas.

Those areas are evaluated in three subsystems such as "industry" "environment" and "public service". The industrial decomposition is performed in six levels from the fourth to the ninth. For "manufactures", eight industries are examined which are largely partitioned to "local" and "newly developed". For "commerce", eleven groups are considered, which are largely divided into "retail trades" and "sightseeing". The functional decompositions are performed for each industry. Each manufacturing industry is evaluated for its "activity" and "stability". "Retail trade" is evaluated for "activity" (amount of sales), and "sightseeing" is for "activity" and "earnings". For manufactures, the "activity" is examined from different points of view of interest. Measurements are the value of shipment per capita from the social point of view, the earnings (value added) per employed person from the point of view of entrepreneurs, and the wage income per employed person from that of laborers. The "stability" is also considered from different points of view of interest. Measurements are the change in share of the value of shipment from the social point of view, the change in the number of business establishments from the point of view of entrepreneurs and the availability of production factors in terms of opportunity of employment from that of laborers.

"Environment" for each subregion is partitioned into "pollution" and "natural environment" in the fifth level. "Pollution" is evaluated for "stench", "noise and vibration", "air pollution" and "water pollution" in the sixth level. "Public service" is partitioned into "sanitation" and "medical care" in the fifth level, and "sanitation is evaluated for "garbage treatment", "sanitary service" and "pervasion of flush-toilet" in the sixth level.

"Prefectural administration" is decomposed in six levels. It is partitioned into "economy", "resources" "environment" and "administration" in the third level. In the fourth level, "economy" is partitioned into "construction", "finance" and "exports". "Resource" is partitioned into "water" and "electric power" supply, and "environment" is partitioned into "cultural", "social" and "natural" environments. In the fifth level, "cultural environment" is evaluated for "preservation of traditional culture" and "university and college education". "Social environment" is evaluated for "stability" which is assessed with the number of criminal cases per capita. "Natural environment" is evaluated for "nature preservation" and "pollution control". "Pollution control" is partitioned into "solid waste treatment" and "water pollution" in the sixth level, and the "water pollution" is evaluated for "rivers", "lakes and marshes" and "estuaries", with the number of cases of unfitness to each environmental standard, respectively, in the seventh level. "Administration" is evaluated for "equity" using an index which shows economic equality compared with the average of national income per capita.

The evaluation process is effectively assisted by the computer program ICOPSS (Interactive Computer Program for Subjective Systems) developed by the authors (Sakawa and Seo, 1980). The ICOPSS program is composed of MUFCAP by Sicherman (1975) and MANECON collection by Schlaifer (1971) and is revised for several devices. With the STRUCT command, the total systems structure can be depicted visually in a tree diagram (Figure 3).

The attribute or UNIF (singleattribute utility function) names and MUF names are formed by combinations of code names corresponding to the hierarchical systems structure. For example, INMNEMCO is an attribute name which represents the value of shipment per capita for measuring "activity" of "electric machinery" in the "newly developed industry" of "manufactures". ENPOW is a MUF name which represents the "water pollution" in "environment". In performing multiattribute utility analysis, 269 attributes in total are selected for assessment. The number of UNIF-values to be assessed also amounts to the same. This numerous attributes is not easy to be evaluated individually. However the system's decomposition into a hierarchical structure shown in the above configuration makes the evaluation work much easier than at first sight.

4.2. Evaluation of UNIF and MUF

Data set for the singleattribute utility functions (UNIF) is constructed with the UNISET command (Figure 4). UNIF is assessed for (1) LINEAR and (2) CONSTANT RISK attitudes. In LINEAR form, evaluation is performed in the form of U(X) = A + BX and, when (2) CONSTANT RISK, U(X) = A + B * EXP(-C * X). (3) DECREASING RISK attitude is also available in the form of U(X) = A * EXP(-B * X) + C * EXP(-D * X) + E. The DISPLAY command demonstrates the function form of UNIF and the results of parameter calculation. Input data sets are listed with the DEBUG command. The scaling constants for the representation forms of multiattribute utility function (MUF) are assessed with the KSET command (Figure 5). The MUF is assessed in the forms of (1) additive: $U(X) = \sum_{s=1}^{m} k_s u_s(x_s)$, $\sum_{s=1}^{m} k_s = 1$, and (2) multiplicative: $U(X) = \frac{1}{K} \prod_{s=1}^{m} (Kk_s u_s(x_s) + 1) - 1$,

```
STRUCTURE FOR KANAZAWA
level 1     level 2     level 3     level 4     level 5       level 6
----IN      ----INM     ----INML    ----INMLFO  ----INMLFOC   ----INMLFOCO
!           !           !           !           !             !
!           !           !           !           !             !--INMLFOCP
!           !           !           !           !             !
!           !           !           !           !             !--INMLFOCW
!           !           !           !           !
!           !           !           !           !--INMLFOS    ----INMLFOSN
!           !           !           !                         !
!           !           !           !                         !--INMLFOSL
!           !           !           !                         !
!           !           !           !                         !--INMLFOSS
!           !           !           !
!           !           !           !--INMLTE   ----INMLTEC   ----INMLTECO
!           !           !           !           !             !
!           !           !           !           !             !--INMLTECP
*
!           !           !           !           !             !
!           !           !           !           !             !--INMLTECW
!           !           !           !           !
!           !           !           !           !--INMLTES    ----INMLTESN
!           !           !           !                         !
!           !           !           !                         !--INMLTESL
!           !           !           !                         !
!           !           !           !                         !--INMLTESS
!           !           !           !
!           !           !           !--INMLFR   ----INMLFRC   ----INMLFRCO
!           !           !           !           !             !
!           !           !           !           !             !--INMLFRCP
!           !           !           !           !             !
!           !           !           !           !             !--INMLFRCW
!           !           !           !           !
!           !           !           !           !--INMLFRS    ----INMLFRSN
!           !           !           !                         !
!           !           !           !                         !--INMLFRSL
!           !           !           !                         !
!           !           !           !                         !--INMLFRSS
!           !           !           !
!           !           !           !--INMLWO   ----INMLWOC   ----INMLWOCO
!           !           !           !           !             !
*
!           !           !           !           !             !--INMLWOCP
!           !           !           !           !             !
!           !           !           !           !             !--INMLWOCW
!           !           !           !           !
!           !           !           !           !--INMLWOS    ----INMLWOSN
!           !           !           !                         !
!           !           !           !                         !--INMLWOSL
!           !           !           !                         !
!           !           !           !                         !--INMLWOSS
!           !           !           !
!           !           !           !--INMLCL   ----INMLCLC   ----INMLCLCO
!           !           !                       !             !
!           !           !                       !             !--INMLCLCP
!           !           !                       !             !
!           !           !                       !             !--INMLCLCW
!           !           !                       !
!           !           !                       !--INMLCLS    ----INMLCLSN
!           !           !                                     !
!           !           !                                     !--INMLCLSL
!           !           !                                     !
!           !           !                                     !--INMLCLSS
!           !           !
```

FIGURE 3 Structure for KANAZAWA.

$\sum_{s=1}^{m} k_s \neq 1$, $-1 < K$. The LISTU and LISTK commands can demonstrate all the data of UNIF and MUF in a tabular form according to all the levels. This device will make it easy to compare the scale-value assessment for each subsystem.

In the process of finding indifferent points among attributes in "industry", the industrial characteristics in each subregion are deliberately taken into consideration by assessing their values relatively in industry × district matrix forms. Similarly, attributes in "environment" and "public services" are relatively assessed according to their characteristics in each subregion.

The evaluation works in every stage have been performed quite intensively with a four-days brainstorming by a special team composed of knowledgeable representative people who are actually engaging in investigation and management in prefectural businesses associations, and MCDM researchers. The predicted UNIF and MUF values are assessed with the unvaried UNIFs and scaling constants in ten years after. ALT1 shows the data set for the current values and ALT2 for the predicted values of the selected attributes.

The evaluation results are as follows.

1) When the Hokuriku Shinkansen begins activities all along the line in 1995, an overall MUF-value for Ishikawa prefecture (ISHIKAWA) will be slightly raised. This effect is larger for prefectural administration (PREFECT) than for each subregion (REGION) although the MUF value is higher for REGION.

```
COMMAND:
=UNISET
INPUT UNIF NAME:
=INMLFOCO
WANT LIST OF UNIF TYPE?
=YES
LIST OF UNIF TYPE


  (1) LINEAR
  (2) PIECEWISE LINEAR
  (3) CONSTANT RISK
  (4) DECREASING RISK
INPUT UNIF TYPE:
=3
INPUT RANGE(WORST & BEST) OF THIS UNIF:
=0.0 50.0
INPUT 50-50 LOTTERY(WORSE PAYOFF;BETTER PAYOFF & C.E.):
=0.0 50.0 21.0
ANOTHER UNISET?
=YES
INPUT UNIF NAME:
=INMLFOCP
WANT LIST OF UNIF TYPE?
=NO
INPUT UNIF TYPE:
=3
INPUT RANGE(WORST & BEST) OF THIS UNIF:
=0.0 1500.0
INPUT 50-50 LOTTERY(WORSE PAYOFF;BETTER PAYOFF & C.E.):
=0.0 1500.0 625.0
ANOTHER UNISET?
```

FIGURE 4 Demonstration of UNISET command.

2) Improvement of the regional MUF-values are the largest in HAKUSAN and the smallest in NOTO.

3) The MUF value in each region will largely increase for "industry" but will decrease for "environment". Particularly large environmental destruction is predicted in NOTO, and the MUF value for "environment" will decrease 8 points (as 1/100 unit) from 0.77 to 0.69.

4) In KANAZAWA, the current degree of satisfaction with "environment" is the lowest. This is because of low UNIF values for noise (ENPOZ) and air pollution (ENPOA) (5 points and 9 points down respectively), and it is predicted that the current situation will be accelerated by the opening of the super-express railway. In "industry", increase of the MUF value for sightseeing (INCH) is the largest (10 points up). Almost all groups of "manufactures" are expected to increase their MUF-values but retail trades are not.

5) In HAKUSAN, the current degree of satisfaction with "environment" is the highest (0.98), but will slightly decline (to 0.97) after the opening because of noise (ENPOZ) (0.71 to 0.60) and air pollution (ENPOA) (0.81 to 0.71). On the other hand, current satisfaction with "public service" is the lowest (0.65). This is because of the law value for medical care (PBMD) (0.5) whose improvement is not at all anticipated. In "industry", increase of the MUF values are anticipated for all groups,

```
COMMAND:
=KSET
INPUT MNAME:
=INMLFOC
WANT LIST OF THE METHOD FOR KSET?
=YES
LIST OF THE METHOD FOR KSET
  (1) INPUT K'S VALUES DIRECTLY
  (2) INPUT ONE INDIFFERENCE PAIR AND LOTTERY
  (3) INPUT TWO INDIFFERENCE PAIRS
WHICH METHOD DO YOU USE?
=2
INPUT REFERENCE UNAME:
=INMLFOCO
INPUT THE FOLLOWING ANS1 AND ANS2:
( INMLFOCO , INMLFOCP ) = ( ANS1           ,      0.         )
        IS INDIFFERENT TO (      0.        , ANS2            )
(INPUT ATTRIBUTE VALUES)
=35.0 1500.0
( INMLFOCO , INMLFOCW ) = ( ANS1           ,      0.         )
        IS INDIFFERENT TO (      0.        , ANS2            )
(INPUT ATTRIBUTE VALUES)
=32.5 800.0
INPUT P SUCH THAT
      LOTTERY --- ALL ARE BEST WITH PROBABILITY P
              !- ALL ARE WORST WITH PROBABILITY 1-P
AND
      CERTAINTY CONSEQUENCE --- INMLFOCO IS BEST
                            !- THE OTHERS ARE WORST
ARE INDIFFERENT:
=0.75
K( INMLFOCO ) =    0.7500
K( INMLFOCP ) =    0.7064
K( INMLFOCW ) =    0.6936
* CAPITAL K =   -0.97266
ANOTHER KSET?
=YES
```

FIGURE 5 Demonstration of KSET command.

particularly for machinery (except electrical) (INMNMC) (14 points up), textile (INMLTE) (10 points up) and sightseeing (INCH) (10 points up).

6) In KAGA, the current degree of satisfaction with "industry" is the highest and is predicted to increase (0.87 to 0.92). "Public service" is the lowest (0.68) and its improvement is not expected. On the other hand, it is predicted that degree of satisfaction with "environment" will be lowered (0.71 to o.66 at the lowest). This is mainly because of noise (ENPOZ)(0.50 to 0.38). Degree of satisfaction with "industry" is the highest for electrical equipment (INMNEM) (0.87 to 0.92), second for ceramics (INMLCL) and third for nonelectrical machinery. Predicted raise of the degree of satisfaction after opening of the Shinkansen is the largest for sightseeing (INCH) (0.80 to 0.93), and second for foods (INMLFO) (0.78 to 0.87).

7) In Noto, the MUF-value for "industry" is the highest under current conditions (0.84). This is because of foods (INMLFO) and textiles (INMLTE). Those industries are expected to increase their MUF-values (0.83 to 0.91 and 0.80 to 0.90, respectively). Sightseeing (INCH) and electrical machinery (INMNEM) are most expected to raise their MUF value (0.65 to 0.78 and 0.75 to 0.87, respectively).

