1-1-1978

Prioritization and pruning in multicriterion mathematical programming.

Joel N. Morse
University of Massachusetts Amherst

PRIORITIZATION AND PRUNING IN MULTICRITERION MATHEMATICAL PROGRAMMING

A Dissertation Presented

By

J. N. Morse

Submitted to the Graduate School of the
University of Massachusetts in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

June 1978

School of Business Administration
Department of General Business and Finance

PLEASE NOTE:

Dissertation contains computer print-outs that have light and broken print.
Filmed as received.

UNIVERSITY MICROFILMS

PRIORITIZATION AND PRUNING

IN MULTICRITERION MATHEMATICAL PROGRAMMING

A Dissertation Presented

By

J. N. Morse

Approved as to style and content by:

Kenan E. Sahin, Chairperson of Committee

Joseph L. Balintfy, Member

Van Court Hare, Member

Hugh J. Miser, Outside Member

Donald Frederick, Director of the
Ph.D. Program in Business Administration

ACKNOWLEDGMENTS

I thank my committee, Professors Balintfy, Hare, Miser, and Sahin for their time, talent, support and patience. In particular I am grateful to my chairman, Kanan Sahin who has spent long hours with me on this project. Professor Sahin's example was one of the main reasons I decided to pursue a doctoral degree.

I thank Commander Rolf Clark of the U.S. Navy because the conference papers we wrote together in 1974 started my speculations that later evolved into this dissertation. I acknowledge Professor Craig Moore, who taught me the difference between David Rockefeller and Milton Friedman. Craig was thus the friend most responsible for my decision to become an academician. Thanks also are deserved by Professor Sidney Sufrin who rounded out my education and made me interested in economics again.

To my colleagues at the University of Delaware I am grateful for financial, intellectual and emotional support during the last year. Mrs. Rita Beasley was far more than a typist. She helped me remember how to write, and her attention to detail was a boon to me in this dissertation process. There are also all the family and friends, too numerous to name, who helped me come to this point.

TO

Megan, my daughter

FRONTISPIECE

Programming sticks upon the shoals
Of incommensurate multiple goals
And where the tops are no one knows.
When all our peaks become plateaus
The top is anything we think
When measuring makes the mountain shrink.

The upshot is, we cannot tailor
Policy by a single scalar,
Unless we know the priceless price
Of Honor, Justice, Pride and Vice
This means a crisis is arising
For a simple-minded maximizing.

A headnote by Kenneth Boulding, contributed to Chapter V of Charnes, A. and Cooper, W. W., "Constrained extremization models and their use in developing system evaluation measures," in Views on General Systems Theory, M. D. Mesarovich, ed. (New York: John Wiley and Sons, 1964).

ABSTRACT

Prioritization and Pruning

In Multicriterion Mathematical Programming

(May 1978)

J. N. Morse, A.B., Williams College
B.B.A., University of Massachusetts
Ph.D., University of Massachusetts

Directed by: Professor Kenan E. Sahin

This dissertation aims at filling some gaps in the treatment of the multiple objective situation, particularly the subset known as the multiobjective linear programming model. First, the appropriateness of multicriterion models is examined, rather than being assumed. Second, both behavioral science and mathematical reasoning are invoked to show that goal programming is not so appropriate as multiobjective linear programming. Third, to deal with the very large size of what is called the nondominated set generated by multiobjective programming, a two-stage pruning algorithm is proposed. In the first stage cluster analysis is employed to minimize redundancy and avoid information overload. In the second stage a series of the duals of a linear programming problem guides the decision maker in his search for one final solution from the nondominated set. The two-stage pruning algorithm appears to be applicable to a wide range of multidimensional choice situations.

The pluralism of modern societies implies the existence of multiple objectives, but the traditional microeconomic theory of the

firm still consists of a single criterion maximizing paradigm. Wealth maximization has served well as an approximation of commercial behavior; it is an inadequate description of basic business processes themselves. To show that the multiple objective situation is characteristic of the firm, a taxonomy is developed. I present three views of the firm: the disaggregation approach, the agency approach and the autopoietic approach. A hypothetical example from the field of corporate finance is used to illustrate that there are situations in which the optimal value of a single objective function (such as profit) is no more important than the values of the decision variables themselves.

The purpose of this dissertation is to examine in depth a subset of this multiple objective situation, namely the multiobjective linear programming model. Although this model can apply to groups as well as to individuals, I concentrate on the case of the single decision maker acting alone or articulating group objectives.

The literature search of multiple criteria decision making is divided into a mathematical chapter and a behavioral chapter. Topics covered include some optimization theories, aspects of human choice theory, and human information processing. This background leads to two mathematical programming techniques for solving multicriterion problems: goal programming and multiobjective linear programming. I use the theory of eigenvalues to show that, for a large subset of decision models, goal programming is less appropriate than multiobjective linear programming. The most difficult task of goal programming is that of prioritization or setting weights on goals.

Having thus justified multiobjective linear programming, I discuss a new optimization concept that it uses. Rather than a single optimal point, there is a nondominated (efficient, or Pareto-optimal) set. This set consists of solutions that are preferred to all solutions outside it; however, within the set no rank order exists unless new evaluative criteria are introduced. Although sequential, information-seeking, and imitative thought processes are discussed, the most attention is paid to algorithmic (mechanistic) methods.

The overwhelming handicap of the multiobjective linear programming approach is the sheer size of the nondominated set of solutions. Since this is true even in the nonlinear case, I attack the more general problem of reducing the size of the nondominated set in any multicriterion mathematical programming problem. This process will be called pruning. The tactic adopted is to divide pruning into two stages. In the first, most of the work is done by the computer. In the second, much more of the task is completed by the decision maker.

In the first stage a data analysis technique called cluster analysis is used to prune the nondominated set down to a representative subset called the generator set. Many clustering experiments are reported, and six other less satisfactory pruning strategies are reviewed.

In pruning's second stage, ranking, weighting, paired comparisons and linear programming are combined to further cut the generators of the nondominated set down to one final solution. The algorithm proposes reducing the amount of data manipulation by the decision maker, but promises to increase its sophistication and usefulness. The goal is to encourage the decision maker to consider compensatoriness, incommensurability and interdependence of attributes and to discover new evaluative criteria.

In sum, this dissertation explores the applicability of multiple criteria decision making to individual decisions. The prioritization (ranking and weighting) of objectives is shown to be more difficult than commonly perceived. This leads from the relatively naive goal programming model to the more appropriate multicriterion mathematical programming paradigm. The unwieldiness of the very large nondominated set produced by the latter is reduced through a two-stage pruning algorithm. The computer is used to do the sort of structured combinatorial reasoning that it does best, while the human mind is liberated for what it does best--unstructured, holistic and creative thinking.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

CHAPTER I

THE MULTIPLE OBJECTIVE SITUATION

This dissertation will examine decision making in the multiple objective situation. Earlier literature has concentrated on a particular segment of the decision process, leaving several gaps in the body of techniques called multiple criteria decision making (MCDM). The aim of the present work will be to create and justify a two-stage pruning process for treating the nondominated set which results from multiobjective mathematical programming (linear or nonlinear).

Two important questions will be posed and responded to:

a) Should the decision maker use goal programming or multiobjective linear programming?

b) What can be done to aid the decision maker in choosing one final solution from the large unwieldy set of nondominated solutions which is generated if he uses multiobjective linear programming?

The area stressed is the multiobjective linear programming model of the individual decision maker. To answer the first question, I shall review behavioral and mathematical support for the rejection of *a priori* ranking methods. The implication of this is that goal programming may be less applicable than multiobjective linear programming, since in using the latter technique it is not necessary to prioritize the objectives. For the second question a two-stage pruning algorithm is proposed; this aids the decision maker to search a large nondominated set for one final solution.

The latter part of this chapter of my dissertation will deal briefly with several theories of the industrial firm, with the multiple objectives inherent in public organizations (i.e., non-profit), and with the need to extend research into the group multiple objective decision situation.

The second chapter will search the literature of mathematical programming. I begin with the origins of the field in general, and then discuss the optimization approach to the special problem of decision making with multiple objective functions. Since models that are not foreign to the mind's perceptual and cognitive habits are more likely to be successfully implemented, Chapter III will treat the individual choice theories of mathematical psychology.

The natural limits on man's ability to process information will be viewed as significant in the choice of a multicriterion model. Several disciplines provide examples which pertain to this. The result of Chapter III is a pessimistic assessment of our ability to set *a priori* weights on multiple goals. This does not mean, however, that MCDM is not practical. It indicates that goal programming (GP) may be far less applicable than multiobjective linear programming (MOLP), because MOLP does not require advance setting of weights on goals. Whereas GP pre-empts the decision by prematurely discarding most of the nondominated set, MOLP relinquishes control to the DM in the pruning stage.

Chapter IV drives in the same direction as the previous chapter;

this time the reasoning is mathematical, rather than behavioral. The way in which the priority weights in GP drive a linear equation is hard to predict. The effects of changing the weight vector can be counterintuitive. This will be illustrated by a discussion of naive weights. Naive weights are obtained when the DM associates the highest weights with the most important goals. I will analyze the unsatisfactory nature of this weighting procedure by working with the eigenvalues and eigenvectors of the criterion matrices of several representative goal programming problems.

In Chapters V and VI we face the problem of the often overwhelming size of the nondominated set (N). Although integer and other nonlinear cases in multicriterion programming are still under development, they do share with the linear case the concept of nondomination. We attack the general problem of reducing the size of this nondominated set of solutions. This process will be called pruning; here it emerges as a two-stage process. In the first stage most of the work is done by the computer. In the second, much more of the task is completed by the decision maker.

The fifth chapter treats the first stage of pruning by using the data analytic technique called cluster analysis. The idea is to generate a representative subset of the nondominated set (N). Cluster analysis partitions N into groups that are relatively homogeneous within their boundaries. Broadly speaking, a very general new evaluative criterion is to be considered implicitly: minimum redundancy. Since there is a threshold of resolution which hinders the decision maker in perceiving the difference between two very similar solution vectors, there is little point in making him waste time in processing all of N as he searches for a final solution.

Chapter VI uses a limited concept of weights once we have reduced the nondominated set to a manageable size. In this second stage of the pruning process weights are used for fine tuning, just as we turn the knobs once a channel has been chosen. By trial and error we use the feedback from our eyes to move toward the best picture (knowing that there might be a better setting that is not easily discoverable). The ranking algorithm of this chapter asks the DM to make binary comparisons between particular elements of the nondominated set. He does not deal directly with criterion weights, just as the person using a television set does not need calibrated knobs. The TV viewer reacts to "better" or "worse" pictures, and moves the knobs as a consequence. The DM in the ranking algorithm of Chapter VI responds to alternatives that are available in decision space as better or worse; the weights merely fall out as a side effect of a proposed linear programming approach. As will be seen, they are not of very general applicability or reliability. It is suggested that they serve to confirm the DM's judgment and increase confidence in the final solution.

Since my viewpoint is that of management science, not mathematics, techniques developed here are designed to be relevant to the management of organizations. Whether an individual executive makes decisions for the organizations or multiple objectives derive from multiple agents, it is important to justify the use of multicriterion methods in organizational processes. From microeconomic theory we have inherited the concept that the firm is a black box that maximizes profit. A refinement is to add that the firm might sacrifice profit in some periods, because its true goal is to maximize the terminal wealth of its shareholders.

In reaction to this orthodox view of the firm a large literature has grown which speculates on alternative goals of the firm. Examples are theories of satisficing [Simon, 1969, pp. 64-65, 75-76], managerial welfare maximizing [Findlay and Whitmore, 1973], the agency theory of the firm [Jensen and Meckling, 1976], strategic goal setting [Quinn, 1977], natural selection [Winter, in Day and Groves, 1975], fatal flaw processes [Kierulff, 1976], partial ignorance [Loasby, 1976], absence of value maximization [Grossman and Stiglitz, 1977] and multiobjective firms [Beedles, 1977]. Three approaches will be discussed: the disaggregation approach, the agency theory approach and the autopoietic approach.

To illustrate the disaggregation approach, we draw on the normative theory of finance which maximizes shareholder wealth over time. When the firm must choose among available projects, it may use well-known tools such as the payback method, the net present value method, the internal rate of return method and the profitability index. When the actualization of the ranked projects would exceed the total funds available, the problem becomes one of capital rationing. A typical decision model of this sort is (cf. Weingartner [1963]):

$$\text{Maximize:} \quad \sum_{j=1}^{n} p_j x_j \tag{1.1.}$$

$$\text{Subject to:} \quad \sum_{j=1}^{n} a_{jt} x_j \leqq M_t \text{ for } t = 1, 2, \ldots, T$$

$$x_j = 0 \text{ or } 1, \; j = 1, 2, \ldots, n$$

where:

x_j is the go/no-go decision on project j,

p_j is the net present value of project j,

a_{jt} is the gross expenditure required for project j in time period t, and

M_t is the limit on capital funding in period t.

This model can be made richer by adding concepts of dividend growth, uncertainty, borrowing, reinvestment of income and logical relationship among projects (the reader who wishes to see how others have treated these extensions is referred to the excellent summary in Reardon [1974]).

Essentially, equations (1.1.) are concerned with profit maximization, adjusted for the time value and availability of money. Can this single objective serve to simulate all the goals of the firm? It can, provided that all corporate desiderata correlate reliably and linearly with profit. I feel that in most cases they do not.

For example, the New York money market banks are restricted from competing in certain fields with "Article XII" banks. These are special entities licensed by the New York State Department of Banking. As such, they can own brokerage houses, underwrite securities, and concentrate lending power for one borrower. According to Bennett [1976], a goal of the money center banks is to encourage federal regulation of the Article XII banks, or to gain similar powers for themselves.

Working toward this could be named as a goal for 1977. Citicorp, hypothetically, could allocate a block of funds for this purpose. But how would all this be worked into the objective function of the capital rationing model above? There is no neat relationship between the goal and profits. Is there a way to support this kind of strategic planning with a rational decision tool?

When the relationship of decision variables to the present value of future profits is fuzzy, managers tend naturally toward articulating <u>multiple goals</u>. For years management scientists forced these multiple goals into a single goal, or performance measure, because they lacked alternative models. To achieve the worthy goal of aiding human minds with the intelligence contained in "models," the subtle goal structures of the real world were aggregated into single goal models.

Recently, a new field called <u>multiple criteria decision making</u> (MCDM) has matured to the point of being useful. In the banking example above, MCDM does not require measuring the progress toward several goals in profit (dollar) terms. For each goal, a separate objective function can be specified, whose coefficients may be <u>incommensurable</u>[1] with those in other objectives. Here the secondary objective function might be:

---

[1] Incommensurable means "not measurable in the same units."

Maximize: $$\sum_{j=1}^{n} k_j x_j \tag{1.2.}$$

where k_j represents the face value of loans "kept away" from Article XII banks by the j^{th} project.

It is claimed here that multiple objectives (goals) should be considered explicitly. An example from international banking has been used to illustrate the case of disaggregation of objectives.

The agency theory of the firm views the commercial organization as a complicated polity of stakeholders. The firm originates by investors pooling their capital and sharing the risks inherent in the commitment to productive activity. Therefore, the dominant group is the holders of the common stock. As owners, this group subcontracts certain tasks to other groups.

Management, for example, serves as the agent of the common shareholders, performing a stewardship function. Groups such as the Securities and Exchange Commission (SEC) and the Financial Accounting Standards Board (FASB) provide a monitoring service for owners. When the owners of a firm seek amplification of their income-providing asset base by borrowing money, they must pay bonding costs, which are incurred in demonstrating the ability to make good on the claims of debt-holders. The SEC and the FASB also provide bonding services, since they attest to the viability of the firm, and to some extent make it harder for new competitors to raise capital. Other forms of bonding costs are restrictions on dividend payout ratios, guidelines for debt/equity ratios and liability insurance premiums.

In some sense the owners hire employees, accountants, government regulators and debtholders as agents to perform services that are exxential for shareholders. An intriguing example is a new view of securities analysis. It is well known that many studies have failed to prove that professional securities analysts add value each period to investors' portfolios. Why then do they continue to exist? Who pays them? According to the agency theory, these securities analysts are funded by society for two reasons. First, there is a large consumption value to capital and money market investors. Second, the analysts discipline corporate managers and promulgate rational thinking throughout an industry as they go to print with detailed comparative analyses of companies. They are an added level of executive talent; the existence of this group reduces the variance of the returns to firms, and hence adds to their capitalized value.

Agency costs consist of monitoring expenses by the owners, bonding expenses which are paid by agents on behalf of owners, and what Jenson and Meckling [1976, p. 308] call "residual loss." The definition of this term will show the relevance of the multiple objective model.

Each owner has a utility function that is presumably monotocally increasing in terminal wealth. But each agent described above exercises some control over the income-producing assets. Although an optimal level of monitoring and bonding costs may exist, the inclusion of multiple stakeholders in the firm as agents and claimants introduces their utility functions. For example, the chief executive officer of the modern corporation is interested in maximizing the share price. But he/she also derives utility from a larger expense account, fancy office furniture, prestige due to span of control (sales maximization versus profit maximization) and reduction of personal risk. Similarly, the utility of each agent diverges, to some extent, from that of the owner(s). That divergence is called the residual loss.

The stakeholders (i.e., owners and agents) must somehow aggregate their objectives into a firm that survives. The political (consensus-building) processes of the modern firm tend to generate multiple objectives; this is often a way to reduce conflict among agents. The communication if of the form "I'll help achieve your objective if you'll help achieve mine." In this way we can see the firm evolving as a set of maximands subject to a set of constraints.

Another justification of the multiple criteria approach is the theory of self-organizing firms. The concept, under the name <u>autopoiesis</u>, developed in the biological systems area (for that history, see Zeleny [1977]). An autopoietic organization is one that maintains itself by the action of particular processes on the system's components. These processes are essentially local rules of interaction which generate the same kind of network from which they came. Examples of these closed systems are cells reproducing cells, social groups producing people that maintain group norms and corporations which perpetuate managerial styles and levels of employee morale.

Zeleny [1977] develops a very lengthy algebra of the behavior of the components and operators that are used to describe autopoietic systems. He then gives examples from biology and moves on to speculate about social organization. Societies maintain continuity and recognizability by rules of conduct that have been generated from their own history. Human systems, he suggests, have evolved into very complex systems; these cannot be changed by simplistic rules such as theory X or theory Y. They can be managed, to some extent, by providing certain environments. An example is making deserts bloom by providing irrigation, thus changing the environment.

The relationship to the firm in multiple objective situations (MOS) follows naturally. We must be cautious about assuming that firms or not-for-profit organizations are purposive. It may only seem that they have multiple objectives. For example, water running down a hill may appear to have that as its "objective." But we also know that water is an inanimate form which is moving as the residual effect of the interaction of gravity and mass. Our intent here is not to determine whether organizations have motives, objectives, or purpose. I have proposed that under the agency theory and the disaggregation paradigm the firm either actually has multiple objectives, or it can be more effectively managed and described by resorting to them as a hypothetical construct. In this section I state, without formal proof, that even if the firm is merely a set of interactive components, the knowledge of MCDM is still of great importance.

If the firm's behavior is not purposeful, but instead autopoietic, we have a collection of individuals and groups which react to stimuli.

Various theories of how they do this are found in Chapter III. In brief, multiple cues are perceived by cognitive agents (such as people or groups of people), which react to each cue with one of a set of multiple responses.