8) For prefectural administration, improvement of the cultural environment (ENU) is most expected, in particular in college and university education (ENUU) which is predicted to obtain the largest UNIF-value (0.60 to 0.80). Raise of the UNIF-value for exports is most expected in machinery (ECEIMC)(0.67 to 0.83), second in silk fabric (ECEDSK) (0.31 to 0.50) and third in lacquer ware (ECEDSH) (0.36 to 0.55). On the other hand, deterioration of the social environment is predicted (0.84 to 0.65). This is because of increase of social strain as a result of permeation of the life styles of the urban society. Aggravation of the nature environment (ENN) is supposed to be large because of degradation of the nature preservation (ENNS) (0.85 to 0.57).

9) For industries, other findings are summarized (Table 2).

(i) Business activities in manufactures are generally predicted to increase. However, some groups will suffer from decrease of stability.

(ii) Business activities in the retail trades are generally predicted to suffer from severe competitive situations introduced by increase of new

TABLE 2 Prediction for business activity of retail trade.

	(increase)	(decrease)
Kanazawa	pottery and porcelain, other utensiles	Japanese draperies, women's and children's clothing, furnitures, machine and tools
Hakusan	(no change)	(no change)
Kaga	pottery and porcelain, other utensiles	Japanese draperies, women's and children's clothing, liquours, furnitures machine and tools
Noto	pottery and porcelain, other utensiles	(no change)

entries from metropolitan areas, particularly in Kanazawa and Kaga district. It is anticipated that local special products having traditional reputations such as pottery and porcelain, lacquer ware and silk products will be placed under pressures by entry of large business enterprises.

(iii) In Kanazawa, more precisely, the current UNIF-value for stability of foods and beverages is the least for the change in share of shipment (INMLFOSS), but predicted to increase after opening of the Shinkansen (0.34 to 0.59). On the other hand, UNIF-value for labor availability in stability of textiles (INMLTESL) will be decreased to the least (0.50 to 0.37). In Hakusan, current UNIF-value for stability of furniture is very low for the change in share of shipment (INMLFRSS), but predicted to be improved largely (0.19 to 0.35). Lumber and wood products are in similar situations, but ceramics is predicted to decrease in the share of shipment (INMLCLSS 0.39 to 0.33). In Kaga, although the current UNIF-value for the change in share of shipment of foods and beverages (INMLFOSS) is the least, large improvement is expected (0.33 to 0.50). In furniture and lumber and wood products, no improvement for stability is predicted. In Noto, the current degree of satisfaction is the highest for labor availability in textiles (INMLTESL), but large reduction is predicted (0.72 to 0.57). On the other hand, a large improvement over all the items for activity of foods and beverages, such as the value of shipment (INMLFOCO), earnings (INMLFOCP) and wage income (INMLFOCW), is predicted.

It should be noted here that the MUF-values have a tendency to be gradually heightened in the upper levels of the preference structure because of cumulative effects of scaling constants for the multiplicative MUF types. Thus MUF-values include upward biases in the process of nesting MUFs and comparison of MUF values should be performed among those in the same levels of the preference structure.

4.3. Proposal of alternative policy programs

One merit of the MUF method is to provide alternative policy programs based on efficiency criteria in terms of the utility values. For this purpose, sensitivity analysis is used for the data set composed of current values of attributes, UNIFs and MUFs. Using numerical results of the sensitivity analysis, the attributes which will bring the most efficient improvement of the utility values are selected and their values are changed discretionally from the different administrative points of view. Alternative data sets are constructed with the changed values of the selected attributes and used as alternative policy programs. It is known that improvement of the attribute values for pollution and public services will raise effectively the degree of satisfaction in each subregion. For MUF-values, increase of activities of nonelectrical machinery and electric and electronic equipment in manufactures and of activity of sightseeing in commerce will effectively raise the degree of satisfaction with regional industries in each district. In addition, increase of activities of textile in Hakusan and apparels in Kanagawa, Hakusan and Noto will also be effective for local satisfaction. Thus enforcement of administrative guidance based on alternative policy programs for selected fields of attributes will contribute to adjusting regional conditions occurred by the opening of the Hokuriku Shinkansen.

5. CONCLUDING REMARKS

Future forecastings for project evaluation with the MUF method depend on predicted UNIF-values which are assessed subjectively and intuitively even though they are obtained depending on various kinds of quantitative and non-quantitative information. For mitigating this judgemental property of prediction, which is unavoided in multiattribute utility analysis, a combined use of the MUF method with other forecasting techniques such as the econometric method will be recommended. While the econometric forecasting can provide a simulation results as a good reference under the given parametric structure of deductive economic models, future predictions with constructed UNIF-values can be performed in fact independently of the model-forecasting results. Then a consistency check is performed. If unconsistency is found, then reevaluation processes are initiated for both the procedures. The present case study brought a consistent result for both the approaches.

In the process of MUF assessment, interactive utilization of a computer program is effective to obtain reasonable results because the assessment should be checked iteratively in the light of experiences and prospects of knowledgeable DM. However, the assessment is performed still deterministically. Thus manipulation of diversity and ambiguity of assessment should explicitly be taken into consideration. This problem is combined with treating multiagent decision making to some extent. Biases of assessment due to a singularity of DM should be resolved on examining multiagent decision problems. Although the research for the collective choice problem has been already presented by the authors (Seo and Sakawa 1984, 1985), there still remains a lot of points to be discussed in future.

Acknowledgement

This research is partially supported by the Kajima Foundations Research Grant.

References

Hokuriku Society of Economic Survey, 1985, *Impact Assessment of Hokuriku Shinkansen for Regional Economy.*

Sakashita, N. 1980, Application of the Spatial Econometric Model (SPAMETRI) to the Evaluation of the Economic Effects of the Shinkansen, in *The Shinkansen High-Speed Rail Network of Japan*, ed. by A. Straszak and P. Tuch, Pergamon Press, Oxford.

Sakawa, M. and F. Seo, An Interactive Computer Program for Subjective Systems and its Application, International Institute for Applied Systems Analysis, Laxenburg, Austria, Working Paper, WP-80-64, 1980.

Sakawa, M. and F. Seo, Integrated Methodology for Computer-aided Decision Analysis, in *Progress in Cybenetics and Systems Research*, vol. X, R.T. Trappl, F. de P. Hanika, and R. Tomlinson, Eds. Washington: Hemisphere, 1982, pp. 333-341.

Seo, F. and M. Sakawa, 1984, An Experimental Method for Diversified Evaluation and Risk Assessment with Conflicting Objectives, *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-14, No. 2, 213-223.

Seo, F. and M. Sakawa, 1985, Fuzzy Multiattribute Utility Analysis for Collective Choice, *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-15, No. 1, 45-53.

Straszak, A. and R. Tuch, (ed.) 1980, *The Shinkansen High-Speed Rail Network of Japan, Proceedings of an IIASA Conference, 1977* Pergamon Press.

Straszak, A. (ed.) 1981, *The Shinkansen Program: Transportation, Railway, environmental, Regional and National Development Issues*, IIASA Collaborative Proceedings Series, IIASA Laxenburg, Austria.

A QUEUEING THEORETIC MODEL FOR PLANNING DIAGNOSIS SYSTEMS IN HOSPITALS

Günter Fandel,
Holger Hegemann, Hagen*

1. Object of the investigation

In the last twenty years hospitals have developed into cost and investment intensive service enterprises confronted with a rising pressure of costs. Apart from an enormous expansion in personnel costs the reasons can be seen in a higher equipment with apparatus based on the rapid developments in the medical-technical progress and the therewith attendant and fast increasing demand for hospital services.

Aspects of medical necessity and effectiveness are decisive for the production of health services in hospitals; but this process must be reviewed on the basis of efficiency considerations as well. This follows from the mere fact that a hospital must be regarded as a consumer oriented facility having to provide certain medical services efficiently by combining different production factors. Although some methodical and legal deliberations have already been made in this direction, the German hospital system does not yet satisfy those economic criteria being already fulfilled in other industrialized nations. In this context the fact is characteristic, that hospital management was not represented at the health economy congress of the Verein für Socialpolitik 1985 by a particular study group.

This argumentation applies correspondingly to the diagnostics being of central significance to the service production in hospitals. For, this subsystem is claimed as the first service area by the patients and the examination results received hereby give important orientation marks for the following therapy and care.

The inexpedient allocation of medical staff and equipment especially in the diagnostics would lead to considerable differences in the load degrees of the different diagnostic service units and to increased waiting times for patients in front of the diagnostic rooms. The latter would entail a retardation of the following therapy and care. In order to avoid such misallocations and their consequences capacity planning in the clinical diagnostics on the basis of the methods of operations research is indispensable.

So far, only subsystems of the clinical diagnostics have been analyzed in the literature by means of network-planning-technique (Fleck 1977, Taylor/Keown 1980), queueing theory (Bailey 1954, Taylor/Templeton 1980) and simulation (Carruthers 1970, Gierl 1976, O'Kane 1981, Revesz/Shea/Ziskin 1972). The above mentioned problem, however, can only be solved by a central capacity planning including the interdependencies existing in the form of patients' ways among the single diagnostic units of a hospital. In addition, under instrumental aspects it must be objected, that the network-planning-technique does not allow the consideration of the patient's arrival behaviour in the diagnostic area and that the simulation as a method without the quality of convergency will remain inferior to the queueing theory as long as the special planning situation allows the application of this exact method.

---

*
Prof. Dr. Günter Fandel and Dr. Holger Hegemann, Chair for Business Administration, Fernuniversität, Roggenkamp 6, D-5800 Hagen 1.

The authors are especially indebted to the former medical director, Dr. Klaus Pieper, and the team of doctors of the diagnosis system at the General Hospital of Hagen City for having made possible the empirical study carried out in this context.

Consequently, a queueing model, which shall describe the structure of the diagnostic system and the diagnosis process by means of patients flows as close to reality as possible, is chosen as the basis for the analytical investigations to be carried out within the scope of capacity planning. In contrast to the approaches in the literature, in which only single diagnostic rooms or units and actually only one patient flow were were objects of the model considerations, here however two modifications are necessary. On the one hand, the entire approach demands that a network-oriented total system of the diagnostics is taken as the basis. On the other hand, attention has to be paid to the fact that this general system is passed through by two patient streams simultaneously resulting from the examination processes of out-patients and in-patients. With regard to the management and optimal configuration of such systems it can be demonstrated, though, that the queueing models wellknown from the literature (Gross/Harris 1974, Jackson 1963) can also be extended easily to such a generalized problem situation under certain assumptions (Hegemann 1985).

The queueing model for capacity planning presented in the following can be applied to the planning of measures for constructing new or extending old buildings in the clinical diagnostics of hospitals. At the same time, it allows to reveal weaknesses within existing diagnosis systems by means of system indices like patients' waiting times, load degrees of the apparatus and the length of the queues and to show alternatives for an improvement of the capacity structure of the diagnosis area. Thereby the long-term optimal capacity structure is often characterized by waiting times of patients in front of diagnostic service units not exceeding a given reasonable limit and reaching a possibly high load degree for the single examination rooms as well as for the total diagnostic system.

Finally, the methodical reflections will be exemplified for the diagnosis system of the General Hospital of Hagen City, which contains 650 beds.

2. Queueing theoretical description of a diagnosis system

2.1 An outline of the diagnostics and their particularities

Beside therapy, care, and medical education and research the diagnostics are part of the medical area of a hospital, which is differentiated from the economic and administrative area containing mainly management tasks. The numerous relations possibly existing between the diagnosis system and other subareas of a hospital are not object of the reflections. The considerations rather concentrate on the diagnostics as a system, their system elements and the interdependencies occurring among them as well as on the derivation of suitable system indices characte-rizing the quality of the diagnosis system under economic aspects. The diagnostics possess their external relations in the flows of out-patients and in-patients entering, passing through and leaving the diagnosis system for reasons of diagnosis on the initiative of private doctor's practices or by order of the care and therapeutical departments of the hospital.

The diagnostics can be viewed as the totality of all those activities necessary for the production of a diagnosis; so these activities delineate the way to a diagnosis. Under system analytical aspects the diagnostics are usually divided into the sections X-ray diagnostics, laboratory diagnostics, function diagnostics, and endoscopy (Eichhorn 1976, p. 281f.), which themselves can be subdivided into single service units or diagnosis rooms. In these service units being elements of the diagnosis system single diagnostic services are produced in the form of examinations by combining medical personnel, diagnostic equipment and means of medical needs as resources. The medical staff consists of doctors, medical-technical assistants, medical helpers as well as, occasionally, laboratory assistants and operating nurses. The actual diagnosis generally emerges from the composition of such individual diagnostic services in different service units; so the diagnosis ensues within the scope of a diagnostic process or is its result, respectively.