In Chapter II the formal multiobjective linear programming model will be introduced as

$$\begin{aligned} &\text{Maximize} \quad && Cx = \{z\} \\ &\text{Subject to} \quad && Ax \leqq b \end{aligned} \tag{1.3.}$$

where C represents a matrix formed by multiple objectives.

This is a paradigm which assumed top-down purposive organization. If the autopoietic view of the firm is more correct, then we would consider the multiple criteria as <u>multiple attributes of cues</u>. The coefficients of C would measure the intensity of attributes and their tentative direction (i.e., good or bad, positive or negative). In other words, MCDM would be used not as a decision-making technology, but instead as a multiattribute response model (MARM). Executives, in this view of the world, have not articulated objectives, but do react for some reason to the attributes inherent in different corporate strategies.

A hypothetical example will now be constructed to illustrate the agency theory, the disaggregation paradigm, and the autopoietic approach to the firm.

To make the above discussions tangible we call on the field of financial management for an example in which the firm must choose a subset of an opportunity set of capital investment projects. For a more thorough treatment of capital budgeting with side criteria, see Bernardo and Lanser [1977] and Morse [1978]. This hypothetical firm will also be referred to in later chapters.

Imagine that there is a firm named Transoceanic, Inc., a large multinational corporation. It operates under some form of capital rationing, and at first it appears that (1.1.) is a good model of Transoceanic. But when the capital outlay planning committee meets in June, one of the members brings of a strategic planning dilemma that does not fit into the single criterion model of (1.1.).

The firm's marketing group is about to close on a $500 million construction contract in a wealthy foreign nation. A certain citizen of that country purports to be the agent of the ruling clique. He has informed Transoceanic that a bribe of $50 million must be paid to him; in turn the agent will guarantee that the company will get the job. Since the ruling clique, for reasons of politics, tradition, and dignity, cannot be open about receiving side payments, they will not even officially confirm that the "agent" works for them.

The $50 million side payment could be added onto the sales price of the contract. But should Transoceanic pay that vast amount of money to a person who might not really represent the ruling group? This uncertainty is hard to represent in the value maximizing objective function of (1.1.).

Will the directors feel that paying bribes is immoral? Does the staff analyst in charge think the same? Or is he/she, by chance, violently sympathetic to the cause of the client nation involved? Already we see two stakeholders in this firm, each having at least one possible objective. This is an example of the ideas of the agency theory of the firm.

Suppose that the bribe were paid and the contract signed. Presumably, the total market value of Transoceanic would rise by an amount equal to the risk-adjusted (time) discounted future cash flows associated with this new business. It is possible that within six months auditors of the Securities and Exchange Commission will discover the bribe. The resultant fines, bad publicity and managerial involvement might reduce the value of the firm by more than the above amount! Several objectives arise naturally: maximize "morality," minimize the probability of being "caught," maximize the probability that the agent truly can guarantee the contract, maximize stability of the firm's employment level (to aid morale and minimize overtime in rush situation), maximize rate of sales growth, and finally maximize earnings per share for the upcoming fiscal year. These optimands do not correlate linearly with the net present value of the firm. There are complicated linkages with the existing activities of the firm.[2] As an example of the contingent nature of the problem, the probability of being caught by the auditors correlates inversely with the size of the contract finally agreed upon. And the perpetuation of the ruling clique of the client state is not independent of the quality of the infrastructure investments it provides for its citizens. But the

---

[2] In this view the firm is a "mutual fund of projects."

likelihood of war with neighboring states is a complicated function of the nation's industrial capacity. In short, how could any alert financial analyst possibly condense this entire discussion into one of the coefficients p_j of (1.1.)? Firms do face decisions like this one; disaggregating the metacriterion of profit into multiple objectives may make it easier to tolerate this complexity.

For specific examples of this complex goal structure, let us look more closely at Transoceanic. It is considering four other capital outlays which are described in Table 1.1. At this point the data are used only to illustrate the disaggregation approach. In later chapters I will make more analytical references to the investment opportunity set of Transoceanic, Inc.

Table 1.1

Investment Opportunity Set of Transoceanic, Inc.

Attribute or objective (criterion)	Project A	Project B	Project C	Project D
Net present value	\$1,000,000	\$1,500,000	\$900,000	\$1,400,000
Percentage of the project's fixed assets that will be in politically unstable countries	30%	40%	30%	40%
Is there significant foreign exchange risk (low, medium or high)?	medium	medium	high	high
Is the project in South Africa (cf. opposition of church groups, etc.)?	yes	yes	no	no

Many people feel that corporate directors will, when interviewed, agree that they are interested in many goals. Almost any goal[3] that the interviewer can name has something good about it. Quinn [1977] points out that chief executives are particularly unwilling to articulate precise operating goals such as per share earnings figures. He says that the more concrete the goal, the more likely is personal disutility if it is not achieved. But chief executives often do articulate metacriteria such as "Be number one in market share for our main products," or "Provide great service," or "Be the small car company" (e.g., American Motors Corporation). More specific goals, although articulated by some companies, can backfire by creating rigidity and focusing opposition within the corporation. These are some of the confusing side effects in the use of management by objectives (MBO) and zero-base budgeting (ZBB).

While objectives may not be precisely articulated, it is certainly true that the management of Transoceanic would have reactions to the four capital investment projects of Table 1.1. Perhaps each project should be characterized as a set of attributes, rather than as criterion-achieving actions. Then the multiple stakeholders in the affairs of Transoceanic would react to the projects. Some form of aggregating, consensus-building political process would ensue; the firm would act. This is what the autopoietic approach is attempting to explain (also see the Brunswik Lens Model in Chapter V).

---

[3] I define objective as an unbounded direction of aspiration, measured by a scale, and goal as a particular target point on that scale.

Having promoted the need for multiple criteria decision making in general, this dissertation henceforth will treat the limited case of the individual decision maker (DM). This DM may act on his own behalf, or he/she may represent the culmination of some group process. For a sample of research on the group, or social decision situation, see Johnson [1968], Keeney and Raiffa [1976], Arrow [1951], and Fishburn [1973].

CHAPTER II

PREVIOUS RESEARCH IN MULTICRITERIA PROGRAMMING

II.1 Historical Remarks

The word optimum was revived by Leibniz. Etymologically the source was Ops, the Roman goddess of fertility and agricultural abundance [Wilde and Beightler, 1967, p. 468]. These happy circumstances seemed like the "optimum," or best, of all worlds. Classical optimizers did not have to face the problem of what best meant. Early applications were directed toward physical science applications. The calculus was developed to find extrema of functions of at most several variables.

To handle initial conditions, or side conditions, the concept of constrained maximization (minimization) was necessary. Lagrange suggested his "multipliers," a simple method which has proved useful over the years in more sophisticated problems, such as integer programming. Then problems involving linear functions were envisioned in which the roots of the system were to be confined to the non-negative domain. But since no pressing application fit this format, no solution techniques were immediately found.

Economic philosophers began to conceive of "homo economicus" allocating scarce resources. According to George Dantzig [1963, p. 16] the first interface between linear models and economics occurred in 1759 in Le Tableau Economique by Quesnay. No known successors to

Quesnay worked on the problem in the nineteenth century. However, men like Cauchy, Jordan, Fourier, Farkas, Minkowski, and Motzkin worked on other aspects of linear equation systems.

During World War II, the British military organized accountants, mathematicians, physicists, and others into a multidisciplinary team which they named the "operational research" group. The United States followed suit; in 1947 Project SCOOP was funded by the Air Force. SCOOP meant "scientific computation of optimum programs"; the tradition of acronyms in mathematical programming had begun.

This research group, whose activities are described by George Dantzig [1963, Ch. 2], generalized the Leontief model of the economy. That model was called the "linear programming problem" (LP). By the end of the summer of 1947, the group had a reliable solution technique which was called the "simplex method," since its finiteness depended on the use of a geometric structure called the simplex. It is curious that a Russian mathematician named Kantorovitch worked on a similar theory in Russia in the late 1930s. But his work was not discovered in the West until much later.

From this seminal point grew a study known first as "activity analysis," and later as linear programming. Literature in this field now includes thousands of citations.

II.2 Linear Programming

The LP approach to economic modeling takes the following well-known form (with superscript T denoting transpose and c, x, A and b properly dimensioned vectors and matrices):

$$\text{Maximize } c^T x \qquad (2.1.)$$
$$\text{Subject to} \quad Ax \leqq b$$
$$x \geqq 0$$

One can see that the objective function is a scalar; one objective, defined by its "price coefficients" c is multiplied by a conformable activity vector x. For example, in classical micro-economics, the firm is presumed to maximize its profit, or minimize its costs, subject to scarce resources (the "right-hand side," RHS). When a model-builder wishes to incorporate other objectives, the standard approach has been to use profit as a surrogate for them. The working assumption, in this approach, is that other objectives such as safety, legality, aesthetics, market phase, longevity, net worth and debt position all have some strong, stable relationship with profit.

Sometimes the firm's decision variables are rather clearly related to these other goals. It may also be that the relationship between decision variables and profit is obscure. In this way multiple objectives become explicitly and separately stated.

II. 3 Multiple Objectives and Commensurability

Koopmans [1951, pp. 33-97] was the first to talk of multiple objectives or goals. He introduced the formalism known as the vector maximization problem (VM):

$$\underset{x}{\text{Maximize}} \quad \{f_1(x), f_2(x), \ldots, f_n(x)\} \qquad (2.2.)$$
$$\text{Subject to} \quad g_k(x) \leqq b, \quad k = 1, 2, \ldots, m$$

where x is an n-dimensional vector of decision variables; $f_i(x)$, $i = 1, 2, \ldots, n$ represent n objective functions; and $g_k(x)$, $k = 1, 2, \ldots, m$ are m constraint functions.

The objective functions $f_i(x)$ can be "commensurable" or "noncommensurable." Whether or not we attempt to aggregate the objectives into a super-objective function, the problem of commensurability is a severe one. We want to avoid adding apples and oranges, as we tell our children.

Price theory in economics deals with this issue. Physically, noncommensurable quantities that trade freely in noncoercive markets are nicely commensurated by monetary units. Another example of commensuration via a pricing mechanism is the attempt of the federal pollution laws to "charge" violators for environmental damage. The government defines certain polluting activities to be non-criminal, but damaging to the environment. By the use of fines, this environmental objective is incorporated into the firm's goal structure via the numeraire of dollars. Some experimental techniques for massaging the coefficients of the objective functions can be found in Morse and Clark [1975] and Widhelm and Doyle [1976].

When faced with absolute noncommensurability, rational decision making is not possible. For example, we teach our children that they can add apples, but not apples and oranges. But can we even add apples when they differ in weight, taste, and dimensions? Of course we can; we make the apples commensurate by disregarding certain attributes that differ. We say, that for the purposes of addition, all apples are identical. The point is that commensurability is a mental construct. In reality, it rarely exists, but we achieve it, when necessary, by modifying our representation of the world.

In practical affairs, fortunately, humans are able to take objectives (or goals) and rank them, trade them off against each other or partition them into types; in short, we can act, within limits, as if our objectives were commensurable to some degree. To the extent that we make them commensurable, they are commensurable. That is how objectives differ from real apples and oranges.

To review this issue we could recall the vector maximization problem (2.2.). If the $f_i(x)$ were perfectly commensurable, linear, and separable, then they could simply be summed up into a super-objective function of the form

$$z = [\sum_{i=1}^{n} f_i(x)] \tag{2.3.}$$

In short, this is a new linear programming problem.

II.4 Weighted Sums

Suppose that the decision maker(s) felt that they could construct a vector

$$\theta = [\theta_1, \theta_2, \ldots, \theta_n]$$

that weighted the n objectives according to their importance. Inserting θ into the objective function would yield the Weighted Sums Problem (WSP).

$$\underset{x}{\text{Maximize}} \quad \sum_{i=1}^{n} \theta_i f_i(x) \tag{2.4.}$$

$$\text{Subject to} \quad g_k(x) \leqq 0, \; k = 1, 2, \ldots, m$$

$$_i \geqq 0, \; i = 1, 2, \ldots, n$$

$$\sum_i O_i = 1$$

A major theme of this dissertation will be that this weight vector cannot be constructed a priori. A technical problem with the parametric approach is the potential for what is called the "duality gap."

Due to its simplicity, the weighted sums method has been popular in applied work [Major, 1969; Haimes, 1970; McGrew and Haimes, 1974; Geoffrion, 1968; Reid and Vemuri, 1971; Krajewski and Ritzman, 1973].

II.5 Vector Maximization and Multiobjective Linear Programming: Work of Historical or Peripheral Importance

Kuhn and Tucker [1951] amplified Koopman's [1951] idea of vector maximization in their oft-cited article on nonlinear programming. They developed, in particular, a duality theory that showed that a point $x°$ in the constraint set is efficient if and only if its dual vector, which is a sort of "price" vector, is positive. They left a void to be filled in by later researchers, since LP algorithms deal with nonstrict inequalities. In other words, these methods cannot be used directly to find a dual vector that is strictly positive.

Bod [1963] gave an algorithm that decides whether a point $x°$ is efficient. Geoffrion [1968] created a whole theory of vector maximization which included nonlinear programming as a special case. He formulated a new version of what Kuhn and Tucker had called "proper efficiency."

DaCunha and Polak [1967] traced vector maximization from Pareto [1896], through Karlin [1959], Debreu [1959], Kuhn and Tucker [1951] and finally to the control theory literature of Chang [1966] and Zadeh [1963]. DaCunha and Polak explained how a vector-valued criterion usually induces a partial ordering on the set of alternatives. This excludes the notion of optimality, and efficiency is discussed instead. Their paper developed a theory of necessary conditions for characterizing efficient points, and ways of knowing when a vector-valued criterion problem could be decomposed into a family of problems with scalar-valued criteria. Philip [1972] reviewed the literature and presented several algorithms for the existence and uniqueness of efficient points. His efforts are somewhat forgotten, in the sense that later authors cite him in a perfunctory way.

Benayoun, Montgolfier, and Tergny [1971] were the first to bring multiobjective linear programming (MOLP) to the decision maker (DM) in a well-documented, practical way. Their STEM method involves a sequential exploration of solutions. This exploration is aided by a carefully defined interaction between the DM and the algorithm. This man-machine symbiosis need not converge. If after p iterations (the number of objectives) the DM is not satisfied, the algorithm concludes that there is no "best compromise solution." It is hoped that the process will enable the DM to learn to recognize good solutions and the "weights" of the various objective functions. In general, this is the approach taken by the active researchers in the field of MOLP. I believe that the reliance on the DM to modify constraints is precisely the kind of job for which the human mind is poorly equipped. This idea will be expanded in Chapter III.

Roy [1971 and 1973] carefully expounds his method called ELECTRE. He explicitly treats aggregation of objectives, progressive definition of preferences, partial orders on alternatives, uncertainty and incomparability. ELECTRE is extremely cumbersome, although it may be very worthwhile to those DMs trained for it.

Belenson and Kapur [1973] add to the growing machinery for MOLP. They form pay-off matrices, as in two-person zero-sum games. Then during interface with the DM, information is elicited, the problem is respecified, and sequential decision making proceeds. This approach has not been adopted by many other authors.

Kornbluth [1974] makes a major contribution. He goes back to the duality theory of vector maximization that was started by Kuhn, Tucker and Geoffrion. He shows that there is a dual solution π^* for every efficient MOLP solution vector x^*. He calls this π^* the "prices" of the resources of the problem. They depend on the subjective weights of the primal problem. He shows how different trade offs between objectives can lead to the same efficient solution but different valuations of resources. He also is the first author to propose a theory of "post-optimality analysis" for MOLP. An interesting concept is that the right-hand side of a MOLP problem is an aggregation, into one vector, of many separate resource vectors.

Silverman and Hatfield [1973] produce outlines of satisficing strategies. A complete ordering of objectives is never required from the DM. Goals are allowed to be either flexible or inflexible. This is done by introducing "preference sets." The preference set identifies all objective functions which are of lesser importance to the DM, a priori. Their iterative procedure first solves a goal programming problem without incorporation of the preference sets. Using this as a starting solution, the feasible space is augmented to include the preference sets.

Lawrence and Koch [1975] and Lawrence and Lawrence [1975] have written many applications papers that have appeared in conference proceedings. There is no algorithmic or theoretical research here. An interesting feature, however, is that real problems are solved both by MOLP and by goal programming.

A tradition of solving MOS problems has developed in management science. A parallel technology has grown up in the field of water resource management, but there has been almost no cross-fertilization!

In the field of water resource management, falling under the rubric of cost-benefit analysis, there is a well-worked-out system called the Environmental Evaluation System (EES) [Dee et al., 1973]. The EES is based on a hierarchical arrangement of quality indicators covering ecology, environmental pollution, aesthetics and human interest. From these four categories 78 parameters are developed, which can be tracked empirically. Next the research team assigns values between 0 and 1, to indicate how each parameter affects each of the four major areas. Then these value functions are summed to yield commensurate units called "environmental impact units." Using these units, the environment is assessed with, and without, each project. This is supposed to indicate whether the proposed projects will have a positive or a negative effect on the environment. Economic effects and constraints are not incorporated into EES. Hence the whole process, despite its sonorous name, is just a careful implementation of a very simple unconstrained maximization concept.

Another very simple effort in this literature is Personal Value Information [Groves and Kahalas, 1975]. This is a way of gathering preference data from the <u>local</u> population.

Somewhat more interesting is the ε constraint approach. This method [Haimes, 1970] replaces (n-1) objective functions with (n-1) constraints as in

$$\begin{aligned} &\underset{x}{\text{Maximize}} \quad F_i(x) \\ &\text{Subject to} \quad g_k(x) \leqq 0 \,, \quad k = 1, 2, \ldots, m \qquad (2.5.) \\ &\qquad\qquad\quad f_j(x) \leqq \hat{\varepsilon}_j \,, \quad j \neq i, \; j = 1, 2, \ldots, n \end{aligned}$$

where the $\hat{\varepsilon}_j$ are maximum tolerable levels. Although Haimes saw that (2.5.) could be transformed into PP, he showed no citation for goal programming. As late as 1973, Gembicki [1973] was claiming independent development of a "goal attainment" method.

Another example which uses maximum tolerable goal level is Sahin [1976]. This paper, which deals with the prediction of corporate bankruptcy, also uses minimum tolerable levels for each goal. This would be like the interval goal programming models, except that Sahin's objective function is quadratic.

The water resources work to date has been collected in a new book [Haimes, Hall, and Freedman, 1975]. For the myriad of ad hoc techniques, I direct the reader to that source. The most interesting part of that book concerns the Surrogate Worth Trade-Off Method, which uses duality theory to assess trade offs between goals.

II. 6 Modern Approaches

II.6.1 The Surrogate Worth Trade-Off Method. The Surrogate Worth Trade-Off Method [Haimes and Hall, 1974] begins with the vector maximization problem (VM) in minimizing form. It finds the minimum value of each objective function, treated alone, subject to the system of constraints. Next the problem is reformulated in ε constraint form (as in 2.5.). Here is how the maximum tolerable levels $\hat{\varepsilon}_j$ will be related to $\overline{f}_j(x)$ (this is the notation for the j^{th} objective function, treated as if it were the only objective function).