Within the diagnosis system the service units, generally united in departments under the disciplinary responsibility of a head physician, are connected by the flows of in-patients and outpatients and the information flux accompanying them. These flows result from the diagnostic processes arranged by doctors of the different medical departments in dependence of the sickness of the individual patient to set up the diagnosis. These processes, normally different for every patient, can be seen as disease-specific parts of the total diagnostic process because of their disjunctive connection of individual diagnostic examinations. Usually they are not fixed a priori, but are mostly made concrete by the doctors during the examination. So the patients' ways via the diagnostic service units, essentially determining the internal relation structure of the total diagnosis system, must be considered as stochastic.

It is obvious to understand the diagnosis system as a waiting system and therefore to make it accessible to a queueing theoretical analysis. The diagnostic rooms can be interpreted as service stations, the patients are the traffic units, the waiting-room is the queue, and the flows of the out-patients and inpatients correspond to the traffic streams. Thus it is possible to describe the formal structure of the diagnosis system as a discrete, dynamic, stochastic, and open system by the number of diagnostic service units, the stochastic arrival behaviour of the out-patients and in-patients, the stochastic examination times in the service units as well as by the stochastic patients' ways with the transition probabilities from one service unit to the next. This, however, is not possible in the scope of a conventional queueing model; on the contrary, two extensions are necessary. Firstly, it has to be taken into account that the demand flow consists of two partial streams. Secondly, the stochastic structure of the diagnosis processes being dependent on the patients' ways has to be regarded. Whereas directly usable clues can be found in the queueing theoretical literature for the first modification, which can be taken over as results here, the necessary second modification is mastered by adding the GERT network-model (Pritsker/Happ 1966, Neumann/Steinhardt 1979) to the queueing model in order to conceive realistically the stochastic arrival processes in the form of patients'ways. To a certain extent, both planning procedures are combined by including the exclusive-or-branchings known from the GERT model and therewith the transition probabilities of the patients between the diagnosis rooms into the queueing theoretical model construction. On this basis realistic system indices can be determined for the total diagnosis system giving information about how well the single service units are coordinated and allowing secure statements about the service quality and the efficiency of the total diagnosis system and its partial elements.

In order to judge about the efficiency and service quality of a queueing system diagnostics the following system indices for the single service units i, i=1,...N, as well as for the total system are of interest:

(1) the expected waiting time of a patient in front of a diagnostic room and in the total system; it may be denoted by $E(V_{q_i})$ and $E(V_q)$.

(2) the expected time a patient spends in a service unit and in the total system; it is indicated by $E(V_i)$ and $E(V)$.

(3) the expected number of waiting patients in a service unit and in the total system; it is expressed by $E(L_{q_i})$ and $E(L_q)$.

(4) the expected number of patients in a service unit and in the total system; it is symbolized by $E(L_i)$ and $E(L)$.

(5) the expected load degree of a service unit and of the total system; it will be denoted by ξ_i and ξ.

These indices serve also as a basis for changes in the capacity structure of a diagnosis system to be carried out eventually.

Among these system indices the expected waiting time of a patient and the expected load degree of the service units and the total system are relevant with regard to health policy. These two criteria, which are to be minimized or maximized, are in conflict with each other being known from analogous areas of industrial production as the dilemma of scheduling. In the field of health policy this conflict is mostly solved in the way that the load degrees are maximized under the constraint that the waiting times do not exceed socially reasonable limits.

2.2 Formal description of the model of a diagnosis system

2.2.1. Premises

The queueing theoretical analysis to be carried out later on is based on the following premises:

(P 1) The outpatients and inpatients arrive at the diagnosis system from outside one by one; the arrival processes are Poisson processes with the intensities λ_o^a (for the out-patients) and λ_o^s (for the in-patients), i.e. the interarrival times are exponentially distributed with the parameters λ_o^a and λ_o^s. The subindex zero indicates the outside of the diagnosis system as the source and sink of the patients flows.

(P 2) The number of potential patients, who can enter the diagnosis system, is regarded as unlimited.

(P 3) The patients remain in the queue of the respective diagnostic service unit they joined after their arrival up to the beginning of the examination.

(P 4) The examinations of the patients in the respective diagnosis rooms take place individually and are not dependent on the arrival processes.

(P 5) The examination times of the diagnostic actions in the service units are exponentially distributed with the parameters μ_i, $i=1,\dots,N$.

(P 6) An examination once started will be finished without interruptions.

(P 7) Breakdown periods of the equipment are not considered.

(P 8) The diagnosis system consists of N service units; each diagnostic service unit i, $i=1,\dots,N$, can be composed of $n_{p_i} \in \mathbb{N}$ parallel and identical diagnostic rooms.

(P 9) If a patient has once left a diagnostic service unit i, he cannot enter it immediately again, i.e. the transition probability r_{ii} is equal to zero for all i $\in \{1,\dots,N\}$.

(P10) The ways of patients from a service unit i to a service unit j are independent of each other for all i, j $\in \{1,\dots,N\}$, $i \neq j$.

(P11) Times needed for the ways between service units may be neglected.

(P12) The way of a patient through the diagnosis system is free of cycles; so he enters each service unit once at most.

(P13) Waiting-room capacities in the diagnostic service units are unlimited.

(P14) Patients enter the diagnosis rooms according to the service discipline "first come - first served" (FCFS).

(P15) There are two independent sources of the flows of outpatients and in-patients.

The premises (P 1) and (P 5) were backed by the results of the empirical study - that will be reported later -, so that their supposition which can also be found in the literature is of no problem. In addition (P 5) allows the arbitrary joining or separation of diagnostic examinations without changing their distribution character. This is relevant for an eventual organizational rearrangement of diagnostic examinations necessary to reach an optimal capacity structure; that will be discussed more detailed below.

(P 9) and (P12) exclude backward loops which would contradict (P 4). Otherwise the arrival processes in front of the service units would no longer be Poisson processes (Walrand 1982, Malamed 1979a, 1979b); this quality, however, is decisively necessary for the following analysis. Moreover, (P12) was backed by the empirical study.

2.2.2 The formal model of a general diagnosis system and its characteristics

Under the premises stated above the diagnosis system can be understood as a network-type queueing system of $N \in \mathbb{N}$ service units i, $i=1,\ldots,N$, arranged in parallel or in series, which is run through by q=2 different patient flows as graphically demonstrated in Figure 1. In this connexion

- λ_i^a, λ_i^s and λ_j^a, λ_j^s, respectively, indicate the flows of out-patients and in-patients going to and leaving the outside of the diagnosis system and the diagnostic service units in the diagnosis system with $i, j \in \{0,\ldots,N\}$.
- r_{ij}^a, r_{ij}^s and r_{ji}^a, r_{ji}^s, respectively, indicate the transition probabilities of the out-patients and in-patients from unit i to unit j and vice versa with $i, j \in \{0,\ldots,N\}$, $r_{ij}^a = r_{ji}^a = r_{ij}^s = r_{ji}^s = 0$ for i=j (premise (P 9)) and

$$\sum_{j=0}^{N} r_{ij}^a = \sum_{j=0}^{N} r_{ij}^s = 1, \quad i=0,\ldots,N.$$

On account of the premises (P 1), (P 5), (P 8), (P13) to (P15) the diagnosis system can be typified extending Kendall's notation (1953) by the queueing model M/M/N/∞/FCFS/2; the first two symbols indicate that the distributions of the patients' interarrival and examination times are Markovian. If there is only one diagnostic service unit with one diagnosis room or n_p parallel diagnosis rooms, one gets the well-known model typifications M/M/1/∞/FCFS/2 and M/M/n_p/∞/FCFS/2, respectively.

Furthermore the general diagnosis system is characterized by the fact, that a branching of the patients' flows in the meaning of an exclusive-or-knode appears behind the outside of the diagnosis system and every diagnosis room and that a fusion or a superimposition of several patients' flows occurs in front of every diagnostic service unit and the outside of the diagnosis system. According to the patient's type these flows are ordered in out-patients and in-patients. Because of the following three basic theorems all (partial) flows of out-patients and in-patients within the diagnostic system are Poisson processes.

Theorem 1 (Burke 1956, 1972):

The departure processes of an M/M/1/∞/FCFS/1- and an M/M/n_p/∞/FCFS/1-queueing system being in equilibrium are Poisson streams and have the same intensities as the arrival streams.

Theorem 2 (Klimow 1979):

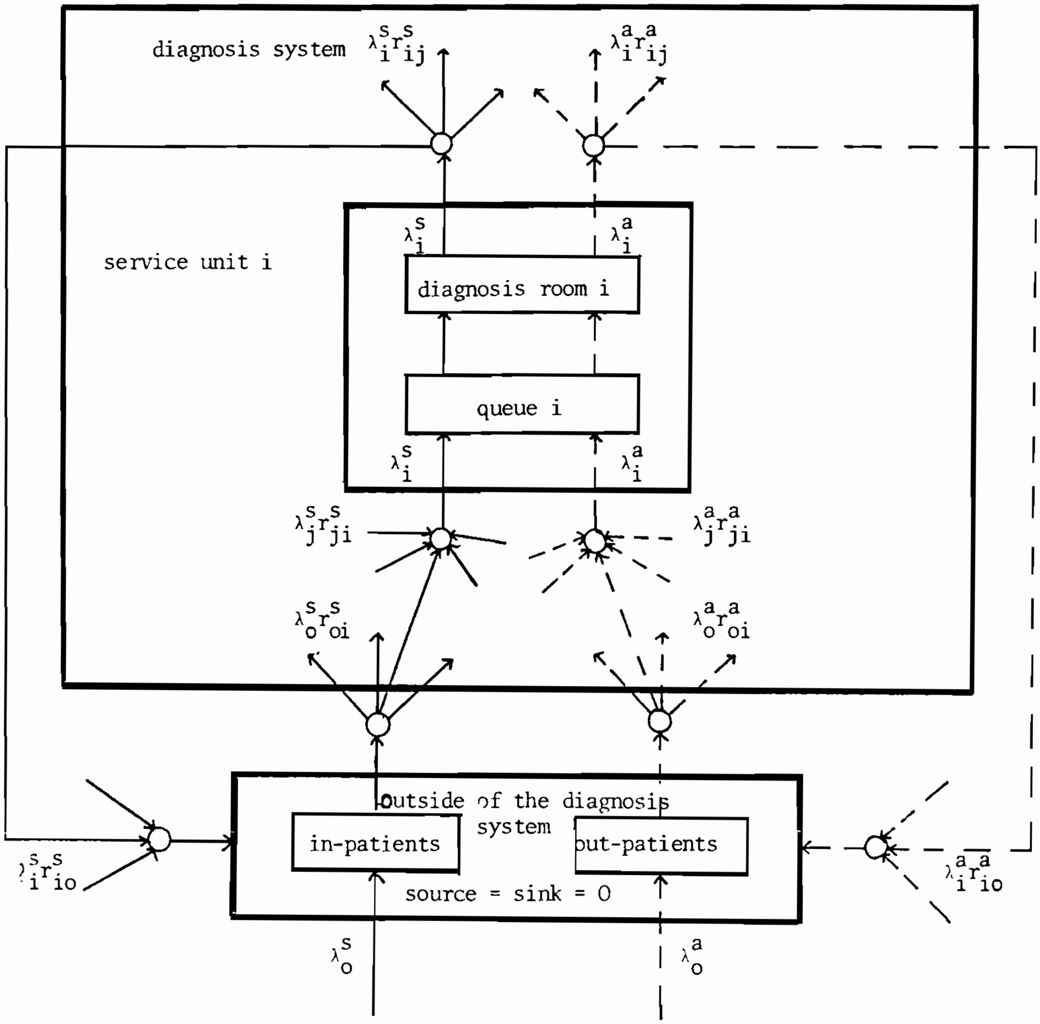
The composition of c independent Poisson streams with the intensities $\lambda_1,\ldots,\lambda_c$, $c \in \mathbb{N}$, is again a Poisson stream with the intensity $\lambda=\lambda_1+\ldots+\lambda_c$.

Theorem 3 (Ferschl 1964):

If the mobile units of a Poisson stream with the intensity λ are stochastically distributed over $c \in \mathbb{N}$ partial streams in such a way that subsequently a unit is within the partial stream i with the probability r_i $(i=1,\ldots,c)$, then the c partial streams are Poisson streams each with the intensity $r_i\lambda$.

This fact is of advantage for the following parts of the analysis. On the one hand it is possible to determine the arrival rates λ_i in front of the single diagnostic service units just by means of the transition probabilities r_{ij}^a and r_{ij}^s together with the total arrival rates λ_0^a and λ_0^s. On the other hand this firstly allows to analyze

Figure 1: The diagnosis system as a queueing system



the single diagnostic service units separatly by applying queueing theoretical considerations and to describe their efficiency by means of indices before, in a second step, one obtains the indices of the total diagnosis system from the combination of the indices of the single diagnostic service units.

Now, $\lambda^a=(\lambda_0^a,\ldots,\lambda_N^a)^T$ may denote the vector of the arrival rates of out-patients and $\lambda^s=(\lambda_0^s,\ldots,\lambda_N^s)^T$ that of in-patients, and

$$R^a=\begin{pmatrix} 0 & r_{01}^a & \ldots & r_{0N}^a \\ r_{10}^a & 0 & \ldots & r_{1N}^a \\ \vdots & & & \vdots \\ r_{N0}^a & & \ldots & 0 \end{pmatrix} \quad \text{and} \quad R^s=\begin{pmatrix} 0 & r_{01}^s & \ldots & r_{0N}^s \\ r_{10}^s & 0 & \ldots & r_{1N}^s \\ \vdots & & & \vdots \\ r_{N0}^s & & \ldots & 0 \end{pmatrix}$$

may indicate the so-called routing matrices of the out-patients and the in-patients with the corresponding transition probabilities. Then the relations

$$\begin{aligned} \lambda^a &= R^{a^T}\lambda^a, \\ \lambda^s &= R^{s^T}\lambda^s, \\ \lambda &= \lambda^a+\lambda^s=R^{a^T}\lambda^a+R^{s^T}\lambda^s \end{aligned}$$

are obviously valid. If λ_0^a and λ_0^s are known it is easily possible to calculate the flow intensities λ_i, λ_i^a and λ_i^s needed for the analysis from these equation systems by means of the transposed routing matrices of the transition probabilities. In case the single diagnostic service units and the total diagnostic system are in equilibrium, i.e.

$$\lambda_i^a + \lambda_i^s < \mu_i \text{ and } \lambda_i^a + \lambda_i^s < n_{p_i}\mu_i, \text{ respectively,}$$

holds for every service unit with only one or n_{p_i} parallel diagnosis rooms, one obtains from the equations the queueing theoretical indices concerning the single diagnostic service units and the total system as mentioned in Paragraph 2.1 and compiled in Table 1. A single service unit corresponds - as already explained - to an M/M/1/∞/FCFS/2- or an M/M/n_p/∞/FCFS/2-model. Furthermore, the total system M/M/N/∞/FCFS/2 is also to be differentiated whether the number of parallel, identical diagnosis rooms of every service unit is arbitrary (n_{p_i} arbitrary) or whether every service unit consists of only one diagnosis room ($n_{p_i}=1$ for all i=1,...,N). As for the derivation of the indices one may be directed to the references Little (1961), Ferschl (1964), Gross/Harris (1974) Lemoine (1977), and Hegemann (1985).

3. The diagnosis system of the General Hospital in Hagen City

3.1 Structure, operating parameters and indices

The diagnostic system of the General Hospital in Hagen City comprises five departments: function diagnostics, sonography, endoscopy, computer tomography, and X-ray diagnostics, which are divided into twelve service units with one diagnosis room each. Under queueing theoretical aspects it is an M/M/12/∞/FCFS/2-model whose structure referring to number and designation of the service units, their services, staff and apparatus is shown in Table 2. The relevant examination times in the service units are between 8:00 AM and 4:00 PM with the X-ray diagnostics being manned 24 hours a day by a medical-technical assistant at least, though.

The parameters of the arrival processes of the out-patients and in-patients at the twelve diagnostic service units - important for the process structure of the diagnosis system - as well as the distribution of examination times in the service units and the routing matrices of the transition probabilities have been investigated within a three-week-period on the basis of diagnosis processes of more than 1,700 out-patients and in-patients by means of a self-recording-method of the medical personnel working

Table 1: Indices of the single service units and the total diagnosis system

indices	single service unit M/M/1/∞/FCFS/2	single service unit M/M/n_p/∞/FCFS/2	total system M/M/N/∞/FCFS/2, $n_{p_i}=1\ \forall i$	total system M/M/N/∞/FCFS/2, n_{p_i} arbitrary
expected waiting time $E(V_{q_i})$ resp. $E(V_q)$	$\frac{\rho_i}{\mu_i(1-\rho_i)}$	$\frac{1}{\lambda_i^a+\lambda_i^s}\cdot\frac{\rho_i^{n_{p_i}+1}}{(n_{p_i}-1)!(n_{p_i}-\rho_i)^2}p_{o_i}$	$\frac{1}{\lambda_o^a+\lambda_o^s}\sum_{i=1}^{N}\frac{\rho_i^2}{1-\rho_i}$	$\frac{1}{\lambda_o^a+\lambda_o^s}\sum_{i=1}^{N}\frac{\rho_i^{n_{p_i}+1}}{(n_{p_i}-1)!(n_{p_i}-\rho_i)^2}p_{o_i}$
expected indwelling time $E(V_i)$ resp. $E(V)$	$\frac{1}{\mu_i(1-\rho_i)}$	$\frac{1}{\lambda_i^a+\lambda_i^s}\left(\rho_i+\frac{\rho_i^{n_{p_i}+1}}{(n_{p_i}-1)!(n_{p_i}-\rho_i)^2}\cdot p_{o_i}\right)$	$\frac{1}{\lambda_o^a+\lambda_o^s}\sum_{i=1}^{N}\frac{\rho_i}{1-\rho_i}$	$\frac{1}{\lambda_o^a+\lambda_o^s}\sum_{i=1}^{N}\left(\rho_i+\frac{\rho_i^{n_{p_i}+1}}{(n_{p_i}-1)!(n_{p_i}-\rho_i)^2}p_{o_i}\right)$
expected number of waiting patients $E(L_{q_i})$ resp. $E(L_q)$	$\frac{\rho_i^2}{1-\rho_i}$	$\frac{\rho_i^{n_{p_i}+1}}{(n_{p_i}-1)!(n_{p_i}-\rho_i)^2}p_{o_i}$	$\sum_{i=1}^{N}\frac{\rho_i^2}{1-\rho_i}$	$\sum_{i=1}^{N}\frac{\rho_i^{n_{p_i}+1}}{(n_{p_i}-1)!(n_{p_i}-\rho_i)^2}p_{o_i}$
expected number of patients in the system $E(L_i)$ resp. $E(L)$	$\frac{\rho_i}{1-\rho_i}$	$\rho_i+\frac{\rho_i^{n_{p_i}+1}}{(n_{p_i}-1)!(n_{p_i}-\rho_i)^2}p_{o_i}$	$\sum_{i=1}^{N}\frac{\rho_i}{1-\rho_i}$	$\sum_{i=1}^{N}\rho_i+\frac{\rho_i^{n_{p_i}+1}}{(n_{p_i}-1)!(n_{p_i}-\rho_i)^2}p_{o_i}$
mean load degree ξ_i resp. ξ	ρ_i	$\frac{\rho_i}{n_{p_i}}$	$\frac{1}{N}\sum_{i=1}^{N}\rho_i$	$\frac{1}{N}\sum_{i=1}^{N}\frac{\rho_i}{n_{p_i}}$

$$\rho_i = \frac{\lambda_i^a + \lambda_i^s}{\mu_i} \qquad p_{o_i} = \frac{1}{\sum_{n_i=0}^{n_{p_i}-1}\frac{\rho_i^{n_i}}{n_i!} + \frac{\rho_i^{n_{p_i}}}{(n_{p_i}-1)!(n_{p_i}-1)}}$$

Table 2: Structure of the diagnosis system

area	room	services	executive personnel	important equipment
function diagnostics	function diagnostics 2 (FCT 2)	long-term-ECG resting-ECG rhythm-ECG pacemaker check phonocardiogram apexcardiogram carotid pulse curve orthostasis-ECG oscillogram	1 or 2 medi-cal-technical assistants (and 1 doctor)	1 long-term ECG computer 1 ECG apparatus 1 ultrasonic doppler appa-ratus
	function diagnostics 3 (FCT 3)	exercise-ECG echocardiogram pulmonary function test dye-dilution	1 medical-technical assistant and 1 doctor	1 ECG apparatus 1 echocardiograph 1 spirometre 1 dye-dilution apparatus
sono-graphy	sonography (SONO)	sonographic exa-mination of the abdominal organs	1 doctor	1 sonographic equipment
endoscopy	endoscopy (ENDO)	gastroscopy colonoscopy ERCP esophageal dila-tion organ biopsy laparoscopy endoscopic sclero-sizing of varices rectoscopy proctoscopy	1 or 2 medi-cal-technical assistants and 1 or 2 doctors	1 operating table and different endoscopes
computerized tomography	computerized tomography (COMP)	abdominal CT spinal CT cerebral CT pancreatic CT thoratic CT kidney CT liver CT pelvic CT	2 medical-technical assistants (and 1 doc-tor)	1 computer tomograph
X-ray diagnostics	X-ray diagnostics 1 (X 1)	fluoroscopy and roentgenogram of all parts of the body upper gastro-intestinal series barium enema phlebography esophagus ERCP fistulography	1 medical-technical assistant and 1 or 2 doctors	1 Bucky table

Table 2 (continued)

X-ray diagnostics	X-ray diagnostics 2 (X 2)	arthrography myelography (otherwise as in X 1 but without phlebo- graphy)	1 medical-technical assistant and 1 doctor	1 Bucky table
	X-ray diagnostics 3 (X 3)	chest X-ray (2 levels) trachea cervical spine paranasal sinuses	2 medical-technical assistants	1 Bucky wall tripod
		tomography of all regions of the body drip infusion cholecystography drip infusion pyelography		1 Bucky table
		panoramic picture of the upper and lower jaw bones		1 status-X-apparatus
	X-ray diagnostics 4 (X 4)	renovasography arteriography of of arms and legs lymphangiography aortography	1 medical-technical assistant and 2 doctors	1 AOT changer
	X-ray diagnostics 5.1 (X 5.1)	bones cholecysto- graphy after intra- venous injection thorax (inclined)	1 medical-technical assistant	1 Bucky table
		chest X-ray (2 levels) paranasal sinuses abdomen trachea spine (up-right- position)		1 Bucky wall tripod
	X-ray diagnostics 5.2 (X 5.2)	bones pyelography by intravenous injection function of arti- culations (knee or ankle) *Scheuba*	1 medical-technical assistant	1 Bucky table
	X-ray diagnostics 6 (X 6)	mammography pneumocystography galactography roentgenogram of the soft tissues of the hands	1 medical-technical assistant	1 mammomat

in the diagnostic system. Table 3 contains the intensities of the arrival flows and the examination rates. Hereby the outside of the diagnostic system is called 0. Then the routing matrices R^a and R^s follow. From these quantities one can calculate the indices of the single service units and the total diagnosis system according to the relations shown in Table 1; they are compiled in Table 4. Among these indices the expected waiting times of the patients $E(V_{q_i})$ and $E(V_q)$ as well as the load degrees ξ_i and ξ are of special interest for the further considerations.

3.2 Weak point analysis of the diagnosis system

Looking at the indices listed up in Table 4 the great differences in the load degree ξ_i of the single diagnostic rooms come up as the first weak point of the diagnosis system of the General Hospital in Hagen. Whereas the load degree of the computerized tomography reaches about 70% and is therefore rather high, the load degrees of the X-ray rooms 2, 5.2 and 6 reaching 13.7%, 25.5% and 6.2% as well as of the sonography reaching 13.6% must be called very low. Furthermore, the considerable differences between the X-ray rooms 1 and 2 equipped nearly identically are especially remarkable.

High divergencies can also be stated for the patients' waiting times $E(V_{q_i})$. A patient in front of the computerized tomography and X-ray room 4 must wait for his examination up to one hour on average, whereas the examination starts mostly immediately in the sonography, in the X-ray rooms 3 and 6 and in the function diagnosis room 2.

The analysis of the indwelling times $E(V_i)$ shows extremely high values for the computerized tomography and the X-ray rooms 1, 2 and 4. In the X-ray rooms the long indwelling times must be attributed to the relatively long examination times, whereas the long indwelling time of nearly 1 1/2 hours in the computerized tomography can be explained by the large part of waiting times.

Comparing the waiting times with the load degrees one meets the well-known dilemma of scheduling in the computerized tomography: the relatively high waiting times definitely result from the rather high load degree of nearly 70%. In the X-ray rooms 5.1 and 3 with comparatively higher load degrees only short waiting times appear because here examinations rarely last over a longer period. In the opposite sense, the dilemma of scheduling comes up in the sonography and in the X-ray room 6, too: the negligible short waiting times here solely result from the low load degrees and thus high dead times.

Concerning the diagnostic system as a whole the average waiting and indwelling times of the patients $E(V_q)$ and $E(V)$ can be considered as reasonable. The expected load degree ξ and thus the efficiency of the existing diagnosis system reaching 33.3% must be regarded as very low. Frequently, it is the aim of restructuring and rationalizing measures to strive for an average load degree of about 50%.