If one objective function is allowed to dominate, and the others are satisfied by being treated as constraints, there will be (n-1) Lagrange multipliers associated with the (n-1) objectives as constraints. When the Lagrange multiplier is non-zero, it means that that particular constraint does limit the optimum. It is shown that the non-zero Lagrange multipliers correspond to the non-inferior[1] solutions, and the zero multipliers are associated with the inferior solutions. Additionally, these non-zero Lagrange multipliers show us the trade-off ratios between the main objective and each of the objectives which are expressed as constraints.

---

[1]Non-inferior solutions are basically the preferred solutions; the DM is indifferent between solution (X) vectors in this set.

These trade-off ratios have the form

$$T_{ij}(x) = \frac{df_i(x)}{df_j(x)} \tag{2.6.}$$

where

$$df_i(x) = \sum_{k=1}^{n} \frac{\partial f_i(x)}{\partial x_k} \; dx_k \tag{2.7.}$$

or equivalently as

$$T_{ij}(x) = \frac{[\nabla_x f_i(x),\ dx]}{[\nabla_x f_i(x),\ dx]} \tag{2.8.}$$

where ∇ is the del operator.

Using concepts of duality, regression analysis, and Kuhn-Tucker conditions, the authors search for an operational equivalent of the T matrix, which they denote $\lambda_{ij}[A(\varepsilon_j)]$. In their published work, they prove that these trade-offs can be obtained for any noncommensurable objective functions. For instance, "let the units of $f_i(x)$ be dollars and the units of $f_j(x)$ be DO (dissolved oxygen). Then the units of λ_{ij} are dollars/DO" [Haimes and Hall, 1974, p. 619].

The decision makers are provided with λ_{ij} for one objective, given certain attained levels of other objectives. Then they are asked the question, "Is the marginal change (λ_{ij}) in the i^{th} objective function $[f_i(x)]$ worth more or less than one unit of change in the j^{th} objective function?" If the first alternative is believed to be true, then objective $f_i(x)$ should be minimal at the expense of $f_j(x)$; if not, the opposite is true" [Haimes and Hall, 1974, p. 620]. The set of λ_{ij} which are simultaneously neutral to this question leads to the desired non-inferior region. This type of interface with the DM is captured in the "Surrogate Worth Function," $W_{ij} = f(\lambda_{ij})$.

Three computational approaches for determining W_{ij} are available. They all require certain assessments from the DM. I will return to this point in Chapters III and IV, where the DM's ability to make such assessments is questioned.

II.6.2 Work of Zionts and Wallenius. In Zionts and Wallenius [1976], Wallenius and Zionts [1975], and Wallenius [1975], an eclectic approach to multicriteria optimization is developed and then tested in a laboratory setting. These authors assume that the objective functions are concave and that they are to be maximized. In addition, the constraint set is assumed to be convex. The DM does not know his own utility function explicitly, but Zionts and Wallenius assume that it is either linear or, more generally, a concave function of the objective functions.

Laboratory experimentation and the intuition of the authors both indicated that European managers prefer to assess their trade-off functions via concrete choices. Instead of setting weights on goals, Zionts and Wallenius propose "trades" to the DM; these will be described below.

Using multidimensional assessment strategies that are reviewed in Raiffa [1969], the original objectives are modified in their arguments so that the condition known as separable additive utility prevails.[3] For example, the profits of a firm may be an obvious objective; but it may be regarded as a nonlinear function of decision variables. The modified objective function might turn out to be a linear function of the square root of profits, or the natural logarithm of profits. This type of modification is common in statistics, where different functional forms are changed so that they can represent data on linear regression equations.

In the 1976 paper, Zionts and Wallenius begin by presenting a "naive" method, which contains some of the ideas and motivations of their actual methods.

The problem is denoted as:

$$\text{Maximize:} \quad \sum_{i=1}^{p} \lambda_i u_i \qquad (2.9.)$$

$$\text{Subject to} \quad Ax = b$$

$$u_i \leqq f_i^j x, \; j = 1, 2, \ldots, p$$

$$Eu - Dx \leqq 0$$

$$u \geqq k$$

$$x \geqq 0$$

where A, x, and b are properly dimensioned matrices and vectors; $\lambda_i \geqq \varepsilon$, ε is a sufficiently small positive number, $\sum_{i=1}^{p} \lambda_i = 1$; and

---

[3] A vector-valued function $f(x)$, $x = [x_1, x_2, \ldots, x_n]$ that can be written in the form $f(x) = f_1(x_1) + f_2(x_2) + \ldots f_n(x_n)$ is "separable additive."

each objective function is of the form $u_i = f_i(x)$, i = 1, 2, . . . p. Zionts and Wallenius approximate this function by appending a set of constraints

$$u_i \leqq f_i^j \text{ , } j = 1, 2, \ldots, p(i) \tag{2.10}$$

where f_i^j is a row vector of appropriate order and u_i is the value of the i^{th} objective function, $f_i(x)$. This can be summarized in matrix notation as

$$Eu - Dx \leqq 0 \tag{2.11.}$$

where E = I, the identity matrix, for the case of linear objective functions, and D is a matrix of the objective function coefficients. If the set $\{f_i(x)\}$ is nonlinear, then E becomes a transposed permutation matrix. This model also incorporates lower bound, on objectives, that is

$$u \geqq h \tag{2.12.}$$

where h is a figure named by the DM or the analyst.

It is known from previously cited work [Philip, 1972; Zeleny, 1974 (a)] that there is a (non-unique) efficient point for any set of weights λ_i. An arbitrary set of weights such that $\sum_i \lambda_i = 1$ is chosen.

This solves the problem of (2.9.) ; as a matter of fact, the solution is an efficient one. It is not possible, holding the λ_i

constant, to improve one objective without devaluing another.[3]

Next the nonbasic variables in the latest tableau are examined. The authors have a way of classifying them into two sets. If the introduction of a nonbasic variable leads to an efficient adjacent extreme point, then it is called an "efficient variable." Otherwise it is called an inefficient variable.

From the theory of linear programming we know that the introduction of a nonbasic variable decreases the objective function. Zionts and Wallenius use this as a way of helping the DM quantify trade-offs between objectives. For each variable that is efficient, according to the definition above, the DM faces this type of questioning:

> Here is a trade. Are you willing to accept a decrease in objective function u_1 of W_{1j}, a decrease in objective function u_2 of W_{wj}, . . . and a decrease in objective function u_p of W_{pj}. Respond yes, no, or indifferent to the trade. [Zionts and Wallenius, 1976, p. 658][4]

Note that W_{ij} is the decrease in the i^{th} objective function due to making x_j basic. Solving a series of LP subproblems, the method constructs the set of λ_i that best approximates the solution to the original vector maximization problem. In my opinion, of all the methods described above, that of Zionts and Wallenius does the best job of asking questions that the DM can handle; it elicits quantitative information without going beyond the information processing ability of the human.

---

[3]If the reader is interested, he should review Yu and Zeleny [1975] and Philip [1972] for uniqueness and existence theorems on efficient points.

[4]The trade is based on switches of efficient variables; the DM is informed that the "decrease" may in some cases be negatively signed, i.e., positive.

These two authors have also used their methods for the case of concave, rather than linear, utility of objectives. Using a cutting plane algorithm, they also have attacked the integer multicriteria problem. They always move along an efficient hyperfrontier, from extreme point to extreme point, asking in essence, "Are new solutions better than the old ones?"

Unlike many authors, Zionts and Wallenius have tested their methods extensively with corporate managers. Most of this experience was with European corporations. In Wallenius [1975] and Wallenius and Zionts [1975, 1976], Zionts and Wallenius [1976], this field work is analyzed, and their own method is rigorously compared with the approaches of Dyer [1972] and Benayoun et al. [1971].

II.6.3 Goal programming. Goal programming (GP) will be covered only very briefly because it is so widely known, and so adequately explained in the literature. As early as 1952 GP was developed to handle regression problems where the least squares methods was inappropriate. This history is reviewed in Charnes and Cooper [1975]. Then GP made its real bow in an article on executive compensation [Charnes, Cooper, and Ferguson, 1955]. In the subsequent two decades, GP provided the technology for thousands of papers.

The main popularizer of GP has been Sang Lee [1972] whose book and articles deal with finance, marketing, school busing, accounting, media selection and many other applications. Lee has added very little to the mathematics of GP, but his eloquent educational activities have done a great deal to make American managers receptive to mathematical techniques for multiple objective situations (MOS).

GP is a clever extension of LP. The objective becomes the minimization of a functional which consists of deviations from targets. The targets are appended as additions or modifications to the RHS. The whole GP problem is formatted

$$\begin{aligned} &\text{Minimize} \quad \sum_{i \in I} (d_i^+ + d_i^-) \\ &\text{Subject to} \quad Ax \leqq b \\ &\qquad\qquad\quad Cx - Id^+ + Id^- = g \\ &\qquad\qquad\quad x, d^+, d^- \geqq 0 \end{aligned} \tag{2.13.}$$

In (2.13.), d_i^+ represents overachievement of the i^{th} goal, and d_i^- represents underachievement, C is the matrix of policy constraints, and I is the appropriately sized identity matrix. Since the simplex methods choose a basis whose columns are linearly independent, we are assured that d_i^+ and d_i^- will never appear simultaneously at a positive level. If absolute priorities are to be indicated, this can be achieved by large weights, or by constraints of the form:

$$d_i^+ \leqq d_j^+ \tag{2.14.}$$

Since the final tableau ensures the relationship $d_i^+ d_i^- = 0$, (2.14.) will ensure the absolute priority even if the objective formation assigns a high weight to d_j^+. Many further subtleties, as well as historical remarks can be found in Charnes and Cooper [1975].

Another way to model priorities is to use preemptive priority weights. This takes the following form. For example:

$$\text{Minimize} \quad z = P_1 d_1^+ + P_1 d_2^+ + P_2 d_3^- \qquad (2.15.)$$

This means that the variables d_1^+ and d_2^+ are to be minimized as much as possible before the algorithm moves on to d_3^-. The P_2 is meant to show that attention to d_3^- is of the second priority, while the first order of priority is reserved for the other two deviational variables. Hence we use the word preemptive. For more on goal programming with this weighting scheme, see the modification to the simplex algorithms in Lee [1972]. Ignizio [1976] extends these to the nonlinear and the integer case.

Chapters III and IV will make comparisons of the applicability of the two popular multicriteria paradigms, goal programming and multiobjective linear programming.

II.6.4 Multiobjective linear programming. Multiobjective linear programming developed from the theories cited in Section II.5. Zeleny [1974(a)], Yu and Zeleny[1975] and Evans and Steuer [1973] independently created variants of the revised simplex method to handle multiple objective functions. All of these authors drew heavily on existing management science theory [see in particular Philip (1972), and DaCunha and Polak (1967)]. But the creation of the multicriteria simplex methods was a great achievement; even the sizes of the tableaux are awesome.

As with linear programming it could be said that the MOLP methods are examples of brilliant bookkeeping. Masses of information are transformed and made useful by the successive tableaux. The several men associated with MOLP use different notations, but the formulation is actually identical:

$$\begin{aligned}
\text{Maximize} \quad & c_1^1 x_1 + c_2^1 x_2 + \ldots + c_n^1 x_n = z_1 \\
& c_1^2 x_1 + c_2^2 x_2 + \ldots + c_n^2 x_n = z_2 \\
& \vdots \qquad\qquad\qquad\qquad \vdots \\
& c_1^p x_1 + c_2^p x_2 + \ldots + c_n^p x_n = z_p
\end{aligned} \tag{2.16.}$$

$$\begin{aligned}
\text{Subject to:} \quad & Ax \leqq b \\
& x \geqq 0
\end{aligned}$$

This can be written succinctly as:

$$\text{Maximize} \quad Cx = z \tag{2.17.}$$

$$\begin{aligned}
\text{Subject to} \quad & Ax \leqq b \\
& x \geqq 0
\end{aligned}$$

In both (2.16.) and (2.17.), x_j, $j = 1, 2, \ldots, n$, is the decision vector; z^q, $q = 1, \ldots, s$, is the decision vector mapped into criterion space; b_j, $j = 1, 2, \ldots, m$, is the RHS of appropriate dimension; A is the technology matrix; and the c_j^i, $j = 1, 2, \ldots, n$ and $i = 1, 2, \ldots, p$, is the criterial coefficient for the j^{th} variable measured along the i^{th} criterion. The $\{c_j^i\}$ aggregate into the C matrix of (II.6.12.).

The solution of these systems results in a set of x vectors which are termed nondominated (or elsewhere in the literature efficient, noninferior, or Pareto-optimal). Formally, a point x is nondominated iff $\nexists\ \bar{x} \in X \ni C\bar{x} \geqq Cx$ and $C\bar{x} \neq Cx$. The concept originated with Pareto [1896], who was defending market economics from the attacks of Karl Marx; it was refined by Kuhn and Tucker [1951] and Geoffrion [1968]. In recent years a similar concept called <u>kernels of preference structures</u> has received some attention [see White (forthcoming)].

Since the set {x} is not singleton, neither is its image in criteria space. We have $Cx = z$, where z takes the form of a set of column vectors, with each one corresponding to a nondominated x^q. In Chapter V it will be convenient to collect the z vectors into a matrix Z.

The essentials of the multicriterion simplex method are as follows:

1) Find a basic feasible solution for one trial set of weights (cf. the parametric problem [2.4.]).
2) From that b.f.s., move on to find an efficient extreme point or terminate if one does not exist.
3) From this efficient extreme point, find all other efficient extreme points.

The computational details of the three steps have never been described in less than ten pages in the literature. I find that Yu and Zeleny [1975] is the best treatment of <u>domination structures</u> and their relationship to MOLP. Zeleny [1974 (a)] is the best description of an approach called <u>the decomposition of the parametric space</u>. This approach is appealing in theory but appalling in (computational) practice. Evens and Steuer [1973] provide the most compact exposition, but they require the most complete theoretical background in optimization theory. Both Zeleny [1974(a)] and Steuer [1974] share their tremendous investment in computer coding with the research community. The ADBASE code of Steuer is easier to implement and offers more options for algorithmic experimentation.

Duality theory for MOLP is currently following several paths. Of note are Kornbluth [1974], Duesing [1976] and Isermann [1977]. The MOLP dual requires the conception of the primal as having multiple RHS. Several b vectors would collapse to only one when weighted; these would be the objective function weight in the dual problem.

Each $x \in N^{ex}$ (the set of nondominated extreme points) is associated with a particular set of weights {ki}. With each $x \in N^{ex}$ is also associated the RHS weights, which we will call u (RHS). But many u vectors may correspond to each $x \in N^{ex}$. Each corresponds to a "pricing" of scarce resources. Since this is multiple pricing, one thinks of corporate applications such as transfer pricing schemes with subsidies. Another application would be pro forma capital budgeting where the parent company is evaluating the physical and financial budgets of several units that might be merged.

II.6.5 Recent dissertations. One of the leaders in the MCDM remarked that "the field was mined." It is no accident that the Multiple Criteria Session at the May 1978 TIMS-ORSA National Meetings is expected to be very applications-oriented. New theories are beginning to re-invent the wheel. The field is currently trying to synthesize, to package itself in a more mature way.

Seven dissertations will be reviewed here. None, with the exception of Bergstresser [1976], seems to be as important theoretically as were the seminal dissertations of Steuer [1973] and Zeleny [1974(a)].

Gomes [1976] offers a multiple criteria framework for the evaluation of forest and road capital projects. Both quantitative and qualitative measures are used to capture the environmental system and the value systems of the decision-making participants. A "project valuation matrix" with columns as indicants of the partial utility of projects is constructed. When pairwise comparisons can be made, they are massaged by an algorithm, in the hope of creating a complete order. Gomes uses a heuristic to estimate the loss of information in his methods. This approach is reminiscent of the very well known (but rarely understood) outranking relationships of B. Roy [1971, 1973, 1976].

Akyilmaz [1976] is concerned with the analysis of trade-offs in the realm of transportation. A detailed typology of these trade-off methods is offered.

Keefer [1976] uses decision analysis for resource allocation in the presence of uncertainty, multiple objectives and organizational constraints.

Lindsay [1976] analyzes alternative sewage sludge disposal systems by means of several multiobjective programming approaches.

Ryan [1976] applies goal programming to actual industrial safety and health problems. A detailed model is built up in the context of the Occupational Safety and Hazard Act (OSHA).

The problem of public investment planning in developing nations is characterized as a multiple objective problem in the dissertation of Sfeir-Younis [1976]. He uses goal programming to study project priorities and macroeconomic impact. The model also serves as a management information system. Projects are ranked on an ordinal scale by computing distances from an optimal set of goal accomplishment. (I question the advisability of the heavy utilization of goal programming and discuss this at length in Chapter IV.)

The dissertation of Bergstresser [1976] seems to be the most advanced in theory. Using domination structures like those in Yu [1974], he explores n-person games. It is shown that traditional solution concepts for single criterion n-person games in both normal and characteristic function form induce domination structures in various spaces such as the pay-off space. These structures indicate the sharing of power by the n players. The author uses balance sets and derives a necessary and sufficient condition for the core of a game to be non-empty. He develops a parametrization process to relate multicriterion games to single criterion games. This is like Zeleny's original work [1974(a)] on the decomposition of parametric space in the MOLP context.

Chapter II has introduced the literature from the mathematical programming standpoint. Chapter III will provide a literature review from the behavioral sciences.

CHAPTER III

PREVIOUS RESEARCH IN HUMAN CHOICE THEORY AND INFORMATION PROCESSING

III.1 Introduction

Classical economics provided the behavioral paradigms that led to the applications of optimization theory in a management context. *Homo economicus* was assumed to be a single criterion maximizer who knew what was available and who knew what he liked. Thus management science evolved as an *information-assuming* discipline. This chapter follows the tradition of Zeleny [1974(a)], and Keeney and Raiffa [1976] in examining the *information-seeking* aspects of management science. The models reviewed here are less tractable than those now implicit in management science, but in this very lack of simplicity lie the seeds of a fruitful direction for research.

III.2 Review of Human Choice Theory

III.2.1 Thurstone. In 1927 Thurstone [1927(a), 1927(b)] developed a mathematical model for describing the way an organism responds to stimuli. He felt that the task of psychophysics was to develop scales or ranges of numbers which represent the subjects' responses to psychological continua such as heat, light or sound. The scale should be constructed so that it captures the *discriminal process* along each dimension of interest.

The discriminal process is random, in the sense that the organism does not always make the same response to the same stimulus.

The goal is to be able to observe an organism and record its responses as points on a scale. It would be nice if each point had an unambiguous meaning. This is usually not possible since the responses do not come out as a metric function.

Thurstone's law of comparative judgment tries to derive a scale from a set of subject responses to paired comparisons among stimuli. We seek a scale denoted by s. Let s_j and s_k correspond to the scale values of two stimuli j and k. Also necessary are σ_j and σ_k, the standard deviations associated with a given stimulus-response set. When two stimuli are presented to the subject, two "discriminal processes," d_j and d_k, occur. The quantity $(d_k - d_j)$ is called a "discriminal difference."

The two stimuli are repeated many times, and Thurstone <u>postulated</u> that the discriminal differences $(d_k - d_j)$ would form a normal distribution on the psychological continuum. The expected value of this probability distribution is equal to the difference between s_j and s_k.

From statistical theory, we know that

$$\sigma_{d_k - d_j} = (\sigma_j^2 + \sigma_k^2 - 2r_{jk}\sigma_j\sigma_k)^{\frac{1}{2}} \qquad (3.1.)$$

in which r_{jk} is the sample correlation coefficient between the values of the discriminal processes associated with s_j and s_k (over some number of repeated trials).

During the presentations of the stimuli, the subject experiences stimuli s_j and s_k and indicates a "comparative judgment" such as softer, louder, brighter, etc.

The number of times that s_j is judged greater than s_k is useful information in creating a scale; some standardization is proposed by Thurstone. He uses x_{jk} to mean the number of times that $s_j > s_k$, measured in $\sigma_{d_k - d_j}$ units. This transformation allows the difference $s_k - s_j$ to be determined from a table of the area under a normal curve.

After a few steps of this sort, Thurstone wrote

$$s_k - s_j = x_{jk}\sigma_{d_k - d_j} \quad (3.2.)$$

Substituting equation (3.1.) we obtain

$$s_k - s_j = x_{jk}(\sigma_j^2 + \sigma_k^2 - 2r_{jk}\sigma_j\sigma_k)^{\frac{1}{2}} \quad (3.3.)$$

This was called the "complete form of the law of comparative judgment." The goal was to achieve scaling as the solution to a simultaneous system of equations.

But, as is shown in studious detail by Thurstone, the "complete" form of the law yields more unknowns than equations. To make a <u>solvable</u> system of equations, nine different sets of simplifying assumptions are used to modify the complete law. For example, Thurstone imposed constant covariance, equal correlations or homoscedasticity of discriminal differences.

He then proceeded to a more subtle version of his general judgment paradigm and called it the "law of categorical judgment." This law begins by assuming that each attribute under examination is associated with a psychological continuum divided into categories. The subject is required to place the stimuli in a particular category, or area of the continuum. In effect, this model of human information processing treats the boundaries of the categories as additional stimuli. As was the case of the complete law, a system of equations results which can be solved only when some special conditions are superimposed.

Thurstone's two laws are combined into a theory of discriminal processes. Together the laws are known as the "general judgment" model and they imply a set of testable hypotheses. The empirical work was carried out, and later researchers were able to use these beginnings to analyze the choices that organisms make.

III.2.2 Luce. Luce felt that it is very difficult to study individual choice behavior apart from the discriminal processes that interested Thurstone. He wrote that psychologists had searched for lawfulness between stimuli and responses [1959, p. 1]. His suggestion was to model choice behavior, whether the organism happened to be choosing among stimuli or among responses.

Utility theories from economics were algebraic in nature. They assumed perfect discrimination of cues, and tried to explain the complexity of response by increasingly subtle functional forms for predicting the alternatives chosen by the human subject. Luce felt that by idealizing discrimination, utility theory became needlessly complex and almost disheartening. He believed that the richness of the response set did not necessarily find its most parsimonious representation by algebraic utility functions, but instead by the inclusion of a probabilistic model of the perception of stimuli.

In his book, Luce tried to use a single axiom that related "the various probabilities of choices from different finite sets of alternatives." Thus, by exploiting whatever lawfulness exists between choice situations, whether these are choices among responses or among stimuli, he was able to extend his famous Axiom of Choice to psychophysics, utility theory and learning theory. Luce had restricted his research (as of 1959) to situations where there existed a well-defined finite set X, whose elements corresponded to crisp alternative choices.

The axiom defines T as a finite subset of a domain of discourse U (possibly infinite), and for every set $S \subset T$, P_S is defined. The notation P_S is a general form of more specific expressions such as $P(x,y)$, the probability that x will be chosen over y from the set $S = \{x,y\}$. Luce's Axiom of Choice is

(i) If $P(x,y) \neq 0,1$ for all $x,y \in T$,

then for $R \subset S \subset T$

$$P_T(R) = P_S(R)P_T(S) \tag{3.4.}$$

(ii) If $P(x,y) = 0$ for some $x,y \in T$,

then for every $S \subset T$,

$$P_T(S) = P_{T-\{x\}}(S - \{x\}).$$

Part (i) is a rather technical condition that has to do with the division of a large set of choices into a sequence of simpler choices from subsets of the original set.

Part (ii) means that if y is always chosen in preference to x, then x may be excluded from T whenever choices from T are under analysis. As Luce says, if a person never selects liver in preference to roast beef, then when the choice between liver, roast beef and chicken is presented, the decision maker should (or does) deal with the smaller problem of roast beef versus chicken.

The beauty of Luce's axiom is that it implies a condition called "independence from irrelevant alternatives." Formally, if

$$P(x,y) \neq 0,1 \quad \text{for all } x,y \in T,$$

then the axiom of choice (3.4.) implies that for any $S \subset T$ such that $x,y \in S$,

$$\frac{P(x,y)}{P(y,x)} = \frac{P_S(x)}{P_S(y)} \tag{3.5.}$$

At first glance this seems to mean that the addition of new elements to a choice set does not change the original choice probabilities. For example, if a certain woman probably would choose a blond man rather than a brown-haired man, then her choice process would not be upset by the addition of a red-haired man to the set of available men. Early applications of Luce's theories were unclear about this point. (3.5.) implies that the ratio of probabilities is unaffected by changes in the feasible choice element, if the changes are truly "irrelevant." In the above example, P(blond, brown) might change with the arrival of the redhead, but the ratio of P(blond, brown) to P(brown, blond) would not change.

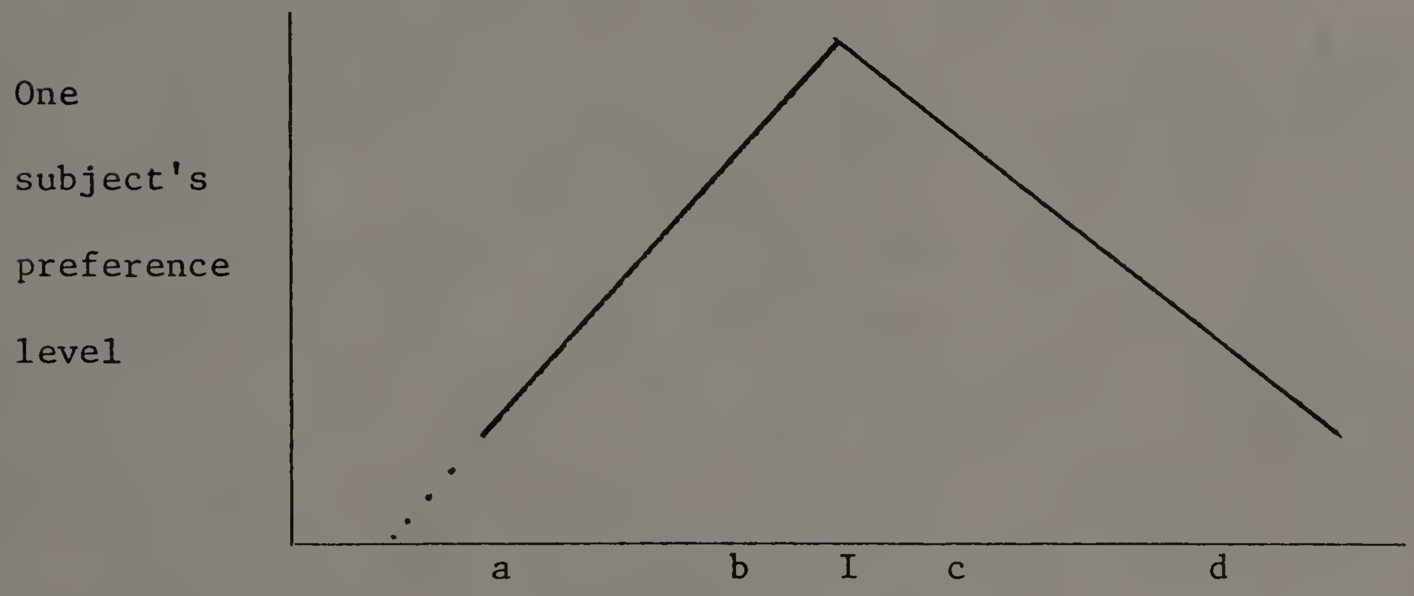
Luce's work was truly seminal, but later researchers became dissatisfied with the implications of (3.5.). Richer models of human choice deal rather intensely with the possibility that new choice alternatives are not truly irrelevant, and that they do affect the ratios of the original choice probabilities.

III.2.3 Coombs. In A Theory of Data (1964) Coombs attempted to unify the existing ideas of choice theory, data types and statistical analysis. He wrote of a theory of unfolding--a way to elicit scales for individual attributes from joint scales of responses to multi-attributed objects. This book categorized the kinds of data that are generated in empirical mathematical psychology. Coombs' data became widely accepted in the fields of mathematical psychology, social psychology and marketing research. Finally a set of techniques was developed to analyze the data. Prominent in this set was metric and non-metric factor analysis, and multidimensional unfolding.

Unfolding is most easily visualized in the one-dimensional case. Several subjects are presented with a stimulus. They are presumed to have the ability to specify an ideal point, which is what they consider an optimal point on the stimulus continuum. Each subject reacts to discrete points on the continuum. The distance of each point from the ideal point varies inversely with the subject's preference for that point. For example, in Figure 1, point b is preferred to point a because it is closer to I, the subject's ideal point. Point d possesses "more" of the stimulus, but it is not preferable to point b, because it is farther from I.

Figure 3.1

A Single-Peaked Preference Function

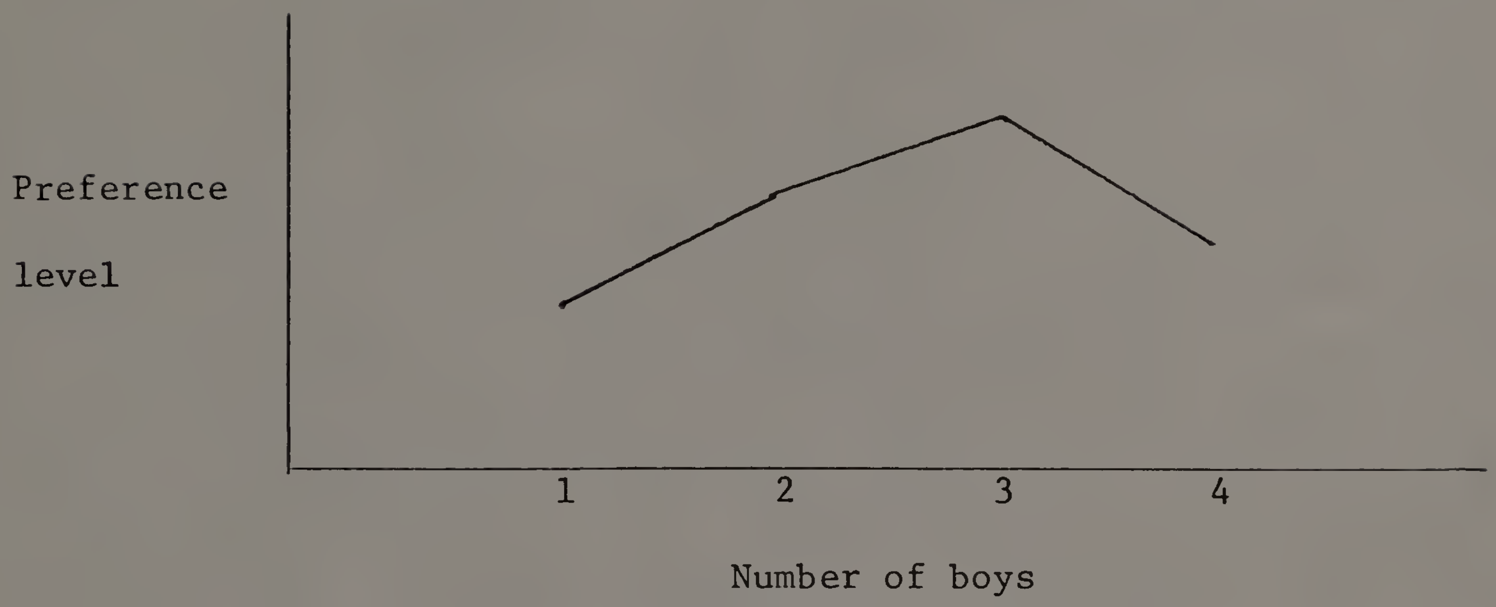


Coombs' studies emphasized that if each subject had a different ideal point I, then each would have a different rank-ordering of the stimulus objects. So if we were to <u>fold</u> the stimulus scale of Figure 3.2 at I, we would create a graph of a preference ordering on a, b, c and d for one subject.

Suppose that instead of beginning with the stimulus scale, a researcher had preference scales available. By <u>unfolding</u> these scales at the ideal point I, the underlying stimulus scales could be derived. The multidimensional analog of this unfolding has been developed and is known as multidimensional scaling. It must be remembered though, that the subjects must have single-peaked preference functions on the stimulus of interest. In other words, they do not follow the economist's assumption that "more is better."

Figure 3.2

An SPF for the Number of Boys

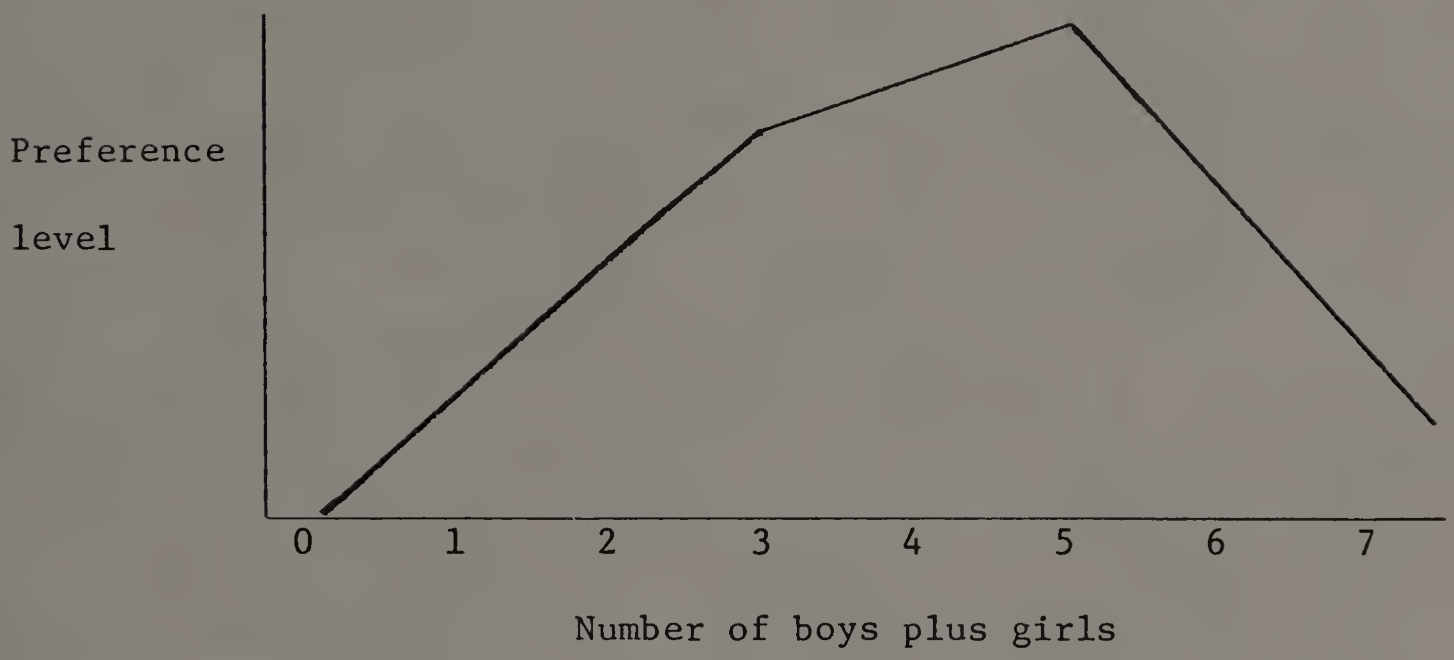


This chapter has been treating how humans react to sets of cues and sets of choices. Coombs uses the concept of single-peaked preference functions, which he attributed originally to Black [1948(a), 1948(b)]. The single-peaked preference function (SPF) is unimodal and need not be symmetic (see Figure 3.1).

Note that the function in Figure 3.2 achieves a unique maximum at three boys. Assume that the context is research into family formation preferences [Coombs, 1975]. Why would the preference in boys drop off so decisively? Perhaps three is considered an optimum family size by the subject. Alternately, the desire to have daughters may become stronger, once boys have been born. Coombs and his wife also elicited, in various countries, SPFs for girls. Then they derived SPFs for both the sum of boys plus girls, and the difference of boys minus girls (see Figure 3.3).

Figure 3.3

An SPF for Total Number of Children



The attractiveness of this approach is that it provides a psychological theory for non-monotone preferences. By considering an observed scale to be the joint effects of several underlying scales, one can "unfold" those scales for closer verification.

Many researchers have proposed choice models where various attributes were weighted and added to form a joint scale. Coombs explains very cogently how these models implicitly assume that there is no interaction among the attributes. Since there are many worldly phenomena where interaction does occur, it is very important to pursue research that incorporates this interdependence.

To be thorough in his review of multidimensional choice models, Coombs reviews (in Chapter 9) several other concepts. One is called the <u>personal compensatory model</u>. The formal definition of this is

$$P = \Big|\, |c| - \text{proj } q \,\Big| \qquad (3.6.)$$

where $|c|$ is the norm or length of the vector from the origin to the <u>ideal point</u>, proj q is the projection of the stimulus vector in the direction of the individual's vector c, and P is the performance measure. Although this notation is unfamiliar to management scientists, this is precisely the kind of model implied by the objective function of a goal programming problem which does not make the weights preemptive.

Another idea that Coombs reviews is <u>lexicographic ordering</u>. Of several objects in the choice set, the one chosen will be that which has the highest level of the most influential attribute. The decision maker disregards the information inherent in the levels of other less important dimensions of the objects of choice. This is known in the optimization literature as goal programming with preemptive priority weights. It is also allied to another MS/OR technique called lexicographic programming. However, Harrald, et al. [1975] show how this lexicographic model is inconsistent with the existence of a utility function in the mind of the decision maker.

The next model discussed, the <u>city block model</u>, is an additive, compensatory model which assumes no interaction. There is a peculiar form to the additivity of dimensions. The distance between choice elements is defined as the arithmetic sum of their differences along each dimension. This model has appeared independently in several fields, but its applicability is hard to prove.

Coombs [1964, p. 206] states that "one of the most desirable consequences of developing alternative models and their (associated) algorithms for data analysis is that their existence destroys any naive complacency with any one model and leads to a search for ways of testing and comparing alternative theories."

This entire section from Coombs has dealt with first-choice preferences. A very different choice behavior may result if the subject is asked to pick his second-best choice. This subtle point has been raised again, over a decade after Coombs' research, by Zeleny [1976(c)] and by Farquhar [1977].

One other issue from A Theory of Data deserves some attention. In Chapter 13 Coombs says that the real world is full of choice sets that include incomparable (noncomparable) elements. Social scientists, he feels, are sometimes criticized for building multidimensional utility functions that crunch these elements together incorrectly, invalidly or haphazardly. But societies do consummate decisions. Since the compression of partial orders into complete orders is often made by a small subset of the population, it is especially important for scientists to improve our methodologies for making these choices.