Table 3: Arrival and examination rates of the diagnostics

room	i	λ_i^a	λ_i^s	λ_i	μ_i
Outside of diagnostics	0	0,089444	0,153750	0,243194	-
FCT2	1	0,017083	0,065694	0,082777	0,193742
FCT3	2	0,007639	0,008889	0,016528	0,058173
SONO	3	0,001806	0,013611	0,015417	0,113483
ENDO	4	0,005357	0,010863	0,016220	0,042030
COMP	5	0,015667	0,010111	0,025778	0,037263
X 1	6	0,003889	0,013889	0,017778	0,037890
X 2	7	0,001667	0,001111	0,002778	0,020243
X 3	8	0,019722	0,050417	0,070139	0,207512
X 4	9	0,002222	0,000278	0,002500	0,008520
X 5.1	10	0,025417	0,032917	0,058334	0,112965
X 5.2	11	0,004861	0,005556	0,010417	0,040794
X 6	12	0,003472	0,000694	0,004166	0,066869

$$R^a=(r_{ij}^a)=\begin{bmatrix}
0 & 0{,}185 & 0{,}005 & 0{,}014 & 0{,}060 & 0{,}172 & 0{,}033 & 0{,}019 & 0{,}160 & 0{,}025 & 0{,}254 & 0{,}054 & 0{,}019 \\
0{,}472 & 0 & 0{,}415 & 0 & 0 & 0 & 0 & 0 & 0{,}106 & 0 & 0{,}008 & 0 & 0 \\
0{,}764 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0{,}200 & 0 & 0{,}036 & 0 & 0 \\
1{,}000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0{,}949 & 0 & 0 & 0{,}051 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0{,}947 & 0{,}009 & 0 & 0 & 0 & 0 & 0 & 0 & 0{,}044 & 0 & 0 & 0 & 0 \\
1{,}000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0{,}917 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0{,}083 & 0 & 0 & 0 & 0 \\
0{,}775 & 0{,}021 & 0{,}007 & 0{,}007 & 0 & 0{,}014 & 0{,}021 & 0 & 0 & 0 & 0{,}063 & 0 & 0{,}032 \\
1{,}000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0{,}951 & 0 & 0 & 0{,}005 & 0 & 0 & 0{,}022 & 0 & 0{,}022 & 0 & 0 & 0 & 0 \\
0{,}914 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0{,}086 & 0 & 0 & 0 & 0 \\
0{,}640 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0{,}080 & 0 & 0{,}280 & 0 & 0
\end{bmatrix}$$

$$R^S=(r^S_{ij})=\begin{bmatrix}
0 & 0{,}232 & 0{,}033 & 0{,}065 & 0{,}063 & 0{,}050 & 0{,}060 & 0{,}005 & 0{,}261 & 0{,}001 & 0{,}196 & 0{,}032 & 0{,}002 \\
0{,}841 & 0 & 0{,}049 & 0{,}011 & 0{,}002 & 0{,}013 & 0{,}002 & 0{,}002 & 0{,}061 & 0 & 0{,}015 & 0{,}004 & 0 \\
0{,}953 & 0 & 0 & 0 & 0 & 0 & 0{,}016 & 0 & 0{,}031 & 0 & 0 & 0 & 0 \\
0{,}735 & 0{,}133 & 0 & 0 & 0{,}031 & 0{,}010 & 0{,}031 & 0 & 0{,}031 & 0 & 0{,}020 & 0{,}010 & 0 \\
0{,}858 & 0{,}077 & 0{,}013 & 0 & 0 & 0{,}013 & 0{,}013 & 0 & 0{,}013 & 0 & 0{,}013 & 0 & 0 \\
0{,}904 & 0{,}055 & 0{,}027 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0{,}014 & 0 & 0 \\
0{,}810 & 0{,}070 & 0 & 0{,}010 & 0 & 0 & 0 & 0{,}010 & 0{,}090 & 0 & 0{,}010 & 0 & 0 \\
1{,}000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0{,}452 & 0{,}419 & 0 & 0{,}052 & 0{,}008 & 0{,}011 & 0{,}022 & 0 & 0 & 0{,}003 & 0{,}019 & 0{,}006 & 0{,}008 \\
1{,}000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0{,}654 & 0{,}131 & 0{,}004 & 0{,}004 & 0 & 0{,}021 & 0{,}076 & 0{,}004 & 0{,}105 & 0 & 0 & 0 & 0 \\
0{,}750 & 0{,}050 & 0 & 0 & 0{,}025 & 0 & 0{,}025 & 0 & 0{,}125 & 0 & 0{,}025 & 0 & 0 \\
0{,}600 & 0{,}200 & 0 & 0 & 0 & 0{,}200 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

Table 4: Indices of the single diagnostic service units and the total system

room	i	$E(V_{q_i})$ [min.]	$E(V_i)$ [min.]	$E(L_{q_i})$ [pat./min.]	$E(L_i)$ [pat./min.]	ξ_i [%]
FCT2	1	3,85	9,01	0,32	0,75	42,7
FCT3	2	6,82	24,01	0,11	0,40	28,4
SONO	3	1,38	10,19	0,02	0,16	13,6
ENDO	4	14,95	38,74	0,24	0,63	38,6
COMP	5	60,23	87,06	1,55	2,24	69,2
X 1	6	23,32	49,72	0,41	0,88	46,9
X 2	7	7,85	57,25	0,02	0,16	13,7
X 3	8	2,46	7,27	0,17	0,51	33,8
X 4	9	48,73	166,11	0,12	0,42	29,3
X 5.1	10	9,45	18,30	0,55	1,07	51,6
X 5.2	11	8,40	32,91	0.09	0,34	25,5
X 6	12	0,99	15,94	0,01	0,07	6,2
total system		$E(V_q)$ 14,9	$E(V)$ 31,3	$E(L_q)$ 3,6	$E(L)$ 7,6	ξ 33,3

Literature:

Bailey, N.T.J. (1954): Queueing for Medical Care, in: Journal of the Royal Statistical Society 3, serie C, p. 137-145.

Burke, P.J. (1956): The Output of a Queueing System, in: Operations Research 4, p. 699-704.

Burke, P.J. (1972): Output Processes and Tandem Queues, in: Proceedings of the Symposium on Computer-Communications Networks and Teletraffic, New York, p. 419-428.

Carruthers, M.E. (1970): Computer Analysis of Routine Pathology Work Schedules Using a Simulation Programme, in: Journal of Clinical Pathology 23, p. 269-272.

Eichhorn, S. (1976): Krankenhausbetriebslehre, vol. II, 3rd edition, Köln.

Ferschl, F. (1964): Zufallsabhängige Wirtschaftsprozesse, Wien/Würzburg.

Fleck, H. (1977): Analyse der Ablauforganisation in einem medizinischen Untersuchungszentrum mit Hilfe der Netzwerktechnik, in: Methods of Information in Medicine 16, p. 81-88.

Gierl, L. (1976): Partialmodelle der Ablaufplanung im Krankenhaus auf empirischer Grundlage, Ph.-D. thesis, Nürnberg.

Gross, D. and Harris, C.M. (1974): Fundamentals of Queueing Theory, New York/London/Sydney/Toronto.

Hegemann, H. (1985): Zur Kapazitäts- und Prozeßplanung in der klinischen Diagnostik, Berlin/Heidelberg/New York/Tokyo.

Jackson, J.R. (1963): Jobshop-like Queueing Systems, in: Management Science 10, p. 131-142.

Kendall, D.G. (1953): Stochastic Processes Occuring in the Theory of Queues and their Analysis by the Method of the Imbedded Markov Chain, in: The Annals of Mathematical Statistics 24, p. 338-354.

Klimow, G.P. (1979): Bedienungsprozesse, Basel/Stuttgart.

Lemoine, A.J. (1977): Networks of Queues - A Survey of Equilibrium Analysis, in: Management Science 24, p. 464-481.

Little, J.D.C. (1961): A Proof of the Queueing Formula $L = \lambda W$, Operations Research 9, p. 383-387.

Melamed, B. (1979a): On Poisson Traffic Processes in DiscreteState Markovian Systems with Applications to Queueing Theory, in: Advances in Applied Probability 11, p. 218-239.

Melamed, B. (1979b): Characterizations of Poisson Traffic Streams in Jackson Queueing Networks, in: Advances in Applied Probability 11, p. 422-438.

Neumann, K. and Steinhardt, U. (1979): GERT Networks and the Time-Oriented Evaluation of Projects, in: Lecture Notes in Economics and Mathematical Systems, vol. 172, Berlin/Heidelberg/New York.

O'Kane, P.C. (1981): A Simulation Model of a Diagnostic Radiology Department, in: European Journal of Operational Research 6, p. 38-45.

Pritsker, A.A.B. and Happ, W.W. (1966): GERT: Graphical Evaluation and Review Technique, in: Journal of Industrial Engineering 17, p. 267-274 and p. 293-301.

Revesz, G., Shea, F.J. and Ziskin, M.C. (1972): Patient Flow and Utilization of Resources in a Diagnostic Radiology Department, in: Radiology 102, p. 21-30.

Taylor, B.W. and Keown, A.J. (1980): A Network Analysis of an Inpatient-Outpatient Department, in: Journal of the Operational Research Society 31, p. 169-179.

Taylor, I.D.S. and Templeton, J.G.C. (1980): Waiting Time in a Multi-Server Cutoff-Priority Queue, and its Application to an Urban Ambulance Service, in: Operations Research 28, p. 1168-1188.

Walrand, J. (1982): On the Equivalence of Flows in Networks of Queues, in: Journal of Applied Probability 19, p. 195-203.

DECISION SUPPORT MODEL FOR PEST MANAGEMENT: A PRELIMINARY REPORT

GERALD L. JUBB
UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND

SUBHASH C. NARULA
VIRGINIA COMMONWEALTH UNIVERSITY
RICHMOND, VIRGINIA

JOHN F. WELLINGTON
GANNON UNIVERSITY
ERIE, PENNSYLVANIA

1. INTRODUCTION

The grape industry of western New York and northwestern Pennsylvania states is an established and a big agribusiness for the region. New York is the nation's second leading producer of grapes which are its second most important crop. It is the most important crop in terms of its overall impact on the economy of the state, Pearson, Seem and Eisensmith (1985). The importance of the grape industry to the region is clear from the presence of permanent agricultural experiment stations in New York and Pennsylvania. The mission of the experimental stations is to provide information and consultation to the grape growers.

According to the New York Orchard and Vineyard Survey of 1980, Suter and Bascom (1981), there are more than 800 grape farms in New York state with nearly 4200 acres under cultivation. The average yield is about five tons per acre. The grapes produced in the region are used for (i) processing juices, jams, or jellies (74%); (ii) fermenting into wines (24%); (iii) dessert and table grapes (2%), Pearson et al. (1985).

The grape growers deal with a number of natural elements that may either destroy or harm a crop; for example, weather, fungi, weeds, insects, mites, and birds. It is important to control or limit the damage due to these causes. The problem is crucial for grape growers since they deal in low profit margin crop in a highly competitive environment. It is expected that the trend will continue. We will address the issue of pest management in this paper.

According to Pearson et al. (1985) more than 200 insects and mite species have been associated with the grapes in this region. Of these 99 feed on grapes and may become pests under proper conditions. However, 15 species (see Table 1 in the Appendix) are common in the vineyards of this region. They state (pp. 12-13):

"Yield losses are likely from 7 to 13 insect pests in the first year and from 1 to 13 pests in the second year. Yield losses in the first year ranged from 0.1% for the grape cane gallmaker to 20% for the rose chaffer.

Other notable pests causing yield loss were climbing cut worms (6.9% yield loss), grape flea beetle (5.2%) and grape berry moth (2.1%). Second year yield loss from the grape borer would be significant and could reach 100% by the fourth year of infestation."

They further state that losses may also arise from the insects as contaminants present at harvest. The high density of such pests at harvest may result in down grading the value of the grapes. In fact, if the infestation is severe enough, the grapes may be rejected by the processors resulting in 100 percent loss. In such cases grapes destined for dessert and table market are also very likely to be a total loss.

Thus the use of chemicals (insecticides and miticides) plays an important role in insuring the economic value of the grape crop. In addition, the chemicals of today are relatively easy to use and are highly effective. Pearson et al. (1985) note that the use of chemicals is more economical than the use of other controls such as cultural control particularly for this region with old established vineyards and cultural practices.

In an effort to find the most "efficient" pest control program, the grape growers in the region could obtain information from a variety of sources, for example, Travis, Jubb, Tetrault, Haeseler, and Obourn (1983), Riedl, Burr, and Tomkins (1983), and Funt, Ellis, Williams and Hall (1983). However until recently the effectiveness of a chemical against a pest or its toxicity to a beneficial arthropod was not known. A recent study referred to herein as the Geneva Project (reported in Pearson et al. 1985) reported the level of effectiveness (low, medium, and high) for fifteen chemicals (listed in Table 2 in the Appendix) against the pests listed in Table 1. They also reported the toxicity (as unknown, relatively nontoxic, low, medium and high) of these chemicals to the beneficial arthropods (listed in Table 3 in the Appendix). From their study, it became clear that some of the chemicals may be used to control more than one pest where another may be effective only against a particular pest. Further, the effectiveness of a chemical depends upon the pest being controlled. The amount of pesticide used and thus its cost is also pest dependent. Another consideration in the selection of a chemical is its effect on the beneficial arthropod. In order to develop a mathematical model for the selection of chemicals to control a number of pests it is important to learn about the interaction among the chemicals as well as the role of arthropods in the control of pests. However, this information is not available at present.

Our objective in this paper is to report a model in which it is assumed that there are no interactions among the chemicals and their effect is linear and additive. The rest of the paper is organized as follows. In Section 2, we give the problem formulation. In Section 3 we present an algorithm to solve the problem. We conclude the paper with a few remarks in Section 4.