This point is reminiscent of Shepard [1964] and Zeleny [1976(c)]. Both review experiments that show the limited information processing ability of the human. The arrival at a decision from among a set of multiattributed alternatives implies that a set of weights has been assigned to the attributes. But that set may be the transient weights of a particular frame of mind. Here is a source of subjective nonoptimality (or satisficing) in individual decisions. A spurious resolution of a conflict situation may be adopted, which, although untenable in the long run, will allow one alternative to be chosen. This is the individual analog of the problem Coombs [1964, Ch. 13] cites from political economy. These grave but chancy situations invite the concern of the social science community. At least let us dignify the choice processes, if we cannot improve them, by demonstrating that they are personal and organizational research processes that provoke learning.

III.2.4 Tversky. Of the many contributions of Tversky to various fields of psychology and utility theory, we will restrict ourselves to his elimination-by-aspects model (EBA). EBA resembles somewhat lexicographic utility and lexicographic programming. They all share a preemptive viewpoint on choice; one goal, or stimulus dimension somehow dominates all others. The lexicographic (often called conjunctive in the field of marketing research) models are deterministic, whereas the dominance of an aspect is a random event in elimination models discussed here.

EBA may also remind some of linear programming in the way it "sweeps out" the set of feasible choices. Tversky added philosophical and experimental rigor to this type of algorithm. Instead of just listing EBA as a human choice model, he developed a probabilistic behavioral theory of human choice. In his paper entitled "Choice by Elimination," Tversky [1972(b)] analyzed the human who is faced with a set of choices as a stochastic elimination operator.

Individual choice behavior has been observed for centuries in art and literature, and for decades by scientific researchers. Apparently, choosing is often inconsistent and random. This could be due to the "unreliability" of human desires, poor experimental designs, or inappropriate models.

To improve the fit of models to experimental results, probabilistic concepts have been explicitly introduced. There are two important recent representatives, the random utility models and the constant utility models. In random utility models the utility or value of each alternative is subject to random fluctuations; the one with the highest value at the moment of choice is selected. In models of constant utility, the scale values associated with each alternative are not random variables. But the choice rule (function) is random. The two basic models differ only in the point at which the stochastic nature of choice is permitted to enter the model.

A third probabilistic viewpoint is proposed by Tversky--the <u>elimination</u> models. These view choice as a multi-stage process, a sequence of eliminations. It is a structure which may prove useful to management science (cf. also Fishburn [1974(b)] and Zeleny [1975, 1976]).

To motivate EBA, Tversky reviews concepts such as simple <u>scalability</u>, <u>order independence</u> and <u>independent random utility models</u>. Then he shows, either by mathematical reasoning or by citing experimental results, that none of the three holds in general.

The formal treatment in Tversky [1972(a)] is highly mathematical.

Even the anecdotal examples used are very subtle and serve to illustrate fine points. But in Tversky [1972(b), p. 297], in addition to reporting some laboratory tests of EBA, the following quotation illustrates that the author's paradigm of covert elimination is easy to grasp.

> The following television commercial serves to introduce the problem. "There are more than two dozen companies in the San Francisco area which offer training in computer programming." The announcer puts some two dozen eggs and one walnut on the table to represent the alternatives, and continues: "Let us examine the facts. How many of these schools have on-line computer facilities for training?" The announcer removes several eggs. "How many of these schools have services that would help you find a job?" The announcer removes some more eggs. "How many of these schools are approved for veteran's benefits?" This continues until the walnut alone remains. The announcer cracks the nutshell, which reveals the name of the company and concludes: "This is all you need to know in a nutshell."

To refine this exposition, notation is introduced [1972(a), p. 346]. Given an offered set A, one chooses a (nonempty) subset of A, say B, with probability $Q_A(B)$. Now, making B the offered set, $C \subseteq B$ is selected with probability $Q_B(C)$, and so on until the selected subset is a singleton set. One alternative has been chosen.

Tversky interprets the transition probability $Q_A(B)$ as the probability of eliminating from A all alternatives that are not included in $B \varepsilon A$. At each of several stages, or states, the subject divides a set of alternatives into "acceptable" and "unacceptable" sets. The relative frequencies of these segmentations could serve as estimates, if one performed a designed experiment, of the transition probabilities. It has been suggested to me informally that, in the context of the prediction of choice behavior, these transition probabilities are non-stationary and inestimable. But, in the prescriptive or normative world of multicriterion optimization, these probabilities are evident (or at least implicit) in the actual behavior of the decision maker, under the EBA hypothesis. Thus, the choice probabilities of the corresponding Markov chain are:

$$P(x,A) = \sum_{B_i \subseteq A} Q_A(B_i)P(x,B_i) \qquad (3.7.)$$

To introduce a theory of choice which is open to experimental validation, an assumption is introduced. A set of transition probabilities satisfies <u>proportionality</u> whenever

$$\frac{Q_A(B)}{Q_A(C)} = \sum_{B_j \cap A=B} Q_T(B_j) \Big/ \sum_{C_i \cap A=C} Q_T(C_i) \qquad (3.8.)$$

when both denominators are constrained to be positive. But if the denominator on the left vanishes, so must the one to the right of the equal signs. A theorem follows:

Main EBA Theorem: Given the general elimination model described above, the proportionality condition holds if there exists a scale U on 2^T (the set of all subsets of T) such that for any $X \varepsilon A \subseteq T$

$$P(x,A) = \sum_{B_j} U(B_j)P(x,A\ B_j) / \sum_{\substack{A_k \\ A\neq 0}} U(A_k) \qquad (3.9.)$$

A proof and some extensions follow in Tversky's paper.

The EBA Theorem treats a set of multiattributed elements of a nonempty set. They are alternatives offered for choice; each alternative consists of several <u>aspects</u> or <u>components</u>. At each Markov transition point one chooses an aspect from those that are evident in the choice set. Let us say that each aspect is ranked in importance by means of a weight (overt or covert); then the probability of an aspect being chosen is proportional to its weight.

Once an aspect surfaces to dominate, all the alternatives that do not include that particular aspect are eliminated. Successively, in this way, alternatives are placed into "acceptable" and "unacceptable" categories, until only one remains. Note that the selection (or more precisely, elimination) criteria are aspects, and the order in which the attributes or aspects surface vary from one time to another. Zeleny [1974 (a), p. 170] states that the weights are "transitory, and not in the possession of the decision maker." Tversky in a slightly different context agrees.

There is more to be said on weights. For $A \subseteq T$, let U(A) be the sum of the weights of all the aspects that are to be found in all the alternatives of A. In addition, no weight is associated with any aspect which exists in an alternative <u>not</u> included in A. Therefore, U(A) is a measure of the unique advantage of those alternatives in A. It should be noted then that U(A) is not the same as "utility" as commonly defined in management science.

This (by now) classical type of utility has not typically dealt with interdependence of aspects. Fishburn [1964] and Keeney and Raiffa [1976] assume that dependence can be eliminated. Farquhar [1974(a)] does deal specifically with dependent aspects in a utility context. But Tversky is mainly responsible for the careful treatment of the decision situation where there is no dependence among aspects but where there is a <u>structural dependence</u> among alternatives.

To talk of structural dependence, we define <u>general scalability</u> as a condition which holds iff $\exists$ a scale U defined on 2^T (as defined before, the set of subsets of T) and functions F_n, $2 \leq n \leq t$, such that for all $x \in A \subset T$

$$P(x,A) = F_a(U[\pi(x,A,1)], U[\pi(x,A,2)], \ldots, U[\pi(x,A,2^t]), \qquad (3.10.)$$

in which a and t denote, respectively, the cardinality of A and T. Also, $\pi(x,A) = [\pi(x,A,1), \pi(x,A,2), \ldots, \pi(x,A,2^t)]$ is a permutation of 2^T. Still following Tversky's treatment of general scalability [1972(a), p. 363] it is assumed that for choice probabilities P(x,A) that are not 0,1

(i) F_a increases in each argument $U[\pi(x,A,i)]$, $i = 1, \ldots, 2^t$, such that $A \cap \pi(x,A,i)$ is $\{x\}$, and

(ii) F_a decreases in each argument such that $A \cap \pi(x,A,i)$ is nonempty and does not include $\{x\}$, and

(iii) F_a is constant in each argument such that $A \cap \pi(x,A,i)$ is either empty or equal to A.

In other words, general scalability is an attempt to explain human choices that seem to depend on the composition of the feasible set of alternatives. Using the example I constructed when reviewing the work of Luce [1959], if "a certain woman would choose a blond man rather than a brown-haired man" for some purpose, then that preference might be upset by the addition of a red-haired man to the choice set. Again this is reminiscent of Zeleny [1974, 1975, 1976] who talks of how weights depend on the particular composition of the available alternatives.

What this means is the disavowal of Luce's condition of independence from irrelevant alternatives. Choice probabilities do, I believe strongly, depend on "irrelevant" but available alternatives (cf. Zeleny [1974(a), 1975, 1976] on ideal points). Our preferences change; we actually relearn them, to some extent, when the choice set is altered.

Specifically, F_a increases with the value of aspects that belong to x and to no other alternative of A. F_a decreases with the value of aspects that belong to some alternative of A but not to x. Also, F_a is independent of aspects that do not belong to any of the aspects of A, or belong to all the alternatives of A (cf. Zeleny [1974, p. 177] on weights as measures of the information that a criterion generates).

EBA modeling of human choice includes the condition called scalability. Bluntly, it is implied that the addition of an alternative to a set of choices "hurts" alternatives that are similar to the one that was added more than those that are less similar to it. Tversky [1972(b), p. 283] cites empirical validation of this phenomenon.

When talking of multicriterion optimization, we are dealing with <u>normative</u> models. But the above <u>descriptive</u> models have been explicated because the chance of a model's acceptance increases with its intuitive appeal.

If a mathematical programming approach asks for input from the decision maker in a natural way, the whole interactive process between man and machine (or DM and analyst) is likely to be improved. Thus the discussion of EBA models versus weighting models is relevant at this point.

When the DM is choosing among a complicated set of multiattributed objects, a vast amount of information must be accommodated. True optimization methods require that weights be assigned to the different attributes, and that these weights be combined by some mathematical aggregation function. Another technique is to assess trade-off rates among the attributes (see, for example, Geoffrion [1968], Geoffrion, Dyer and Feinberg [1972], Haimes and Hall [1974] and Zionts and Wallenius [1976]).

Human choice in large multiattribute or multicriterion situations is probably going to develop using EBA models or paired comparison models. These will fall more naturally within the limits of human perceptual and cognitive ability than do the weighting models.

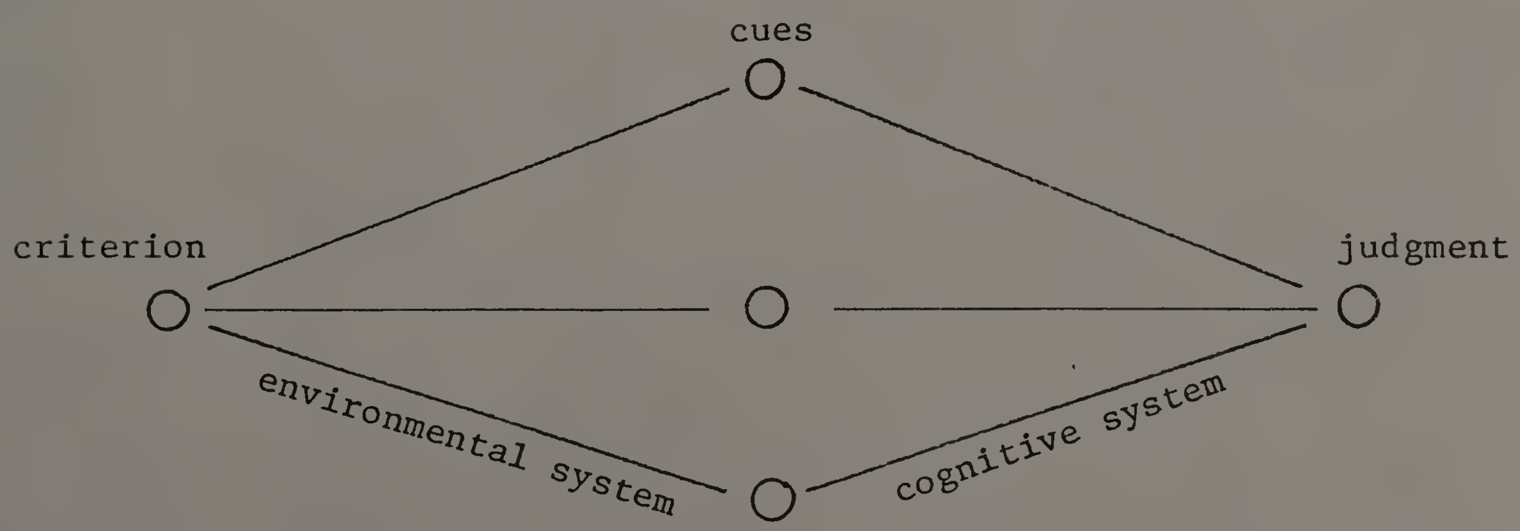
The above studies of mathematical psychology are in some danger of being outdated by the process-oriented models of artificial intelligence. It is too early to see whether this will be the case. Since the methods developed in this dissertation are really a hybrid algorithm of searching and branching processes, they are akin in spirit to artificial intelligence. Arbib [1972] provides a good sense of this field. The next section covers some of what we know about how the human mind processes information.

III.3 Review of the Difficulties of Human Information Processing

III.3.1 The Brunswik Lens Model. The Brunswik Lens Model will be surveyed as a view of the environment in which humans process information. This treatment follows Rappoport and Summers [1973, p. 16 ff.]. (See Figure 3.4.)

Figure 3.4

Brunswik Lens Model



When the decision process is represented by a structural or process scheme, it is defined as consisting of a list of variables or attributes (morphology), and a list of functional relations among them (mechanisms). The process's morphology is divided into two classes of variables: external and internal to the decision maker.

Another way to say this is that a cognitive system is any minimally organized set of relationships between an individual's judgments and the information ("cues") on which judgments are based. In connection with judgment phenomena, cognitive systems can be thought of as policies. That is, to the extent that an individual finds meaning in a body of uncertain information, he does so through application of an implicit or explicit policy concerning the causal relationships indicated by the information and the relation of his outstanding goals or purposes to that information. Technically, policy may be seen as a set of rules for using available evidence to reach a decision in an uncertain situation. The environmental system is the situation itself.

The Brunswik Lens Model provides a useful framework for describing the judgment process. It allows us to see that environmental systems consist essentially of an object, event or variable, i.e., a criterion, that is presented to cognitive systems as an array of cues. The criterion usually generates multiple cue information, and this is denoted by showing three cues, but the number can be larger or smaller. The lines running from the criterion to the cues denote the relationship of each cue to the criterion. Virtually all environmental systems to which man tries to adapt are uncertain, rather than fully determined. Thus man rarely encounters situations in which cues are perfectly related to a criterion.

Similarly, the cognitive system depicted by the lens scheme consists essentially of a set of relationships between judgments about the criterion and the multiple probabilistic cues upon which these judgments are based. The lines connecting the cues to the judgment denote the relationship between each cue and the individual's judgments. These are also not fully determined.

Three specific advantages of the lens model methodology are: 1) Any environmental system may be described by assessing cue-criterion relationships (environmental "cue validities") and overall uncertainty (variance in the criterion not accounted for by all the cues). 2) Any cognitive system can be described by determining the manner in which each cue is used by the individual ("functional cue validities") as well as by the uncertainty in the individual's policy (variance in the policy not accounted for by all the cues). 3) The general efficiency or utility of any policy can be described by establishing the accuracy of the judgments it produces. The efficiency of that policy can be explicated by comparing functional cue validities with environmental cue validities.

<u>III.3.2 Examples of problems with cues and judgments</u>. This early reference is useful to give a sense of the complex nature of human information processing. Hermann von Helmholtz (in Boring [1950] p. 311)

dealt with a theory of perception. This was that "the bare sensory pattern, as directly dependent upon the stimulus object, was called a Perzeption. A pure Perzeption is comparatively rare; it is nearly always supplemented and modified by an imaginal increment dependent upon memory and induced by unconscious inference." Von Helmholtz believes that the key to perception lies in the "Anschaung" which involves both sensation and imagery, both stimulation and unconscious inference. Von Helmholtz concluded, "Thus the properties of objects are merely their effects upon our senses, the relations of the objects to the organs of sense."

The best known reference on the information overload problem is Miller [1956, 1965, and 1967]. He reviewed a multitude of experiments that showed the human brain capable of holding only about seven symbols active in short-term memory. He corroborated this number seven by citing its frequent appearance in the literature and customs of many cultures. For example, the week has seven days and (before area codes) telephone numbers were seven digits long.

Miller said that most people can recode primitive symbols to increase their span of attention. For example, one could use A as the symbol for all states that begin with that letter. But, in general, to assess either weights or trade-offs, vast numbers of symbols must be simultaneously considered. Miller's work leads to a cautious view of our ability to do this--hence the oft-reported instability of weights, and the sense of frustration as the trade-off methods stop converging. Perhaps our confidence is highest when weights or trade-off rates are used ex post facto to justify or explain a decision that really resulted from a far more complex compensatory model.

Fixation (Scheerer [1963]) is often an obstacle to problem solving. Fixation is caused by an inaccurate perception of the type of solution needed to solve the problem; often the wrong assumptions are made. Learning to use familiar objects in an unusual way is often impeded by fixation. Over-motivation, such as strong ego-involvement, can also hinder solution. Insight occurs and may overcome fixation when one suddenly perceives the problem or its methods of solution in a different way. An important lesson of pivoting in mathematical programming is that problem solving can be viewed as successively changing the representation of a problem. Certain transforms cause fixation to subside, since they stimulate new assumptions and new trial solutions.

In Bass, Pessemier, Teach, Talarzyk [1969] it was found that brand preference is related to the attributes of brands as perceived by individuals. Another intuitive result was that in some cases there was an additive weighting structure on the attributes. Bass et al., "policed," as they put it, for non-additivity due to interaction affects and noise. When it became necessary, they replaced the derived scales of subjects by market share data. This is reminiscent of Samuelson's theory of revealed preference.

Mason and Moskowitz [1972] found that human subjects are more conservative than "reconstructed logics," such as Bayesian statistics. Conservatism refers to revising estimates in the direction indicated by the data, but not as far in that direction as Bayes's Law would indicate. Conservatism was also found to be an increasing function of "quality" and "primacy." Quality refers to the informativeness (or surprise value) of the data. Primacy is a word used to describe data to which more attention is paid because it comes early in a sequence.

Bass, Lehmann, and Pessemier [1972] support the Fishbein attitude model which says that attitude is a function of one's perception and evaluation of attributes in an object. Despite constant attitudes and preferences, brand switching occurs, probably due to the need for occasional variety. Switching to similar brands occurs more often than switching to dissimilar brands.

In Wilkie and Weinreich [1973] it is shown that increasing the number of attributes to be evaluated can be detrimental to decision making. Decision making is also affected by the type and order of attributes presented.

Wright [1974] described how time pressure caused subjects to weight negative evidence more heavily. Also as time allotted to decision tasks decreased, individuals used less of the data that was available to them.

A large scale data base on preferences has been developed in the field of mathematical nutrition and menu planning. Balintfy, et al. [1974] worked with large numbers of college students and military personnel; utility functions were assessed and used in advanced optimization models. One of the most interesting results was the discovery and quantification of <u>time-preference functions</u>. The utility of a particular food was found to be a decreasing function of its frequency of consumption.

Scott and Wright [1976] attempted to evaluate weights used by buyers in evaluating products through multiple regression. It was found that the subjects' self-reported weights differed from the weights found through regression. Three or fewer cues accounted for most of the explained variance in the decisions. Increasing the number of cues to six caused the decision maker's weighting process to become unstable.

Gehrlein and Fishburn [1976] have developed an algorithm which exploits a subset of information available to the decision maker and demonstrates that better decisions can be made when some information is withheld. Although information overload is well known among human information processors, the authors prove that even machines suffer in this way.

Troutman and Shanteau [1976] discovered favorable information was weighted more heavily than neutral information. This is like Zeleny's [1974(a)] entropy concepts and Tversky's [1972] attention levels. According to this article, a functional form to capture this behavior is an averaging model, as opposed to an additive model.

In "Funes the Memorious," the Argentinian writer Jorge Luis Borges describes the alien situation of a boy who is very good at a certain kind of information processing. The boy is literally incapable of forgetting anything. He says, "I have more memories in myself alone than all men have had since the world was a world" [Borges, 1962]. Forgetting trivia is also important.

There is an overlap here with a concept from psychoanalysis. Although the limits on information processing may at first seem to impede multicriterion optimization, perhaps they aid it by preserving sanity and rationality. Becker [1973, p. 178] talks of an idea that appears in Freud, Rank, and Kierkegaard. It is alternately called <u>partialization</u> or <u>fetishization</u>. To avoid being "crippled for action," man "narrows his world, shutting off experience, developing an obliviousness both to the terrors of the world and to his own anxieties." This repression, a small part of which the above literature search describes, prepares the mind for the decision making stage.

III.4 Implications for Multicriterion Optimization

The human choice and information-processing literature concludes that choosing is very difficult. Large amounts of data aggregate poorly, cause human frustration and are associated with weights and trade-offs that are unstable over time. They also cannot be captured by separable additive utility functions.

This means that the goal programming (GP) of Charnes and Cooper [1961], which uses cardinal weights on goals, is probably at best a trial representation of the decision environment. GP reveals too little of decision space to the decision maker. If the weights chosen are wrong, then they will drive the linear equation system in an undesired direction. But the "solution" will have the appearance of validity, since it came from an optimization algorithm.

In addition, the GP of Lee [1972], which uses cardinal weights nested within ordinal weights, also underrepresents decision space. In Harrald et al. [1975, p. 13], the following insight is presented for the first time to the management science community: "Pre-emptive priority levels are equivalent to a lexicographic preference ordering of these levels. As is shown in note 2 on pages 72 and 73 of Debreu [1959] the existence of a lexicographic preference ordering is inconsistent with a utility function structure of these preferences." Thus Lee's GP suffers from the usual difficulty in weighting in addition to mathematical representation that distorts utility maximization. A related mathematical proof showing the inconsistency of GP weights is in Chapter IV of this dissertation.

To replace GP, we have multiobjective linear programming (MOLP) as developed independently by Yu and Zeleny [1975] and Evans and Steuer [1973]. In MOLP it is not necessary to provide point estimates of the weights on objectives. Of course, if the set of nondominated solutions turns out to be large, there are problems of choice to which this paper is devoted.

These methods may be embedded in a man-machine symbiosis. This was proposed by Zeleny [1975(a) and 1975(b)]. Why not let man do what he does best, and let the computer do what it does best? Man can perform a sequence of paired ordinal comparisons, and the computer can do the bookkeeping and the aggregation. The combinatorial work on the elementary human judgments can often yield a matric preference structure. When no such structure is apparent, owing to interdependent attributes, learning, etc., then computers still free man to do the kind of broad, synthetic, exploratory, information-creating reasoning in which he excels.

The DM can show varying levels of attention to different stages of the decision process. Referring back to partialization [Becker, p. 178], by a judicious sharing of tasks with the computer, man is prepared to concentrate his constrained abilities on the holistic aspects and externalities of the problem.

CHAPTER IV

PRIORITIZATION: A THEORY OF NAIVE WEIGHTS

IV.1 Introduction

Two of the most popular techniques for multiple criteria decision making (MCDM) are goal programming (GP) and multiobjective linear programming (MOLP). GP implicitly depends on a multiattribute, separable additive utility function. It moves through the decision space to an optimum by exploiting a known preference structure. MOLP encourages a search of the preference structure because the decision maker must compare several nondominated solutions. The purpose of this chapter is to help the decision maker (DM) choose between GP and MOLP. Specifically, the general applicability of GP will be questioned.

IV.2 Two Useful Results

IV.2.1 The difficulty of setting weights. If goal programming is to work, the DM must be able to set weights on goals. Considerable research in psychology indicates that the information processing capacity of the human brain is rather limited. For a survey of the choice theory literature (especially Thurstone, Luce, Coombs and Tversky) see Chapter III. That chapter also reviews studies of information overload in humans.

For example, Miller [1956] demonstrates that the brain can accommodate only seven objects simultaneously in short-term memory.

Shepard [1964] cites examples of the difficulties of weight-setting. Troutman and Shanteau [1976] experiment with the idea that the additive (compensatory) model does not capture the consumer's choices among multiattributed objects. Zeleny [1974(a), p. 170] states that weights on goals are learned throughout the choice process and are not independent of the feasible set of alternatives.

In summary, weight assessment is a difficult search/learning task. Nevertheless, more effective use of the DM's time may exist. For example, weights can be considered as outputs of a man-machine interactive model; but usually in GP they are inputs.

In MOLP, learning about the preference structure occurs while reducing the set of nondominated solutions to one. This pruning process (see Zeleny [1974] and Chapter V), which is a sort of _a posteriori_ weighting, is recommended for capturing the unstructured experience of the expert decision maker.

IV.2.2 Debreu's result. The above behavioral section relates most easily to the nonpreemptive weights used by Charnes and Cooper [1961] and Ijiri [1965]. Sang Lee [1972] and Ignizio [1976] use preemptive weights which they claim are consistent with the DM's desire to cluster his multiple goals at several priority levels. This scheme of preemptive, or nested goals is subject to all the problems discussed in section IV.2.1 above. In addition, Harrald et al. [1975, p. 13] present a relatively unknown result:

> Pre-emptive priority levels are equivalent to a lexicographic preference ordering of these levels. As is shown in note 2 on page 72 and 73 of Debreu [1959] the existence of a lexicographic preference ordering is inconsistent with utility function structure of these preferences.

Thus Lee's objective function optimizes something, but it does not optimize a utility (preference) function.

IV.3 A Theory of Naive Weights

It is difficult to predict how the coefficients of the objective function will drive a linear equation system. We have a mature theory of post-optimality analysis in linear programming (LP), but it extends at best to two coefficients at a time. In GP with non-preemptive weights, we have an LP system with the goal weights as objective function coefficients. The management science literature treats these weights as numbers from scales (cardinal or ordinal) that correspond to priorities among objectives.

Definition IV.3.1. Weights are a set of scalars k_i, $i \in I$, that can be used to drive a suitably designed multiple criteria system toward a desired state.

Remark IV.3.1.1. By "suitably designed" it is implied that the system under study can be modeled as a mathematical programming problem. Weights will appear naturally as coefficients associated with either deviations in a GP format, or objective functions in a MOLP format.

<u>Remark IV.3.1.2</u>. To associate the highest weight with the highest goal may seem logical. But does the system (or its mathematical representation) move toward its target state? Under certain conditions it may be naive to assume that it does.

<u>Definition IV.3.2</u>. <u>Naive weights</u> are a set of scalars, k_i, $i \in I$, such that the highest weight is attached to the most important of p goals, and so forth. More formally, let $\{k_i\}$ be the ranked weights and $\{c_i\}$, $i \in I$, be the ranked goals. If $\exists$ a mapping $\alpha: \{k_i\} \to \{c_i\}$ $\exists$ the indices of the ordered pairs of α are identical, then the weights $\{k_i\}$ are said to be <u>naive</u>. <u>Non-naive weights</u> are defined as weights that are not naive.

The weights $\{k_i\}$ appear in the literature of GP and MOLP. In MOLP, there are many combinations of weights that can lead to non-dominated points in decision space. The assumptions and the theorem that follow are not directed toward MOLP but only apply to GP. Particular assumptions follow.

<u>Definition IV.3.3</u>. An <u>objective</u> is an unbounded aspiration, a desired direction as measured along a criterion c_i.

<u>Definition IV.3.4</u>. A <u>goal</u> is a particular setting along the scale defined by the criterion c_i. It is an objective operationalized by a target level.

<u>Definition IV.3.5</u>. (from Tversky [1972(a), p. 296]). A <u>state of mind</u> is a transient psychological state that causes the DM to pay attention

to a set of goals or attributes in a certain order.

Definition IV.3.6. (from Zeleny [1974(a), p. 177 ff.]). Contrast intensity is a measure related to the average intrinsic information that a given set of decision vectors generates via the i^{th} criterion c_i. For example, profit might be the most important goal in some situation. But if the decision vectors under consideration all yielded the same profit, that criterion would be of no use in choosing among the alternatives. Attention would shift to another goal (or criterion), despite its lower importance.

Definition IV.3.7. A satiety point is a level of goal achievement at which the DM's attention would shift. This might be due to Tversky's state of mind ideas, to Zeleny's contrast intensity or to external strategic or cultural influences.

Now we can define the goal programming problem as

$$\text{Minimize} \quad \sum_{i=1}^{p} k_i \left(d_i^+ + d_i^-\right) \tag{4.1.}$$

$$\text{Subject to:} \quad Ax \leqq b$$

$$Cx - Id^+ + Id^- = g$$

$$x,\ d^+,\ d^- \geqq 0$$

Where A is m X n, b is m X 1, C is p X n, I is the identity matrix of order p, d^+ and d^- are p X 1, and g is p X 1. In addition, the k_i, i = 1, 2, . . . , p are non-preemptive weights on the p goals. If the preemptive quality is desired, it can be achieved by constraints of the form $d_i^+ \geqq d_j^+$. What distinguishes (4.1.) from the standard GP problem is that g, the right-hand side corresponding to the goal constraints (also known as policy constraints), is a vector of satiety points.

Typically g is vaguely defined, and the student of the model is forced to search for the intent of the model builder by examining the weights for symmetry. For example, is $k_i(d_i^+) = k_i(d_i^-)$? Lee [1972] and Ignizio [1976] stress achieving goals; in other words, deviation on the upside is penalized by the objective function, but is generally acceptable to the DM. Thus the g vector represents "enough" goal achievement, and restricting it to a vector of satiety points simply makes that practice explicit.

Problem (4.1.) lends itself to changing the weight structure and examining the solutions. Due to the satiety point concept, the word better is partially defined. Comparing two solutions, for example, along two criterial dimensions, the one with less downside deviation on the first goal is preferred even if its upside deviation is larger. Due to satiety, the DM is indifferent to overachievement of a goal. Thus, it may be possible to shift scarce "resources" from b and g by changing the weights $\{k_i\}$.[1]

---

[1]This idea came from an off-hand remark by Ralph Steuer at the Philadelphia TIMS-ORSA meetings in 1976. He said something about "burning weights," by which I think he meant using up weights on the wrong goals, i.e., ones that would be achieved even with low weights.

Weights in GP have no meaning unless they are derived from a finite scale. A weight of 10, for example, is high if 10 is the highest weight that can be placed on a goal. Usually practitioners of GP choose weights from a finite scale, such as (0-10) or (0-1). Standardizing to the (0-1) range makes it evident that the $\{k_i\}$ form a <u>convex combination</u>. Thus, if there are three goals, setting a weight of .5 on the most important goal means that only .5 is left to be shared between the second and third.

Recalling the definition of naive weights (Definition IV.3.2), a theorem can be stated.

<u>Theorem IV.3.1</u>. In the context of Problem (4.1.), naive weights $\{k_i\}$ do not uniquely generate good solutions.

<u>Proof</u>. Proof will be by construction. Problems (4.2.) and (4.3.) will be constructed and solved. Then it will be demonstrated that both naive weights and non-naive weights lead to good solutions.

<u>Problem (4.2.)</u>

$$\text{Minimize} \quad \sum_{i=1}^{4} k_i(d_i^+ + d_i^-)$$

$$\text{Subject to} \begin{bmatrix} 6 & 10 & 2 \\ 4 & 7 & 14 \\ 12 & 5 & 3 \end{bmatrix} x \begin{matrix} \le 100 \\ \le 30 \\ \le 40 \end{matrix} \tag{4.2.}$$

$$\begin{bmatrix} 12 & 4 & 8 \\ 13 & 6 & 8 \\ 1 & 8 & 2 \\ 14 & 6 & 8 \end{bmatrix} x - Id^+ + Id^- = \begin{bmatrix} 35 \\ 78 \\ 60 \\ 70 \end{bmatrix}$$

and x, d^+ and d^- are properly dimensioned vectors for a problem with p = 4 goals.

Problem (4.3.)

$$\text{Minimize} \quad \sum_{i=1}^{4} k_i(d_i^+ + d_i^-)$$

$$\text{Subject to} \quad \begin{bmatrix} 4 & 6 & 8 \\ 7 & 12 & 34 \\ 16 & 2 & 19 \\ 5 & 20 & 10 \end{bmatrix} x \leq \begin{bmatrix} 50 \\ 100 \\ 89 \\ 120 \end{bmatrix} \tag{4.3.}$$

$$\begin{bmatrix} 12 & 4 & 8 \\ 2 & -4 & 9 \\ 1 & 8 & 2 \\ 14 & 6 & 8 \end{bmatrix} x - Id^+ + Id^- = \begin{bmatrix} 32 \\ 30 \\ 60 \\ 45 \end{bmatrix}$$

and x, d^+ and d^- are properly dimensioned vectors.

Conforming to Definition IV.3.7, the g vectors (the RHS for the policy constraints) are satiety points. Therefore when the weights are set in the objective function, zeroes are assigned to the $\{d_i^+\}$, which are the deviation variables associated with over-achieving goals. The priority weights for under-achievement of goals $\{d_i^-\}$ may vary between 0 and 1. In addition, they are chosen to sum to 1.

Problem (4.2.) was run with various sets of weights. It is assumed that the four goals are listed in descending order of importance in all cases; thus the reader can readily identify both naive and non-naive weights. For $\{d_1^-, d_2^-, d_3^-$ and $d_4^-\}$ equal to, in turn

$$\begin{array}{lllll} & \{.5 & .4 & .1 & 0\ \} \\ & \{.2 & .1 & .4 & .3\ \} \\ & \{.1 & .2 & .4 & .3\ \} \\ & \{.1 & .3 & .4 & .2\ \} \\ & \{.1 & .15 & .4 & .35\} \\ & \{.2 & .15 & .3 & .35\} \\ & \{.2 & .15 & .2 & .45\} \\ & \{.25 & .1 & .2 & .45\} \\ \text{and} & \{\ 0 & .4 & .6 & 0\ \} \end{array}$$ [2]

the solution was, <u>in every case</u>,

$$x_1 = 2.03 \qquad x_4 = 1.88 \qquad x_9 = 32.97$$
$$x_2 = 3.13 \qquad x_7 = 32.84 \qquad x_{11} = 22.81 \qquad \text{(Solution 4.2.1.)}$$

For priority weights {.2 .4 .1 .3} the solution was

$$x_1 = 3.01 \qquad x_4 = 11.41 \qquad x_9 = 54.42$$
$$x_3 = 1.28 \qquad x_7 = 28.58 \qquad x_{11} = 17.56 \qquad \text{(Solution 4.2.2.)}$$

Next, considering problem (4.3.) with priority weights

$$\begin{array}{llll} \{.4 & .35 & .15 & .1\} \\ \{\ 0 & .2 & 0 & .8\} \\ \{\ 0 & 1 & 0 & 0\} \end{array}$$

The solution was, in all these cases,

$$x_1 = 2.74 \qquad x_4 = 19.89 \qquad x_9 = 52.51$$
$$x_3 = 2.38 \qquad x_7 = 3.13 \qquad x_{11} = 12.37 \qquad \text{(Solution 4.3.1.)}$$

---

[2]Nine GP runs with these nine different weight specifications were performed.

For weights $\{d_i^-\}$ set at

$$\begin{array}{llll} \{.1 & .35 & .45 & .1\} \\ \{\ 0 & .35 & .55 & .1\} \end{array}$$

the solution was

$$\begin{array}{lll} x_1 = 4.56 & x_4 = 44.29 & x_9 = 17.26 \\ x_2 = 4.69 & x_7 = 36.48 & x_{10} = 49.78 \\ x_3 = .35 & & \end{array} \qquad \text{(Solution 4.3.2.)}$$

For weights

$$\begin{array}{llll} \{0 & 0 & .9 & .1\} \\ \{0 & 0 & .7 & .3\} \\ \{0 & 0 & .6 & .4\} \\ \{0 & 0 & .5 & .5\} \\ \{0 & 0 & .4 & .6\} \\ \{0 & 0 & .3 & .7\} \\ \{0 & 0 & .2 & .8\} \end{array}$$

the solution was

$$\begin{array}{lll} x_1 = .72 & x_5 = .08 & x_9 = 12.72 \\ x_2 = 5.82 & x_7 = 51.84 & \end{array} \qquad \text{(Solution 4.3.3.)}$$

Finally, for weights {0 0 1 0} the solution was

$$\begin{array}{lll} x_2 = 6 & x_7 = 54 & x_{11} = 9 \\ x_5 = 8 & x_9 = 12 & \end{array} \qquad \text{(Solution 4.3.4.)}$$

It is evident that solutions (4.2.1.) and (4.3.1.) are generated identically by naive weights and by non-naive weights. Being optima for weights that, for the moment, are considered valid indications of the DM's preference structure, we will say they are "good" solutions, and Theorem IV.3.1 is proved.

These two problems demonstrate that the DM, in some cases, should not spend much time setting weights. Since a weight is a scalar that is used to drive a system toward a desired state, weights are not very useful when they drive a system in the wrong direction, or do not change it much at all. For example, as we move from solution (4.3.3.) to (4.3.4.), the weight on the third goal goes up from .2 to .1. However, x_7, which represents underachievement of the third goal, goes up from 51.84 to 54.

Zeleny [1974(a), p. 405] writes, "Before the nonlinear case may be approached efficiently, more experience is needed not only with multiobjective programming, but also with interpretations of nondominated solutions." This important point can be addressed by referring to problems (4.2.) and (4.3.).

From Table 4.1 it is seen that for problem (4.2.), solution (4.2.1.) is worse than solution (4.2.2.) in terms of Goals 2 and 4, but better in terms of Goal 3. For problem (4.3.), solution (4.3.1.) is better than (4.3.2.) in terms of Goal 2, but worse in terms of Goal 3, and indifferent in terms of Goals 1 and 4 because of satiety. The DM in the present context, shifts his attention to other goals precisely at the target level g_i. That is modeled by using a weight of zero on all d^+ variables.

Similarly, from Table 4.1 pairwise comparisons can be made

Table 4.1

Criterial Achievements of Problem (4.3.)