2. PROBLEM FORMULATION

To formulate the grape grower's chemical selection problem, we define the following parameters and decision variables:

Parameters

C_{ijk} = cost per acre of using chemical i rated at effectiveness level j for treating pest k.

$$e_{ijk}^{bt} = \begin{cases} 1, & \text{if chemical } i \text{ rated at effectiveness level } j \text{ for treating pest } k \text{ has toxicity level } t \text{ toward a beneficial arthropod } b, \\ 0, & \text{otherwise} \end{cases}$$

n_j = minimum number of pests to be treated at the effectiveness level j

n_{bt} = number of times a beneficial arthropod b is exposed to toxicity level t

D = desired upper bound on the cost.

Decision Variables

$$X_{ijk} = \begin{cases} 1, & \text{if chemical } i \text{ rated at effectiveness level } j \text{ is used for treating pest } k \\ 0, & \text{otherwise.} \end{cases}$$

The model for the pesticide selection problem (PSP) can be stated as:

$$\text{Minimize} \sum_i \sum_j \sum_k C_{ijk} X_{ijk} \qquad \text{(PSP)}$$

$$\sum_i \sum_j X_{ijk} \geq 1 \quad , \forall k \qquad (1)$$

$$\sum_i \sum_k X_{ijk} \geq n_j \quad , \forall j \qquad (2)$$

$$\sum_i \sum_j \sum_k e_{ijk}^{bt} X_{ijk} \leq n_{bt} \quad , \forall b,t \qquad (3)$$

$$\sum_i \sum_j \sum_k C_{ijk} X_{ijk} \leq D \qquad (4)$$

$$X_{ijk} = 0 \text{ or } 1 \quad \forall\, i,j,k \qquad (5)$$

In (PSP), the objective function minimizes the total cost of the chemicals per acre; constraints (1) assure that each pest is treated by at least one chemical; constraints (2) guarantee that at least n_j pests will be treated at effectiveness level j; constraints (3) assure that a beneficial b will be exposed to toxicity level t no more than n_{bt} times; constraint (4) controls the total cost to be less than or equal to a specified level; and constraints (5) state that the decision variables can take on a value equal to 0 or 1.

It may be observed that (PSP) is a zero-one integer programming problem and reduces to a standard set covering problem if constraints (2), (3), and (4) are ignored.

3. SOLUTION PROCEDURE

We use branch and bound procedure to solve the PSP problem. We represent the problem as a tree with p+1 levels, where p denotes the number of pests. Level 0 denotes the root of the tree and level k corresponds to pest k, k=1,...,p. Each level consists of nodes which are connected to nodes at the preceding and the succeeding levels by arcs. A node at level k of the tree represents a decision variable for treating a pest k. Each node is identified by an integer which corresponds to the pair (i,k), i.e., the chemical i used to treat pest k. A branch consisting of a series of p+1 connected nodes, one at each level, insures that each pest is treated by a chemical and thus represents a solution to the problem. All possible solutions of the problem are represented by the branches of the tree. A branch starting at level 0 and made up of fewer than p+1 nodes represents a <u>partial solution</u> to the problem.

The total cost of a branch, starting at level 0 and consisting of a series of p+1 connected nodes, one at each level, is equal to the sum of the C_{ijk}'s along the branch: It may be noted that this sum is non decreasing as we move from level k to k+1 in the tree. Thus, it makes it unnecessary to complete the delineation of a branch for which the total cost of the partial solution exceeds the total cost of some branch delineated previously. Such a branch is said to be <u>fathomed</u>. The ability to fathom in general makes it unnecessary to examine every branch of the tree explicitly. A branch can also be fathomed if a partial solution does not satisfy an effectiveness or a toxicity constraint.

A step by step procedure to find an optimum solution can be stated as follows:

<u>Step 0</u>: Construct the list.

Let p denote the number of pests and n the number of available pesticides. Identify the number of pesticides m_j available to treat pest j (=1,...,p). Let $-i_1,\ldots,i_{m_j}$ denote the indices of the decision variables arranged in order of non decreasing cost ($C_{i_1} \leq \ldots \leq C_{i_{m_j}}$). The sequence of ordered indices is referred to as field j. To separate various fields, the left most index of each field is preceded by a minus sign. Label the left most index of field 1.

Set $Z^* = \infty$ and $r = 1$. Go to Step 1.

<u>Step 1</u>: Let $J = \{j_1,\ldots,j_r\}$ be the set of indices labelled in the current list for fields 1,...,r, where j_i is the index currently labelled in field i, i=1,...,r.

For j = 1,...,n, set

$$X_j = \begin{cases} 1, & \text{for } j \in J, \\ 0, & \text{otherwise.} \end{cases}$$

Set $Z = \sum_{j=1}^{n} C_j X_j$ and $\mu = \sum_{j=1}^{n} X_j$. Go to Step 2.

<u>Step 2</u>: Perform the fathoming test.

If $Z > Z^*$, fathom the current solution, go to Step 5 (i.e., take a backward step); otherwise, go to Step 3.

Step 3: Check the feasibility of the current solution.

If the current solution is feasible and if $Z < Z^*$, set $Z^* = Z$ and save the current solution as the best solution found so far. If the rightmost index in field r is labelled, go to Step 5 (i.e., take a backward step). If further continuation of the current solution will not meet the feasibility requirements of the problem or if $\mu = p$, set $s = 0$; otherwise set $s = 1$. Go to Step 4 (i.e., take a forward step).

Step 4: Take a forward step.

Set $r=r+s$. If field r is currently labelled, advance the label in field r one position to the right. If field r is currently unlabelled, label the leftmost index. Go to Step 1.

Step 5 Take a backward step.

If $r=1$, examine field 1; if $r>1$, decrease r by 1 and continue to do so until the first field r is found whose label may be advanced one position to the right and remain within that field. If none can be found, stop; the procedure is completed. Otherwise, drop the labels in fields $r+1,\ldots,p$ and go to Step 1.

4. SOME REMARKS

The model proposed here can be used for selection of chemicals to control pests if the simplifying assumptions of the model hold. However, under these conditions, the usefulness of the model lies really in the fact that it allows us to study the effect on the solution as we change the parameters of the model. We can study the effect on the solution to change in the level of effectiveness required against a pest, or the change in the level of toxicity desired towards a beneficial arthropod or limit the number of chemicals used, etc. These changes can be studied easily. We have developed a FORTRAN computer program for the mainframe computers and a program in BASIC for the microcomputers.

As mentioned in the introduction, the model is based on some very simplifying assumptions. Clearly, the problem is a multicriteria optimization problem. However, before a more sophisticated decision support system can be developed, we need further information from the agricultural experiment stations.

REFERENCES

Funt, R. D., Ellis, M. A., Williams, R. N., and Hall, F. R. (1983). Ohio Commercial Fruit Spray Guide. Ohio State University Cooperative Extension Service Bulletin 506B.

Pearson, R. C., Seem, R. C., and Eisensmith, S. P. (1985). Assessment of Pests, Losses They Cause and Pest Management Strategies for Great Lakes/Eastern Grapes. Special Report, The New York State Agricultural Experiment Station, Geneva, Division of the New York State College of Agriculture and Life Sciences, Cornell University, Ithaca, New York.

Riedl, H., Burr, T. J., and Tomkins, J. P. (1983). Grape Pest Control Guide. Cornell University Cooperative Extension Service, Ithaca, New York.

Suter, G. W. and Bascom, P. F. (1981). New York Orchard and Vineyard Survey. New York State Crop Reporting Service, Albany, New York.

Travis, J. W., Jubb, G. L., Tetrault, R. C., Haeseler, C. W., and Obourn, T. H. (1983). Grape Disease, Insect, and Weed Control Suggestions. Cooperative Extension, The Pennsylvania State University, University Park, Pennsylvania.

APPENDIX

TABLE 1 Fifteen species of possible pests in vineyards of northwestern Pennsylvania and western New York

Insect/Mite Name
Climbing cutworms
Drosophila (fruit flies)
European red mite
Eightspotted forester
Grape berry moth
Grape blossom midge
Grape cane gallmaker
Grape cane girdler
Grape flea beetle
Grape leafhopper
Grape phylloxera (foliar)
Grapevine tomato gall
Japanese beetle
Redbranded leafroller
Rose chaffer

TABLE 2 Fifteen chemicals used to control pests in the vineyards of northwestern Pennsylvania and western New York

Chemical Name
Azinphosmethyl
Carbaryl
Demeton
Diazinon
Dicofol
Encapsulated methyl parathion
Endosulfan
Hexakis
Malathion
Methomyl
Methoxychlor
Naled
Parathion
Phosmet
Phosolone

TABLE 3 Some beneficial arthropods in the vineyards of northwestern Pennsylvania and western New York

Beneficial Name
Apis mellifera
Chrysopa
Geocoris
Nabis
Neoseiulus fallacis
Orius

DECISION SUPPORT FOR DEVELOPMENT STRATEGY IN THE AREA OF INTERACTION OF THE INDUSTRIAL AND AGRICULTURAL SECTORS

T.Ryś, M.Skocz, M.Zebrowski, W.Ziembla
System Research Department of the Institute for Control and Systems Engineering, Academy of Mining and Metallurgy, Cracow, and of the Industrial Chemistry Research Institute, Warsaw, Poland

1. INTRODUCTION

Design of the Industrial Development Strategy /IDS/ is a natural field for development and application of Decision Support System /DSS/.

The key factor in IDS design is to embed a phenomenon of interaction between sectors which are respectively a supplier and a consumer. Here we go far beyond a simple demand-oriented approach.

The subject of this paper covers a DSS, and more specifically its structure and basic procedure, proposed as an aid for design of IDS. The case considered here concerns the pesticide industry. The interacting sector is agriculture.

The research was initiated by two related but originally independent projects. First was aimed at decision support for development of the pesticide industry /sponsored by the Institute for Organic Chemistry, Warsaw, Poland/. This project covers a so-called Production-Distribution Area /PDA/ [1, 2], delimited by the domestic pesticide industry. The second project /sponsored by the Institute for Plant Protection, Poznan, Poland/ is aimed at analysis and design of plant protection programs for the Polish agriculture. The experience gained in the course of these projects led inevitably to the research reported here. Owing to this new approach, the Decision Maker /DM/ responsible for the IDS of the chemical sector, is able to consider a demand expressed in terms of a goal of the consumer /i.e. effects expected by agriculture/. This also includes influence of competitors on the pesticide market.

2. OBJECTIVES AND SCOPE

2.1. Decision Maker Objectives

The goal of DM is to transform his holistic view on how his industry should develop into the sound IDS. This means that his experience, intuition and creativity with a help of relevant information should enable him to arrive at the state of knowledge which can be cathegorized as follows:

- what market volume is expected to be achieved?
- what array of products, existing and new technologies is to be selected in order to form the industrial structure being able to meet the challenge coming from the assumed market expansion?
- what would be necessary investment and other economic consequences of assumed IDS in terms of intensity and volume of operation?
- what would be the position of the industry among competitors?

An answer to the above questions leads to formulation of IDS representing a concept where all relevant assumptions are to be consistent. The assistance in creating of such an IDS should be a natural role of DSS.

2.2. System Description

Let us look into the system from the point of view of DM's objectives. The system consists of production and consumption sectors i.e. of the pesticide industry and agriculture. The production sector includes the domestic industry and pesticide imports that represent other producers. The sectors of the system, as well as intersector flows are illustrated in Fig.1.

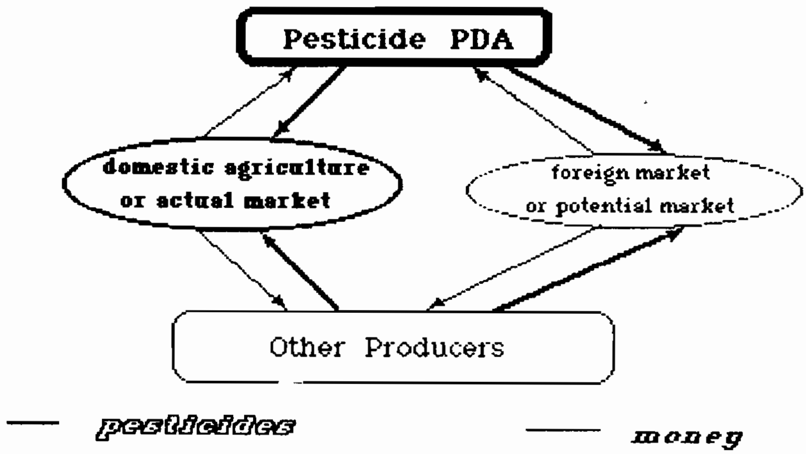


Fig 1. The Battle Field

In a natural way the interaction between the sectors goes on through the market where the goods /pesticides/ and money are exchanged. Therefore, for our purpose, there is no need to consider other variables representing this interaction. From the point of view of agriculture, money flow to industry represents expenditure which is supported to bring benefit coming from plant protection. On the other hand, the same money flow as received by industry represents resource that can be acquired by the industry only provided that it is efficient above a certain critical level. This means that both partners have to be satisfied, and of course this satisfaction can be achieved by agriculture simply by switching to other producers.