	Goal Attainment	
	Solution (4.2.1.)	Solution (4.2.2.)
Goal 1	Over by 1.88	Over by 11.41
Goal 2	Under by 32.84	Under by 28.58
Goal 3	Under by 32.97	Under by 54.42
Goal 4	Under by 22.81	Under by 17.56
	Solution (4.3.1.)	Solution (4.3.2.)
Goal 1	Over by 19.89	Over by 44.29
Goal 2	Under by 3.13	Under by 36.48
Goal 3	Under by 52.51	Under by 17.26
Goal 4	Over by 12.37	Over by 49.78
	Solution (4.3.3.)	Solution (4.3.3.)
Goal 1	Under by .08*	Exactly met
Goal 2	Under by 51.84	Under by 54
Goal 3	Under by 12.72	Under by 12
Goal 4	Exactly met	Under by 9

*This will be interpreted as a zero; i.e., Goal 1 was achieved.

among all four solutions to (4.3.). In a real decision situation the goals would have names and the DM, presumably, could either choose his favored solution or go on experimenting with additional sets of weights. Even in this abstract problem setting two things are clear. First, for these two problems, there seems to be little correspondence between priority weights and goal attainment. Second, even though we have a formal definition of nondominated (good) solutions it is very difficult to determine the best one. This is because in the general multiple objective situation (MOS) there exists <u>incommensurability</u>.

This term means that the objectives are not measurable in the same units. Whether GP or MOLP is employed, it is not a simple mechanical task to reduce a set of good solutions down to a single one for action.

This task, which can be called pruning, is discussed by Zeleny [1976] and will be the subject of Chapter V below. As discussed above, setting weights in a GP context does not obviate the need for pruning. The following section is concerned with why, in the GP environment, naive weights do not uniquely span the space of good decisions.

IV.4 Eigenvalues of Criterial Matrices

Naive weights do not always suffice to implement the intentions of the DM. In this section three simple reasons for this will be proposed. The reasons will be unified by the use of eigenanalysis. Finally, since there are certain GP situations where prioritization by means of naive weights is possible, a tentative typology of these situations will be suggested.

As the weight vector is changed through the DM's experimentation and learning, the initial feasible set of solutions is enlarged. The classical theory of LP offers a means--called post-optimality analysis--of partially characterizing the shape and size of this expanded feasible set.

In the GP context, however, the actual decision variables have zero objective function coefficients. The audit trail on the information inherent in the problem's coefficients leads primarily through the RHS of the policy constraints into the objective function. But even here GP adds new complexity. These policy constraints [the C and the g of problems (4.1.), (4.2.) and (4.3.)] have both slack and surplus variables in them. This makes the optimal basis degenerate (see Simmonard [1966, p. 144]). They in turn are manipulated by a set of objective function coefficients which are not independent. For one to change, at least one other must adjust. Even if the model-builder does not restrict himself to weights that are convex combinations, changing one weight absolutely changes the relative relationships of all the $\{k_i\}$.

There are several reasons for the difficulty in weighting. One is the natural scaling of the deviation goals. In a road-building context, for example, one goal might be to minimize salt run-off and another to minimize capital budgets. The deviations for the first goal would typically be denominated in parts per million, and those of the second goal in millions. The second goal would dominate, even if its importance were secondary. Of course rescaling the $\{k_i\}$ can sometimes take care of this problem. But the choice of a right-hand side is a complicated matter that influences the size and directionality of the deviations.

Also, if the variance within each row of the criterion matrix C is low, that goal will be insensitive to basis changes.

A third distorting factor will be called <u>amplification</u>. Imagine a criterial matrix

$$C = \begin{bmatrix} 5 & 7 & 19 \\ 3 & 6 & 18 \\ 9 & 1 & 0 \end{bmatrix}$$

Each row represents the policy constraint of a goal; assume that all three goals are equally weighted. The three columns of C correspond to three decision variables x_1, x_2 and x_3.

The 19 that appears as C_{13} would penalize x_3. Since the smallest element of the first row of C is 5, the use of units of x_1 would be encouraged. In row 2, the ratio of 18 to 3 also forces the system toward x_1. The two goals represented by rows 1 and 2 are non-orthogonal; they amplify each other along. It is possible for goal 1 to be achieved with a zero weight if other goals "amplify" it sufficiently. This extreme case actually happened in problems (4.2.) and (4.3.).

LP theory has not addressed this issue because in LP the objective function coefficients are related to columns, not rows, of the constraint matrix. The coefficients are estimated independently. In GP, however, these coefficients are convex combinations. Therefore, to change one weight requires changing at least one other. This dependence among weights is an artifact of the mathematical representation. This phenomenon called amplification distorts the weight-setting process, causing the DM to unwittingly "use up" his weights on goals that may not need such attention.

Three issues have been raised as distorting factors in weighting; they are scaling, variance and amplification. In the next section eigenanalysis will be used as a rough way of unifying and tracking these three factors. In practice, using convex combinations of weights, one often changes three or four weights at a time. Since sensitivity analysis fails here, we will initiate an analysis that is both formal and ad hoc. The C matrix will now be discussed in depth, ceteris paribus.

We are concerned here with certain pathologies of the criterial matrix. Due to its structure in particular situations, the intent of weights set a priori can be lost.

Let C be a matrix of goals, or criteria, as in (4.1.). Thus C is appended to A.

$C = \|c_{ij}\|$, for goals $i = 1, 2, \ldots, p$ and decision variables x_j, $j = 1, 2, \ldots, n$

$\overline{C}$ is the "mean score" matrix of criteria, i.e.,

$$\overline{C} = \|c_{ij}\| = \frac{\sum_{j=1}^{n} c_{ij}}{n}, \text{ for each } i,\ i = 1, 2, \ldots, p$$

Next define

$D = C' - \overline{C}'$, where the superscript ' means "transpose."

$D = \|d_{ij}\|$ and can be thought of as a matrix composed of deviations about the mean, for each goal. (Note: goals are now columns, not rows.)

Following this paradigm reminiscent of regression analysis,

$$D'D = \begin{bmatrix} \Sigma d_1^2 & \Sigma d_1 d_2 & \cdots & \Sigma d_1 d_p \\ \Sigma d_2 d_1 & \Sigma d_2^2 & & \Sigma d_2 d_p \\ \vdots & & & \vdots \\ \Sigma d_p d_1 & \Sigma d_p d_2 & \cdots & \Sigma d_p^2 \end{bmatrix}$$

D'D is a sum of squares and cross products matrix. If we now divide each d_{ij} by n, the number of original decision variables (as opposed to the 2p deviation variables), we obtain

$$E\,[D'D] = VC, \text{ a variance-covariance matrix.}$$

The eigenvalues of VC can be readily computed; since VC is square and Gramian it is very tractable. A Gramian matrix G has the following properties:

1) Every principal minor determinant is nonnegative. If G is nonsingular, then the minors are positive.
2) G^{-1} is also Gramian, and nonsingular if G is nonsingular.
3) The quadratic form $Q = x'Gx$ is positive semidefinite.
4) The eigenvalues of G are nonnegative; if G is nonsingular, then they are positive.

Recall also that in the expression $Tx = \lambda x$, the eigenvectors x are the vectors of a linear transformation T which are stretched, contracted or reversed by T. The vector x makes T "act like" a scalar, and as the right-hand side of $Tx = \lambda x$ shows, that scalar is λ, the associated eigenvalue.

It is true that

$$\sum_i \lambda_i = \sum_i a_{ii}, \text{ where } \sum_i a_{ii} \text{ is the "trace" of A, the matrix of the transformation T.}$$

In addition, if A is the matrix of T, its determinant D is available; that is, $D(A) = \prod_i \lambda i$. The rank of A is equal to the number of positive eigenvalues. Therefore, if any eigenvalue is zero, then A is singular. Another convenient property is that the rank of VC is equal to the rank of C.

All these transformations were simply to obtain a square matrix for easy analysis. One cannot work directly on C when it is rectangular.

The eigenvalues of VC show the inherent variance of the criterial system if one were to diagonalize it. The eigenvectors form a matrix, V, and V' VC V = diag. In this form we have orthogonal goals, and any collinearity or correlation among the original goals is eradicated. Typically, there is enough misspecification of goals to induce collinearity. And thus the entire criterial force can often be explained in fewer than p dimensions.

In organizational life, the participants often allow each other's goals to be stated in the trial model, to avoid human conflict. These goals are often minor variations on each other. For example, maximizing profit, maximizing liquidity ratios, maximizing end of period net worth and maximizing return on equity are highly correlated goals. In theory this is evident through a thorough knowledge of accounting. In practice, for management science support groups, it is convenient to have a mathematical way of perceiving all this.

The phenomenon of collinear goals is indicated by a small number of positive eigenvalues in VC. In other words, if VC is "very" singular that means that there are non-orthogonal criteria.

Another way of looking at this is through the actual correlation matrix of the p goals. We had VC = D'D. Now let

$$\Delta = \begin{bmatrix} d_1^2 & & & \bigcirc \\ & d_2^2 & & \\ & & \ddots & \\ \bigcirc & & & d_p^2 \end{bmatrix}$$

It can be shown that R, the correlation matrix among the goals is

$$R = \Delta^{-\frac{1}{2}} \, VC \, \Delta^{-\frac{1}{2}}$$

Therefore, the phenomenon of goal collinearity is also to be seen as an R matrix with large off-diagonal elements.

We can review the eigenanalysis procedures by working on the C matrix of problem (4.2.). For that example,

$$VC = \begin{bmatrix} 10.67 & 9.33 & -9.33 & 10.67 \\ 9.33 & 8.67 & -7.33 & 10 \\ -9.33 & -7.33 & 9.55 & -8.22 \\ 10.67 & 10 & -8.22 & 11.55 \end{bmatrix}$$

and

$$R = \begin{bmatrix} 1 & .97 & -.92 & .96 \\ .97 & 1 & -.8 & .99 \\ -.92 & -.8 & 1 & -.78 \\ .96 & .99 & -.78 & 1 \end{bmatrix}$$

From R it is clear that the goals are extremely correlated. The point is made more clearly by the calculation of the eigenvalues of VC. There are only two, and the second one is insignificant compared to the first one. For the matrix VC, $\lambda_1 = 37.75$ and $\lambda_2 = 2.69$. This means that the entire criterial variance of (4.2.) can be expressed in two dimensions.

The additional eigenvalues λ_3 and λ_4 are equal to 0. Nevertheless, associated with each of the four λ_i is an eigenvector v. Construct V, a matrix with v as its columns. For (4.2.) this would be

$$V = \begin{bmatrix} .53 & -.10 & .02 & .84 \\ .47 & .34 & .77 & -.27 \\ -.45 & .80 & .05 & .38 \\ .54 & .48 & -.64 & -.27 \end{bmatrix}$$

<u>Conjecture IV.4.1</u>.

$$N' = (C'V)' = \begin{bmatrix} 19.57 & 4.54 & 11.4 \\ 10.75 & 10.95 & 7.38 \\ 1.24 & 1.24 & 1.24 \\ 3.15 & 3.15 & 3.15 \end{bmatrix}$$

is what the "principal components" of C are. In orthogonal goal space, almost all the driving power resides in the first row of N'. Rows 3 and 4 of N' are not equal to zero because of the scaling of the original rows of C.

Now we can see why (4.2.) was so insensitive to changes in the weight vector. Viewed after transformation to orthogonal goal space, (4.2.) has an "essential dimensionality" = 1. Therefore, changes in the $\{k_i\}$ have no more effect on the x vector than do multiplicative transformations on the single criterion of the classic LP problem.

To use the language of Zeleny, if the goals do not exhibit "contrast intensity," they should be collapsed to a lower dimensionality. This is why, in the terminology of this paper, naive weights can be "simulated" by non-naive weights.

A rough measure of goal collinearity would be

$$\alpha = \frac{\text{the number of positive eigenvalues}}{\text{the number of original goals}}$$

As α drops well below 1, the GP model-builder should be cautious about the validity of naive weights. He should experiment with many sets of weights, and attempt to redesign a more parsimonious set of goals. A serendipity is that while doing this, and while pruning the set of GP trial solutions, he will truly begin to learn his preference structure.

A subject for further research is to determine what type of C matrices will be plagued by scaling, variance and amplification. Problems (4.2.) and (4.3.) were doomed to have C matrices of less than full rank because the number of goals exceeded the number of decision variables. Two questions need attention:

1) By what factor must the number of decision variables exceed the number of goals to allow naive weights to function intuitively? As n becomes significantly greater than p, uncorrelated goals are much more likely.
2) For typical problems in functional areas like finance, marketing and engineering design, what is the typical relationship of p to n?

Another response to collinear goals would be to change the model from GP to MOLP. In that case no *a priori* weights need to be set. The duals in GP lose much of their interpretability. In MOLP a duality theory is currently being studied.

The RHS of the dual of a MOLP problem can be thought of as a *multiple* right-hand side. This raises interesting practical possibilities. For example, in financial capital budgeting the multiple RHS might be interpreted as budgets for different years or states of nature. The reader could review Lee [1972] and Ignizio [1976] to review the many applications of goal programming. Wherever GP seems to have been indiscriminately utilized, there is a reason to experiment with the more general multicriterion mathematical programming model.

IV.5 Summary

Psychological and mathematical reasoning has been used to demonstrate that naive weights are sometimes inappropriate. When this occurs, the DM can resort to multiple GP iterations, redesign of the problem, or MOLP. We see weights as metacriteria which cannot be revealed; this leads to mechanistic improvements to MOLP which are discussed in Chapter V.

CHAPTER V

PRUNING NONDOMINATED SETS IN MULTIOBJECTIVE LINEAR PROGRAMMING

V.1 Introduction

Howard Raiffa [1970, p. 155] said,

> Personally, I feel that this quest for a "scientific" and "mathematically objective" rule is all wrong! When there is a paucity of objective evidence at hand, we require a methodology that brings information, however vague and imprecise, into the analysis, rather than a methodology that suppresses information in the name of scientific objectivity . . . ; we should limit formal analysis to the characterization of the efficient set and let unaided, intuitive judgment take over from there.

This statement is referring to relatively small problems which are often characterized by independent choice attributes or goals. In other cases Raiffa [1970] and Keeney and Raiffa [1976] are extremely sophisticated about eradicating interdependency. But the "vague and imprecise" information referred to above is often the decision maker's intuitive sense of dependency among the goals of the MOS or of the dependency that is due to the structure of the feasible set (see Chapter III for a discussion of dependency).

Is it true that once the nondominated (N) set has been discovered, all mechanistic approaches are to be discarded? Shepard [1964] argues against this at length. Some man-machine symbiosis is fruitful when N is very large.

With the advantage of hindsight, it can be said that Raiffa's

1970 remarks became dated when the multicriteria simplex method was invented independently by Evans and Steuer [1973] and by Yu and Zeleny [1975]. Under the MOLP paradigm N can become immense. For example, Zeleny [1974(a), p. 117 ff.] shows that a certain problem with eight decision variables, five objective functions and eight constraints has seventy nondominated extreme points. If points along nondominated faces are considered, the size of N obviously moves toward infinity.

To represent the process of reducing N to one action, we will use the word pruning. The DM's perusal of N can be pictured as a search on a tree; thus the etymology of pruning becomes clear. Whether pruning is heuristic or aided by algorithmic devices, it is a search process through which the DM may discover new goals, relationships and cultural or strategic externalities. To be precise, pruning involves, by definition, the addition of new evaluative criteria to the problem. Sometimes the new criterion simply minimizes redundancy among the decision vectors of N. The methods of this chapter are applicable to both N^{ex} and to any other finite subset of N.

In particular multiple objective applications new criteria will surface to force the DM toward a choice. Or some of the original criteria can be dropped, which also reduces the size of N. Here I am concerned with methodology; the present criteria are not from a real-world problem. Therefore, minimizing redundancy is the criterion added for general pruning process. Cluster analysis is proposed for this task.

V.2 Cluster Analysis

The formation of m groups or clusters of similar objects, directly from the n original objects, is called cluster analysis (see Anderberg [1973], Bijnen [1973], Hartigan [1975], Sneath and Sokal [1973], and Sokal and Sneath [1963]). The clusters are usually mutually exclusive, or non-overlapping. In some measurable way, the clusters are to be composed of objects that are similar, either qualitatively or quantitatively. Operationally, a cluster can be defined as whatever results from a clustering algorithm. Or it can be one of the set of groups that emerges when some objective function, such as within-groups sum of squares, is minimized [Rao, 1971].

Formally speaking, a clustering is a mapping β: OTU $\rightarrow$ C, when OTU is the set of objects to be clustered (called <u>operational taxonomic units</u>) and C is the set of clusters. Each element of {C} is naturally a non-empty subset of {OTU}, but the mapping β is of very general form. If β is such that each element of {OTU} is mapped into one and only one cluster, then β is said to be a <u>partition</u>. If β forms clusters that are nested (i.e., some may be disjointed but others may be within larger clusters) then it is called a <u>tree</u>.

The most appealing view of a cluster comes from Carmichael and Julius [1968]. They define <u>natural clusters</u> as areas where the OTUs are dense, surrounded by areas where they are sparse. Implementing this computationally led to a myriad of attempts as described in the literature; many are, in some way, optimization algorithms themselves.

Ling [1973] proposes a rigorous way to define clusters, based on notions of geometrical probability theory. This work, unfortunately, is too theoretical to have an immediate impact on the practice of clustering.

In defining natural clusters homogeneity becomes an objective. Since the n objects are represented by a p-element vector, this leads to measures of similarity. Metric and nonmetric coefficients of similarity are discussed at length by Anderberg [1973].

There are three types of clustering strategies. First, the divisive approach starts with the n objects as a cluster, and successively refines this into smaller clusters. Second, the agglomerative approach starts each trial cluster as one object. Then these are linked together into larger and larger clusters. Third, the iterative approach begins with an arbitrary choice of clusters and moves objects from cluster to cluster until some homogeneity criterion is met (see Anderberg [1973, p. 156 ff.]).

In clustering large sets of data, optimal search techniques converge too slowly to be of practical use. Heuristic methods exist; the obvious drawback is that an optimal set of clusters is not guaranteed.

In this chapter we are using cluster analysis as an aid in multiobjective programming problems.[1] There is an important difference

---

[1] An ironic reformulation appears in Hanson [1977] where multiple criteria methods are used as an aid in clustering problems. This is called <u>bicriterion cluster analysis</u>.

between the treatment of outliers in most statistical applications and that of the present chapter. Statisticians normally develop methods to reduce the influence of outliers because they are interested in measures of central tendency and higher moments which characterize dispersion. But in pruning N, we wish to stress the outliers and the boundaries of the dense central areas.

The $\{z\} \varepsilon N$ are (geometric) rays that are "interpreted" by the preferences that are elicited from the DM. One ray will end up being the final solution vector chosen at the end of a two-stage pruning process. Some of the z vectors are so close to each other that the finite sensibility of the human mind cannot distinguish them. This near-collinearity is most apparent in the dense central areas of N. That is precisely why cluster analysis is useful to eliminate this redundancy.

But the outlying z vectors carry much information. They are solutions to the MOLP problem that will not join any cluster. They represent unique combinations of criterial achievement, or unique weights on those same criteria. Therefore it is important systematically to collect and display these points, even though they do not fuse any clusters. The RESIDUE facility of CLUSTAN does exactly this. In Chapter VI the set composed of the outliers plus a representation of each cluster (this could be the centroid or any other element), will be the subject of attention. Letting z be any solution vector which joins a cluster, and z" be any one that doesn't, we can form the set

$$\{g\} = G(N) = \{z'\} \cap \{z''\} \subseteq N^{ex} \subset N.$$

The sixth chapter will refer to this set as the generators of N (or N^{ex}).