It follows from the above that in the case of agriculture we consider only some of its activities related to application of pesticides. The problem of development strategy comes here from a strong dependence of crops upon appropriate plant protection. Thus, one of possible strategies would be this of maximizing of profit coming from crop increase as a function of protection program.

The key factor in this is a great variety of possible alternatives of protection programs according to a variety of substitutes among pesticides.

On the side of the chemical industry we also deal with a great variety of substitution due to alternative production processes which can run on any given installation, and consequently to alternative technological routes that lead to final products.

Although in principle the industry should supply its products according to demands of agriculture, technological interdependencies and differences in profit margins of both sectors, as well as their different technological progress make their development strategies seem to be a kind of compromise. Through the analysis of properties of their development, this compromise can be recognized and arising conclusions can help in design of appropriate IDS. A search for concordance between the assumed market expansion /our DM's goal/ and the plant protection efficiency /goal of agriculture/ is a common rule of development. This stems from the fact of competition between pesticide producers who eventually influence a state of the market. Despite its common sense, this rule of development is quite difficult to be put into real life.

The market conditions may be conveniently described by introduced Pesticide Application Characteristic /PAC/. By this characteristic we understand a relationship between plant protection cost and crop increase expressed in monetary units. Such characteristics can be constructed for any pesticide producer that exists on the market. Differences between the characteristics corresponding to the particular producer and other companies naturally determine a market share, i.e. allocation of money spent by agriculture for plant protection. From this follows that any desired change of the market share to the benefit of the industry represented by DM demands modification for improvement of its PAC. Moreover, by analysis of the relevant Pesticide Application Characteristics, DM who's problem is discussed may recognize the production ability and marginal economic effectiveness achieved by competitors[1]. The above mentioned parameters are to be expressed in terms of volume of operation /monetary/ and of intensity ratio /production effect vs production cost/. These two parameters form an Industry Effectiveness Characteristic /IEC/ which may be used as the basic determinant for IDS design.

---

[1] It might be very difficult, if not impossible to construct PAC for individual competitors. But it is sufficient to have an aggregate PAC representing all competitors.

Concluding, from the analysis of market conditions described by Pesticide Application Characteristic /PAC/, the DM can derive a new PAC /alternatively called market expansion function/ which leads to modification of the Industry Effectiveness Characteristic /IEC/ for development strategy.

Below, a theoretical framework for our consideration will be presented. The introduced models and proposed procedure to DM are aimed at design of DSS according to the objective and scope given above.

3. DSS STRUCTURE

3.1. Mathematical Models Involved

The structure displayed in Fig.2. reflects results of the identification presented in Section 2.2. In other words, Fig.2. shows a mapping of the system into the corresponding system of models. As can be seen, in our approach two basic models are used. PDA Model of the Chemical Industry [1, 2] and Plant Protection Model [3].

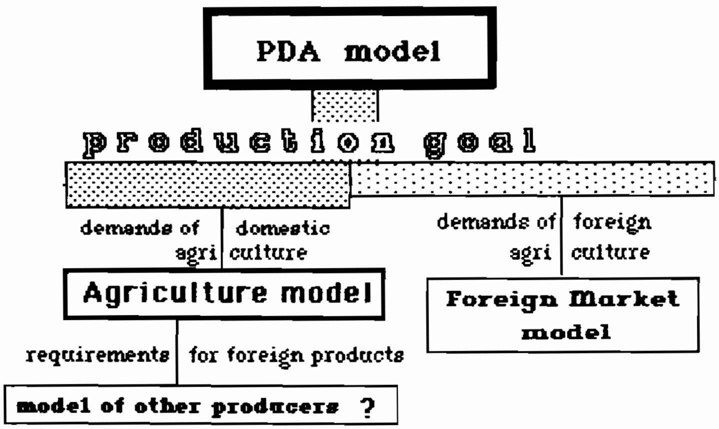


Fig. 2 The System Description

Though the core of the problem concerns the industry, the more natural way will be to start our description from the agriculture model.

Plant Protection Model. The model aims at evaluation of economic effects of pesticide application /pest management/ in a regional- or country-wide scale and does not incorporate a detailed description of pest dynamics. The value of increased yields to pesticide use is determined through consideration of the most important factors and relations such as:

a/ the potential crop yield resulting from the use of a

particular pesticide or a sequence of pesticides,
b/ the average crop production per ha /expressed in physical and monetary units/,
c/ the cost of pesticides and their application,
d/ parameters that describe possible methods of protection /doses etc./,
e/ the acreage of plants,
f/ the pesticide availability.

The relation /a/ and information /c/ are used to construct objective function/s/. Relation /b/, /d/, /e/, /f/ constitute constraints of the model. The model has an LP formulation where decision variables are acreage of plants protected and amounts of pesticides demanded to fulfil a protection program. For the sake of illustration, let us present the model in the following compact form:

$$\max_{y \geqslant 0} \langle d, y \rangle$$

subject to:

$$Cy \leqslant r_a \qquad (1)$$

$$F_y = e$$

where:

y - is a vector of pesticide demand,
d - is a vector of unit benefits and costs of pesticides applied,
r_a - is a global sum of money spent on a protection program,
e - is a vector of agricultural constraints /acreage of plants, availability of pesticides etc./,

Hence, our model represents the problem of crops increase subject to constraints on methods of protection and agricultural

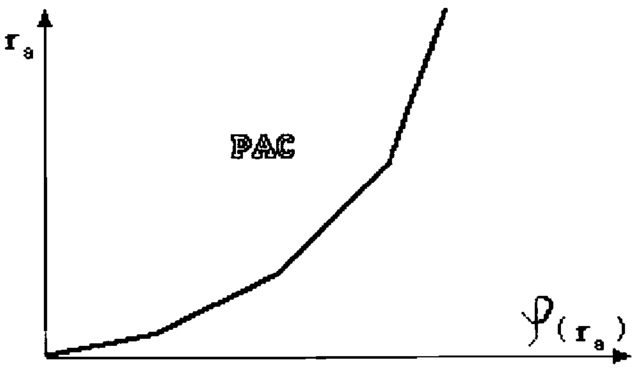


Fig.3 Pesticide Application Characteristic

production conditions. From the model one may derive the relationship of special interest: the Pesticide Application Characteristic /PAC/. Let us consider the following function:

$$\varphi(r_a) = \max_{y \in Y(r_a)} \langle d, y \rangle \tag{2}$$

where:

$$Y(r_a) = \left\{ y \, : \, Cy \leqslant r_a \, , \, Fy = e \, , \, y \geqslant 0 \right\}$$

For our convenience, the PAC is defined as the inverse function of /2/. Since /1/ is a LP problem, PAC will look as in Fig. 3.

<u>PDA Model of the Pesticide Industry</u>. The model represents the chemical plant/s/ level viewed as a network of production processes aggregated into simple production functions and distribution flows for a group of chemicals specified within a given Production Distribution Area. The version of the model describes all possible modes of production including alternative ranges of products made at a given installation, recycling of semiproducts and coupled production of a number of chemicals at each plant considered. This must therefore take into account:

- the processing and flows of chemicals,
- the flows of chemicals into and out of other areas and industries representing the marketing or business activity of the PDA,
- the flow of investment, revenue and utilities /energy, manpower, etc./.

The model based on the mass balance principle has LP representation that, for the sake of simplicity, we will write in the form:

$$\max_{x \geqslant 0} \langle c, x \rangle$$

subject to:

$$\begin{aligned} Ax &\leqslant r_i \\ Bx &= p \end{aligned} \tag{3}$$

where:

x - denotes production levels of chemicals produced or consumed within the PDA,

c - is a vector of sale or purchase prices of chemicals,

r_i - is a vector of resources allocated to the PDA /including capital utilized in production and investment/,

p - is a vector of pesticide production goal.

For our purposes it is also convenient to have an aggregated model which easily results from (1). This aggregated model will be used as the Industry Effectiveness Characteristic of the branch being considered. It is defined as follows:

$$f(r_i) = \max_{x \in X(r_i)} \langle c, x \rangle \tag{4}$$

where:

$X(r_i) = \{x : Ax \leqslant r_i,\ Bx = y, x \geqslant 0\}$

For the same reasons as in the case of $\varphi(r_a)$, $f(r_i)$ looks as in Fig.4.

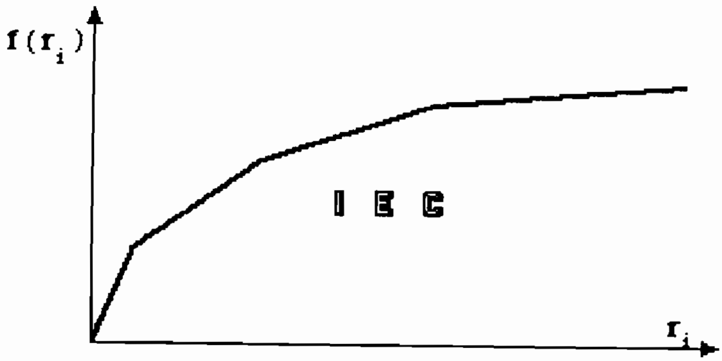


Fig.4 Industry Effectiveness Characteristic

As r_i may be a multidimensional vector, $f(r_i)$ is generally a hypersurface over r_i .

The interconnection between both vectors may be described by the equation:

$$Gy \leqslant p \qquad (5)$$

3.2. DSS Procedure

DSS enables DM to go through a procedure which consists of some basic steps:

Step 1. Market investigation,
Step 2. Choice of Market Expansion Goal /MEG/,
Step 3. Matching MEG with the potentially available technological structure of the industry,
Step 4. Evaluation of IDS that results from the procedure,
Step 5. Interactive modification of basic assumptions imposed on IDS design in steps 2 and 3 or choice of IDS.

Before describing DSS procedure let us make some important remarks. Although steps 1 - 5 suggest an algorithmic type of the procedure, it must be underlined that its idea is of the DSS type. This means that DM is free to use the steps in a flexible interactive way and he may supply any information necessary for given steps from various sources including his imagination. Let us describe now steps 1 to 5. In the phase of market investigation DM has to estimate the present state of the market. The simplest aggregate way of doing it is to use a set of PACs exemplified in Fig.5.

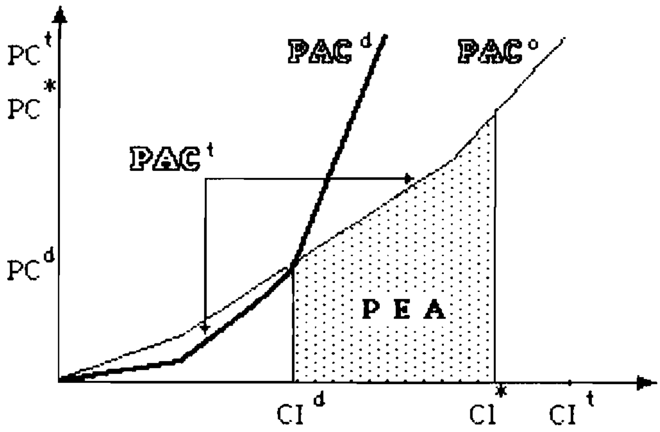


Fig.5 The state of the market

where:

PAC^d - corresponds to pesticide supply from DM's industry,

PAC^o - corresponds to pesticide supply from other sources,

PAC^j - joint characteristic corresponding to the whole market,

PC^d - level of DM's industry sale,

PC^t - level of the total market sale.

It follows from properties of the market that the joint characteristic i.e. PAC^j represents the most efficient plant protection which agriculture can get from the pesticide offer.

Assuming that the state of the market was recognized adequately, the next action of DM would be to set a goal for market expansion. In Fig.5. an assumed Potential Expansion Area /PEA/ was drawn according to MEC expressed by a new level of protection cost PC* and a corresponding crop increase CI*. Of course to achieve such a goal a new PAC^d must be positioned below the PAC^j.

Having perhaps an ambitious MEG, DM should match it with technological reality of his industry. This may be achieved by utilizing a natural flexibility of plant protection /substitution of pesticides/ and flexibility of technological structures /substitution of technologies/. Mathematically, using models (1) and (3) we can formulate the following problem for IDS design:

$$\max_{x,p \geqslant 0} \langle c, x \rangle$$

subject to:

$$\begin{aligned} & Ax \leqslant r \\ & Bx = p \\ & Gy \leqslant p \\ & PAC^d(PC) \subset PEA \quad \text{for} \quad PC \leqslant PC^* \end{aligned} \tag{6}$$

As follows from the interpretation of models (1) and (3), as a solution to problem (6) we obtain a hypothetical structure of the industry /represented by x / and an array of pesticides that are to be sold to agriculture.

Now, time comes to evaluate the newly emerging structure in the terms of its efficiency. Again, a function of DSS is to compute the new Industry Economic Characteristic /IEC/. From the IEC one may draw various conclusions. DM can be either satisfied and take the IDS into further processing towards implementation, or otherwise again with help of DSS he may interactively modify basic assumptions imposed on IDS design at step 2 and/or step 3.

This procedure has been tested on the PDA which represents the domestic pesticide production. Domestic supplies constitute only a fraction of agricultural consumption - the remainder is imported. Relevant models of the industry and agriculture were used in the procedure. This enabled gaining positive enough experience to claim practical value of the DSS. The exact description of the case study goes, however, beyond the scope of the paper and will be published separately.

4. IMPLEMENTATION

According to the experimental stage of the project, the DSS has been implemented by using the existing software packages developed during our studies on IDS design for the chemical industry and for the plant protection program in agriculture. These packages and procedures have been adapted for this case of DSS and their original version has been documented [4, 5].

5. CONCLUSIONS

Our experience confirms that practical use of DSS is possible provided that it is really oriented towards support as opposed to algorithmization of the decision making process. Further on, this experience will be used for implementation of a new version of DSS with emphasis on the interactive mode.

ACKNOWLEDGEMENTS

We are grateful to our collegues at the Institute for Plant Protection in Poznań, Poland and at the Chemical Industry Research Institute for support, discussions and information without which this work could not have progressed.

REFERENCES

1. Dobrowolski G., Kopytowski J., Lewandowski A., and Żebrowski M. /1982/ Generating Efficient Alternatives for Development in the Chemical Industry. Collaborative Paper, CP-82-54, International Institute for Applied Systems Analysis, Laxenburg, Austria.
2. Górecki H., Dobrowolski G., Kopytowski J., and Żebrowski M. /1982/ The Quest for Concordance between Technologies and Resources as a Multiobjective Decision Process, Proceedings of IIASA Workshop on Multiobjective Decision Modelling.
3. Skocz M. /1984/ Optimization Method for Pesticide Application. Proceedings of the IX CIEC Congress, Budapest,Hungary.
4. Kreglewski T., and Lewandowski A. /1983/ MM-MINOS - An Integrated Interactive Decision Support System. Collaborative Paper, CP-83-63 International Institute for Applied Systems Analysis, Laxenburg, Austria.
5. Dobrowolski G., Hajduk K., Korytowski A., and Ryś T. /1984/ POSTAN - A Package for Postoptimal Analysis /An Extension of MINOS/. Collaborative Paper CP-84-32 International Institute for Applied System Analysis, Laxenburg, Austria.

ABSTRACT

The scope of this paper covers a problem of Industrial Development Strategy design /IDS/ for the pesticide industry which includes as a key factor interaction between the industry and agriculture. It follows from our research that the problem should not be solved by adoption of a simple demand-oriented strategy of development.

The purpose of our broader study was two-fold. First, we aimed at development of Decision Support System which could be used in design of IDS for the pesticide industry as well as for the analysis and design of plant protection programs. Second, it was to assist and take part in applying of DSS for the two mentioned sectors.

Here we focus on findings related to DSS for the Decision Maker /DM/ responsible for development of the industry. Accordingly, DM objectives and results of identification of the problem area are presented. From this a structure and functions of DSS are envolved. Relevant models of both sectors are briefly shown. Then, an essence of IDS procedure is described.

The approach used for the currently implemented version of DSS is also mentioned. Up-to-date experience gained so far is summarized.

List of Participants

Workshop on Large-Scale Modelling and Interactive Decisions

Wartburg Castle, Eisenach, G.D.R.

November 17-22, 1985

Mrs. A.M. Barnikov
Academy of Sciences of the GDR
Institute for Chemical Technology
Rudower Chaussee 5
DDR-1199 Berlin
G.D.R.

Prof. E. Bischoff
University College of Swansea
Department of Management Sciences
and Statistics
Swansea SA2 8PP
United Kingdom

Prof. F. Breitenecker
Technical University Vienna
Gusshaustrasse 27
A-1040 Wien
Austria

Prof. M. Despontin
Free University Brussels
Pleinlaan 2
1050 Brussels
Belgium

Dr. Y. Dubov
Institute for Systems Studies
Prospect 60 Let Octyabria 9
117034 Moscow
U.S.S.R.

Prof. K.H. Elster
Technische Hochschule Ilmenau
Sektion Mathematik Rechentechnik
und ökonomische Kybernetik
Postfach 327
DDR-6300 Ilmenau
G.D.R.

Prof. J. Ester
Technische Hochschule
Sektion Automatisierungstechnik
DDR-9010 Karl-Marx-Stadt
Postfach 964
G.D.R.

Prof. G. Fandel
Fernuniversität
Gesamthochschule
Fachbereich Wirtschaftswissenschaft
Postfach 940
D-5800 Hagen
F.R.G.

Prof. A. Göpfert
Technische Hochschule
Carl Schorlemmer
Sektion Mathematik
Leuna-Merseburg
Otto Nuschkestrasse
DDR-4200 Merseburg
G.D.R.

Dr. W. Gabrisch
Technische Hochschule
Carl Schorlemmer
Leuna-Merseburg
Sektion Mathematik
Otto Nuschkestrasse
DDR-4200 Merseburg
G.D.R.

Mrs. R. Gehne
Academy of Sciences of the GDR
Institute for Informatics
Rudower Chaussee 5
DDR-1199 Berlin
G.D.R.

Dr. R. Genser
GUK-OE
Osterreichisches Gesamtverkehrskonzept
Bundesministerium für öffentliche
Wirtschaft und Verkehr
Operngasse 20b
A-1040 Vienna
Austria

Dr. R. Gollmer
Academy of Sciences of the GDR
Institute for Informatics
Rudower Chaussee 5
DDR-1199 Berlin
G.D.R.

Prof. M. Grauer
Academy of Sciences of the GDR
Division of Mathematics and Informatics
Rudower Chaussee 5
DDR-1199 Berlin
G.D.R.

Prof. J. Guddat
Humboldt Universität zu Berlin
Sektion Mathematik
Unter den Linden 6
DDR-1086 Berlin
G.D.R.

Prof. Hartmann
Technische Hochschule
Carl Schorlemmer
Sektion Werkstoff und
Verarbeitungstechnik
Leuna-Merseburg
Otto Nuschkestrasse
DDR-4200 Merseburg
G.D.R.

Dr. St. Kaden
Water Policies
IIASA
A-2361 Laxenburg
Austria

Prof. V. Kaftanov
Deputy Director
IIASA
A-2361 Laxenburg
Austria

Dr. M. Kok
Delft University of Technology
Department of Mathematics and Informatics
P.O. Box 356
NL-2600 AG Delft
The Netherlands

Prof. P. Korhonen
Helsinki School of Economics
Runeberginkatu 14-16
00100 Helsinki 10
Finland

Dr. M. Kunsch
Belgonucleaire
25 Rue du Champ de Mars
B-1050 Brussels
Belgium

Prof. A. Kurzhanski
Chairman
System and Decision Sciences Program
IIASA
A-2361 Laxenburg
Austria

Prof. A. Lewandowski
System and Decision Sciences Program
IIASA
A-2361 Laxenburg
Austria

Dr. K. Lindner
Institut für Energetik
Zentralstelle für rationelle
Energieanwendung
Postfach 27
DDR-7024 Leipzig
G.D.R.

Prof. F. Lootsma
Delft University of Technology
Department of Mathematics
and Informatics
P.O. Box 356
NL-2600 AG Delft
The Netherlands

Dr. Dinh The Luc
Computer and Automation Institute
Hungarian Academy of Sciences
P.O. Box 63
Kende ut. 13-17
H-1111 Budapest
Hungary

Dr. S. Messner
Energy Group
IIASA
A-2361 Laxenburg
Austria

Dr. B. Metev
Bulgarian Academy of Sciences
Institute of Industrial Cybernetics
and Robotics
1113 Sofia
Acad. G. Bonchev Street, Bl. 2
Bulgaria

Prof. H. Nakayama
Department of Applied Mathematics
Faculty of Sciences
Konan University
Higashinada
Kobe 658
Japan

Prof. S.C. Narula
Virginia Commonwealth University
School of Business
1015 Floyd Avenue
Richmond, Virginia 23284-0001
U.S.A.

Prof. R. Nehse
Technische Hochschule Ilmenau
Sektion Mathematik, Rechentechnik und
Ökonomische Kybernetik
Postfach 327
DDR-6300 Ilmenau
G.D.R.

Dr. L. Polmer
Technische Hochschule
Carl Schorlemmer
Leuna Merseburg
Sektion Mathematik
Otte Nuschkestrasse
DDR-4200 Merseburg
G.D.R.

Prof. Reinisch,
Technische Hochschule Ilmenau
Sektion Technische und Biomedizinische
Kybernetik
Am Ehrenberg
DDR-6300 Ilmenau
G.D.R.

Prof. M. Sakawa
Department of Systems Engineering
The Faculty of Engineering
Kobe University
Rokkodai Nada
Kobe 657
Japan

Prof. H.P. Schwefel
Universität Dortmund
Fachbereich Informatik
Lehrstuhl Systemanalyse
Postfach 50 05 00
D-4600 Dortmund 50
F.R.G.

Prof. H.J. Sebastian
Technische Hochschule Leipzig
Sektion Mathematik und Rechentechnik
Postfach 66
DDR-7030 Leipzig
G.D.R.

Prof. F. Seo
Kyoto Institute of Economic Reseach
Kyoto University
Kyoto
Japan

Dr. G. Slavov
Bulgarian Industrial Association
134 Rakovski Street
1000 Sofia
Bulgaria

Prof. J. Spronk
Erasmus University Rotterdam
Faculty of Sciences
P.O. Box 1738
NL-3000 DR Rotterdam
The Netherlands

Prof. I. Stantchev
Institute for Systems Analysis
and Management
Economic University 'Karl Marx'
Bul. G. Dimitrov 53
Sofia 1000
Bulgaria

Prof. R.E. Steuer,
The University of Georgia
Department of Management Sciences
College of Business Administration
Brooks Hall
Athens, Georgia 30602
U.S.A.

Prof. G. Stiller
Academy of Sciences of the GDR
Institute for Informatics
Rudower Chaussee 5
DDR-1199 Berlin
G.D.R.

Dr. R. Straubel
Zentralinstitut für Kybernetik
und Informationsprozesse
Abteilung Systemanalyse
Kurstrasse 33
DDR-1080 Berlin
G.D.R.

Dr. M. Strubegger
Energy Group
IIASA
A-2361 Laxenburg
Austria

Prof. H. Tamura
Osaka University
Department of Precision Engineering
Faculty of Engineering
2-1 Yamada-oka, Suita
Osaka 565
Japan

Dr. T. Tanino
System and Decision Sciences Program
IIASA
A-2361 Laxenburg
Austria

Dr. I. Valyi
System and Decision Sciences Program
IIASA
A-2361 Laxenburg
Austria

Dr. M. Voigt
Academy of Sciences of the GDR
Institute for Informatics
Rudower Chaussee 5
DDR-1199 Berlin-Adlershof
G.D.R.

Prof. J. Wallenius
Arizona State University
College of Business
Department of Decision and Information
Systems
Tempe, Arizona 85287
U.S.A.

Dr. K. Wendler
Humboldt Universität zu Berlin
Sektion Mathematik
Unter den Linden 6
DDR-1086 Berlin
G.D.R.

Prof. A. Wierzbicki
Technical University of Warsaw
Institute of Automatic Control
Ul. Nowowiejska 15/19
00665 Warsaw
Poland

Dr. L. Winkelbauer
Advanced Computer Applications
IIASA
A-2361 Laxenburg
Austria

Prof. P.L. Yu
The School of Business
University of Kansas
307 Summer Field Hall
Lawrence, Kansas 66045
U.S.A.

Dr. Ziembla
Institute of Automatic Control
and Systems Engineering
Academy of Mining and Metallurgy
Al. Mickiewicza 30
Krakow
Poland

This series reports new developments in mathematical economics, economic theory, econometrics, operations research, and mathematical systems, research and teaching – quickly, informally and at a high level. The type of material considered for publication includes:

1. Preliminary drafts of original papers and monographs
2. Lectures on a new field or presentations of a new angle in a classical field
3. Seminar work-outs
4. Reports of meetings, provided they are
   a) of exceptional interest and
   b) devoted to a single topic.

Texts which are out of print but still in demand may also be considered if they fall within these categories.

The timeliness of a manuscript is more important than its form, which may be unfinished or tentative. Thus, in some instances, proofs may be merely outlined and results presented which have been or will later be published elsewhere. If possible, a subject index should be included. Publication of Lecture Notes is intended as a service to the international scientific community, in that a commercial publisher, Springer-Verlag, can offer a wide distribution of documents which would otherwise have a restricted readership. Once published and copyrighted, they can be documented in the scientific literature.

Manuscripts

Manuscripts should be no less than 100 and preferably no more than 500 pages in length. On request, the publisher will supply special paper with the typing area outlined and essentials for the preparation of camera-ready manuscripts. Manuscripts should be sent directly to Springer-Verlag Heidelberg or Springer-Verlag New York.

Springer-Verlag, Heidelberger Platz 3, D-1000 Berlin 33
Springer-Verlag, Tiergartenstraße 17, D-6900 Heidelberg 1
Springer-Verlag, 175 Fifth Avenue, New York, NY 10010/USA
Springer-Verlag, 37-3, Hongo 3-chome, Bunkyo-ku, Tokyo 113, Japan

---

ISBN 3-540-16785-4
ISBN 0-387-16785-4