Summing up, clustering involves choosing distance and joining rules so that the DM and the analyst can draw a picture of the infrastructure of N. The {z"} must be accentuated by means of the RESIDUE facility because they carry more information (surprise value) than do the {z'}. This is analogous to simplifying and interpreting a painting by choosing the points that show what the artist wishes us to see and feel. For example, in a picture of a mountain, the point that marks the top is more crucial than one of the thousands of points that convey the details of the foothills. If the pruning process that accompanies MOLP does not expose the DM to {z"}, it fails in the same way that goal programming does. Steuer [1975, p. 5] has written that GP presents the DM with "too sparse" a representation of the nondominated set.

V.2.1 Direct clustering. Direct clustering was introduced in Hartigan [1972], with slight amendments added in computer codes written by the author in Hartigan [1975] and BMD [1975]. This type of clustering is initially very appealing to the researcher in fields where taxonomic methods are relatively new. This is because direct clustering (also known as block clustering) does not require the specification of a distance function.

The distance function describes inter-object distances which are used to create a similarity matrix. An example is the Euclidean distance

$$d_{ij} = [\sum_{j=1}^{n} (x_{ij} - x_{kj})^2]^{\frac{1}{2}}, \text{ for } i = 1, 2, \ldots, m \text{ and } j = 1, 2, \ldots, n. \tag{5.1.}$$

When one considers both binary, continuous and mixed data, the literature includes at least fifty different distance coefficients. Since the resulting clusters are very dependent on the definition of distance, the analyst must be wary of spurious clusters. In disciplines such as zoology, where the use of cluster analytic methods dates back to the first brush with twentieth-century empirical research methods, there are theories of distance. It is known, for instance, that particular distance coefficients perform well with particular types of zoological and biological data.

In the administrative and policy sciences, only marketing research has passed the beginning of the learning curve with cluster analysis. The application at hand, pruning the nondominated sets in MOLP problems is the first attempt in the literature. Ralph Steuer [1976(a), p. 11] uses a filtering technique that is equivalent to forming clusters from given cluster centroids. But he does not call the procedure cluster analysis, nor is the implied algorithm one that really finds clusters.

Since clustering in this context is new, there is no theory of distance. The question remains "Of all the possible distance measures, which best describes the DM's (psychological) reactions as he compares elements of N^{ex}?" Zeleny [1974(a)] addresses this issue. His distance functions are part of a rather mechanistic approach requiring the use of an ideal point. This point is that intersection of rays (usually not feasible) when each objective function finds its maximum in z space (criteria space).

In this section we will assume that the DM wishes to do some searching before the biases inherent in the display of an ideal point are introduced. Owing to a lack of such an anchor and to interdependencies of the structure of the feasible set, direct clustering is very appealing because it is not necessary to specify a distance function (see Appendix III for a description of the direct clustering algorithm).

For a data set two of the few solved problems in the literature were chosen. One will be referred to as the Zeleny problem (see Appendix II) and the other will be called the Steuer problem (see Appendix III). Both are MOLP problems, and the appendices contain both a statement of the problem and a complete list of N^{ex}. These nondominated points form the inputs for the various clustering attempts discussed in this chapter.

Seventy computer runs of direct clustering were performed. As in dynamic programming we face the curse of dimensionality; seventy-way cross tabulation, or chi-square experiments are impossible to interpret. Therefore, for this section we will not concentrate on the actual clusters of nondominated points. Since the objectives of the test problems are not real-world problems, the clusters have little intuitive meaning.

What will be examined is the behavior of the algorithm itself, over a large number of permutations of the discretionary parameters.

In Tables 5.1 through 5.4 "intervals" refers to the number of segments into which each z vector is divided. For example, "2,3,6,6,1" means that the range exhibited by the z vectors (N^{ex} points) in Appendix II was divided into two intervals for the first objective, three intervals for the second objective, six for the third objective, and so on.

The column dealing with blocks refers to the basic output of this type of clustering, which is data blocks. By permuting rows and columns of the data matrix (which has been transformed into intervals) areas of similar values appear. These, in turn, indicate case clusters along the vertical margin of a case-by-variables matrix. Simultaneously there will be variable clusters created as the marginal clusters along the horizontal axis. This can be seen in the part of Appendices IV and V labeled "Block Diagram." For a complete explanation refer to Hartigan [1972, 1975].

Blocks "not single" means that more than one case (OTU) entered the block. In the present application, if the number of "single" data blocks is large, no meaningful clustering has been achieved. With our

Table 5.1

Zeleny Data (10 Passes)

# of intervals within each obj.	# of blocks total	# of blocks not single	# of blocks single	Pass # of join	Goals that join	Pass # of join	Goals that join
1, 1, 1, 1, 1	1	1	0	2	3rd,4th	10	2nd,4th,3rd
1, 1, 6, 6, 1	14	11	3	8	2nd,5th	10	2nd,5th,4th
2, 2, 2, 2, 2	15	10	5	8	3rd,2nd	10	3rd,2nd,1st
2, 2, 4, 4, 2	45	22	23	10	1st,2nd,5th		
*2, 2, 4, 4, 2	28	26	2	14	1st,2nd,5th		
3, 3, 4, 4, 3	55	36	19	8	3rd,5th	10	3rd,5th,2nd
3, 3, 3, 3, 3	51	23	28	8	5th,2nd	10	5th,2nd,3rd
4, 4, 3, 3, 4	58	30	28	8	3rd,5th	10	3rd,5th,1st
4, 4, 4, 4, 4	66	32	34	8	3rd,5th	10	3rd,5th,1st
*4, 4, 4, 4, 4	56	32	24	12	3rd,5th	14	3rd,5th,1st
5, 5, 5, 4, 5	81	42	39	8	3rd,5th	10	3rd,5th,1st
5, 5, 5, 5, 5	90	39	51	8	3rd,5th	10	3rd,5th,1st
6, 6, 6, 6, 6	95	46	49	8	3rd,5th	10	3rd,5th,1st
7, 1, 1, 1, 1	8	7	1	2	2nd,5th	10	2nd,5th,4th
7, 7, 7, 7, 7	102	52	50	8	3rd,5th	10	3rd,5th,1st
8, 8, 8, 8, 8	102	55	47	8	3rd,5th	10	3rd,5th,1st
9, 9, 9, 9, 9	107	50	57	8	3rd,2nd	10	3rd,2nd,1st
10,10,10,10,10	123	55	68	8	3rd,2nd	10	3rd,2nd,1st
11,11,11,11,11	125	56	73	8	3rd,2nd	10	3rd,2nd,1st
12,12,12,12,12	137	51	86	8	3rd,2nd	10	3rd,2nd,1st

* Indicates that a run with 14 passes was performed.

Table 5.2

Zeleny Data (16 Passes)

# of intervals within each obj	# of blocks total	# of blocks single	# of blocks not single	Pass # of join	Goals that join	Pass # of join	Goals that join
1, 1, 1, 1, 1	1	1	0	2	3rd,4th	16	2nd,4th,3rd
1, 1, 6, 6, 1	14	11	3	12	2nd,5th	16	2nd,5th,4th
2, 2, 2, 2, 2	12	10	2	8	3rd,2nd	16	3rd,2nd,1st
2, 2, 2, 2, 2	27	21	6	16	5th,1st,2nd	16	
3, 3, 4, 4, 3	46	38	8	12	3rd,5th	16	3rd,5th,2nd
3, 3, 3, 3, 3	42	27	15	12	5th,2nd	16	5th,2nd,3rd
4, 4, 3, 3, 4	50	29	21	10	3rd,5th	16	3rd,5th,1st
4, 4, 4, 4, 4	56	34	22	12	3rd,5th	16	3rd,5th,1st
5, 5, 5, 4, 5	76	43	33	12	3rd,5th	16	3rd,5th,1st
5, 5, 5, 5, 5	77	41	36	12	3rd,5th	16	3rd,5th,1st
6, 6, 6, 6, 6	91	43	48	14	3rd,5th	16	3rd,5th,1st
7, 7, 7, 7, 7	95	51	44	12	3rd,5th	16	3rd,5th,1st
7, 1, 1, 1, 1	7	7	0	2	4th,5th	16	3rd,5th,4th
8, 8, 8, 8, 8	95	45	50	14	3rd,5th	16	3rd,5th,1st
9, 9, 9, 9, 9	106	55	51	14	3rd,2nd	16	3rd,2nd,1st
10,10,10,10,10	118	57	61	14	3rd,2nd	16	3rd,2nd,1st
11,11,11,11,11	122	55	67	14	3rd,2nd	16	3rd,2nd,1st
12,12,12,12,12	127	53	74	14	3rd,2nd	16	3rd,2nd,1st
13,13,13,13,13	137	53	84	14	3rd,2nd	16	3rd,2nd,1st

Table 5.3

Steuer Data (10 Passes)

# of intervals for each obj.	# of blocks total	not single	single	Pass # of join	Goals that join
1, 1, 1, 1	1	1	0	10	3rd, 2nd
2, 2, 2, 2	33	10	23	10	1st, 2nd
3, 3, 3, 3	41	19	22	10	2nd, 3rd
4, 4, 3, 4	55	18	37	10	4th, 1st
5, 5, 5, 4	69	22	47	10	2nd, 4th
5, 5, 5, 5	75	29	46	10	1st, 3rd
6, 6, 6, 6	75	30	45	10	2nd, 1st
7, 7, 7, 7	84	26	58	10	3rd, 1st
8, 8, 8, 8	84	27	57	10	3rd, 4th
9, 9, 9, 9	100	32	68	10	2nd, 1st
10, 10, 10, 10	102	24	78	10	3rd, 1st
11, 11, 11, 11	109	29	80	10	2nd, 1st
12, 12, 12, 12	101	29	72	10	3rd, 4th
13, 13, 13, 13	112	27	85	10	3rd, 1st

Table 5.4

Steuer Data (16 Passes)

# of intervals within each obj.	# of blocks total	# of blocks not single	# of blocks single	Pass # of join	Goals that join
1, 1, 1, 1	1	1	0	16	3rd, 2nd
2, 2, 2, 2	18	16	2	16	1st, 2nd
3, 3, 3, 3	29	20	9	16	2nd, 3rd
4, 4, 3, 4	43	25	18	16	4th, 1st
4, 4, 4, 4	41	22	19	16	4th, 2nd
5, 5, 5, 4	61	27	34	16	3rd, 1st
5, 5, 5, 5	69	28	41	16	1st, 3rd
6, 6, 6, 6	63	38	25	16	2nd, 1st
7, 7, 7, 7	81	34	47	16	3rd, 1st
8, 8, 8, 8	83	31	52	16	3rd, 4th
9, 9, 9, 9	93	32	61	16	2nd, 1st
10,10,10,10	97	31	66	16	3rd, 1st
11,11,11,11	98	36	62	16	2nd, 1st
12,12,12,12	98	47	51	11	3rd, 4th
13,13,13,13	109	39	70	16	3rd, 1st

managerial focus, we eschew this sort of disaggregation.

The remaining two columns of Tables 5.1 through 5.4 refer to "pass numbers," which are fully explained in Appendix I. Each problem was run many times with 10 passes and with 16 passes (an occasional run with 14 passes was also performed as part of this empirical experiment). More passes allow the algorithm to represent the data in fewer blocks, which is desirable. Of course, more passes increases the computation time, which is not desirable on large data sets.

Direct clustering is a two-way splitting technique. That means that clustering occurs simultaneously for the rows and for the columns. Revealed clusters in the columns show a bad kind of goal congruence, namely goal collinearity, or correlation. It is necessary for the analyst to check this; the methods of Chapter IV help do that job in a reliable way.

Tables 5.1 through 5.4 show that for the Zeleny data, the third, fifth and first objectives join with great frequency when the number of intervals is three, four or five. Arbitrarily assuming that ten data blocks are considered a good number by the DM, we are forced to reduce to two intervals for all objectives for both the Zeleny data and the Steuer data. It is more likely that the DM would find greater interpretability in ten case clusters. By starting sufficiently to the left on the joining diagram for cases (see Appendices IV and V), the analyst can be assured a wide range of the total number of clusters. At any setting of the parameter intervals the tree begins with many case clusters on the left, and moves towards the right with fewer and fewer clusters.

Two things can be said at this point. First, since the joining structure of the objectives is so sensitive to the number of intervals, we are led to more classic mathematical methods. Second, the reduction of intervals, on both data sets, to two on each objective indicates a very gross division of the criteria achievement, represented by the z_i^q, q = 1, 2, . . . , s and i = 1, 2, . . . , p. This could lead to relatively non-homogeneous clusters. This is often referred to in the literature as clusters that are not "tight."

Direct clustering performs a degradation on the data; ratio-scaled data is coded into ordinally-scaled data. For example, the set of z_i^q values (7, 9.41, 3, 27) could be coded into three intervals on the range 3 through 27. The new data would be (1, 1, 1, 3). But there is a huge amount of cluster analysis literature on distance functions for non-ratio-scaled data. Therefore the claim that direct clustering does not require the user to define a distance function is very misleading. He does define a distance measure when he chooses the level of aggregation (number of intervals) for each objective.

In the section on "thresholds" in Appendix I, it is implicit that distance concepts are used in the algorithm. Perhaps this is why Hartigan [1972] said he found "disappointingly few" differences between direct clustering and average distance linkage (to be explained in section V.2.2) when both were applied to state voting data.

For Hartigan's data the average distance linkage method produced more tiny clusters (a failure to achieve the pruning we desire) and used more computer time. But computer time is of less importance each year, since unit computation costs are still dropping. As will be seen in section V.2.2 below, the average distance linkage method was one of the best of all the algorithms for Zeleny data.

There is one more way in which direct clustering behaves implicitly like the hierarchical distance methods. Hartigan [1972, p. 126] explains that direct clustering "splits up" the data matrix in such a way that the downside change in error sum-of-squares (SSQ) is maximized at each stage. Ward's method (described below) is one of the best hierarchical methods, and fuses clusters at each stage using essentially the same minimizing objective function.

V.2.2 Hierarchical clustering. As the attractiveness of direct clustering diminishes, the use of the hierarchical methods is increased. This type of clustering system does not split data blocks as does direct clustering. Instead it proceeds only on cases in the present application, to join OTUs into successively larger clusters.

Eight different hierarchical algorithms were applied to the Zeleny N^{ex} points. In the following paragraphs, brief summaries of these methods are given. The original references can be found in the bibliographies of Anderberg [1973], Hartigan [1975], and Wishart [1975].

First, there are three linkage methods. In single linkage algorithms the similarity between any two clusters is equal to the highest single similarity coefficient between two cases, each of which resides in a different cluster. In other words, when searching for two intermediate clusters to fuse, the nearest neighbor rule prevails. The single linkage method finds long serpentine clusters very well. But on larger data sets this can lead to chaining. Chaining occurs when most of the OTUs join one cluster, and the rest form a meaningless residue.

The complete linkage method uses the smallest similarity (or highest dissimilarity) to amalgamate two points or two clusters into a new one. In other words a farthest neighbor rule is employed. This strategy produces spherical clusters. But these may be spurious clusters, since group structure is not considered. Outlines may be too influential in forcing clustering.

The third method is the average linkage method, which considers the average of all the similarity coefficients for all pairs of individuals across clusters. Thus group structure is recognized. This algorithm, which has been introduced independently by several authors, is a well-behaved search strategy which creates spherical clusters.

The fourth hierarchical option employed was centroid sorting (also known in the literature as the weighted group method of Sokal and Michener). Clusters are joined by finding the center of gravity, or the mean of the clusters. Chaining occurs, but less perniciously than in the single linkage method.

The fifth approach is the median method. Valid mainly with distance type coefficients, this method computes a distance s (r, p+q) where r is an arbitrary cluster and p+q is a cluster fused during the last stage of the process. The distance s (r, p+q) is measured from the centroid of x and the midpoint of the line which joins the centroids of p and q. Again, chaining is likely for large populations.

The sixth type is known as Ward's method, although Orloci and Wishart have also written about it. First, the distance of each OTU to the centroid of its parent cluster is recorded. For each cluster these distances are added up and called the error sum of squares. The next two clusters to be joined are those for which the increase in this error sum of squares is minimized. Recall that the direct clustering technique of section V.2.1 reversed this; a split was performed which maximized the decrease in the error sum of squares. Ward's method computes minimum-variance spherical clusters.

The Lance-Williams flexible beta method seems to be very promising. It offers the chance to dilate or contract space by varying only one parameter of a joining measure. In the expression

$$s(r, p+q) = \{[s(r,p) + s(r,q)(1-\beta)]/2s(p,q)\}\beta \qquad (5.2.)$$

it is β that acts like the dilation and contraction operators of Zeleny [1974(a), p. 172] or Zadeh [1963]. With $\beta = -0.25$, the flexible beta method is reported to act like Ward's method.

The last major variant of hierarchical clustering is <u>McQuitty's similarity analysis</u>. Two clusters are joined to form one new one by minimizing the coefficient

$$s(r,\ p+q) = [s(r,p) + s(r+q)]/2. \qquad (5.3.)$$

Notice that this is equivalent to the Lance-Williams method with their $\beta = 0$. Again, chaining occurs with large populations.

All eight of the hierarchical clustering algorithms were used on the Zeleny N^{ex} points. As Table 5.5 indicates, no less than five of these methods caused terrible chaining problems. About 60 of the 70 undominated points tended to join one cluster, with the remainder forming clusters each with only one, two, or three elements. This outcome was useless managerially, since it does not reduce the redundancy among the N^{ex} set. The DM would still have to process all the points in the "chained" cluster heuristically. Chapter III showed that the human brain may be poorly suited to this task.

Table 5.5

Hierarchical Clustering Methods that Caused Chaining

1. The single linkage method. NOTE: This method was implemented by both BMDPO2M (Biomedical Computer Programs [1975]) and Clustan 1C [1975].
2. The complete linkage method (i.e., the farthest neighbor rule).
3. Gowers median method.
4. The Lance-Williams flexible beta method.
5. McQuitty's similarity analysis.

Three of the methods performed well on the Zeleny test data. They were Ward's method, the group average method, and the centroid method. As Appendix VI shows, the three methods produce clusters with fairly similar composition across the methods. In addition, the clusters are all of a size that is managerially useful (as will be explained in section V.2.3). Appendix VI contains too much data to be subjected to a contingency analysis, or chi-square tests. But since each method was tested with between ten and five clusters "requested," the behavior of the three clustering strategies is not extremely difficult to follow by inspection.

There are commonly used methods to help absorb the masses of information provided by a good computer code for clustering. For example, in Figure 5.1 the positions of the N^{ex} points are plotted and numbered. Since the portrayal occurs in the plane, only objective function #1 versus objective function #2 can be visualized in this way. Naturally, in a real business decision all pairs of objective functions could be so plotted.

Figure 5.2 reduces some of the visual clutter of Figure 5.1 by drawing circles around the points in each cluster, and eliminating the actual case numbers. Since the computer, in this particular case, was coded to produce seven clusters, the reader will note seven circles. Except for cluster #6, the clusters are reasonably tight, spherical and disjointed as seen along <u>the first two criterial dimensions</u>.

Figure 5.1

Scatter Diagram of Zeleny Problem

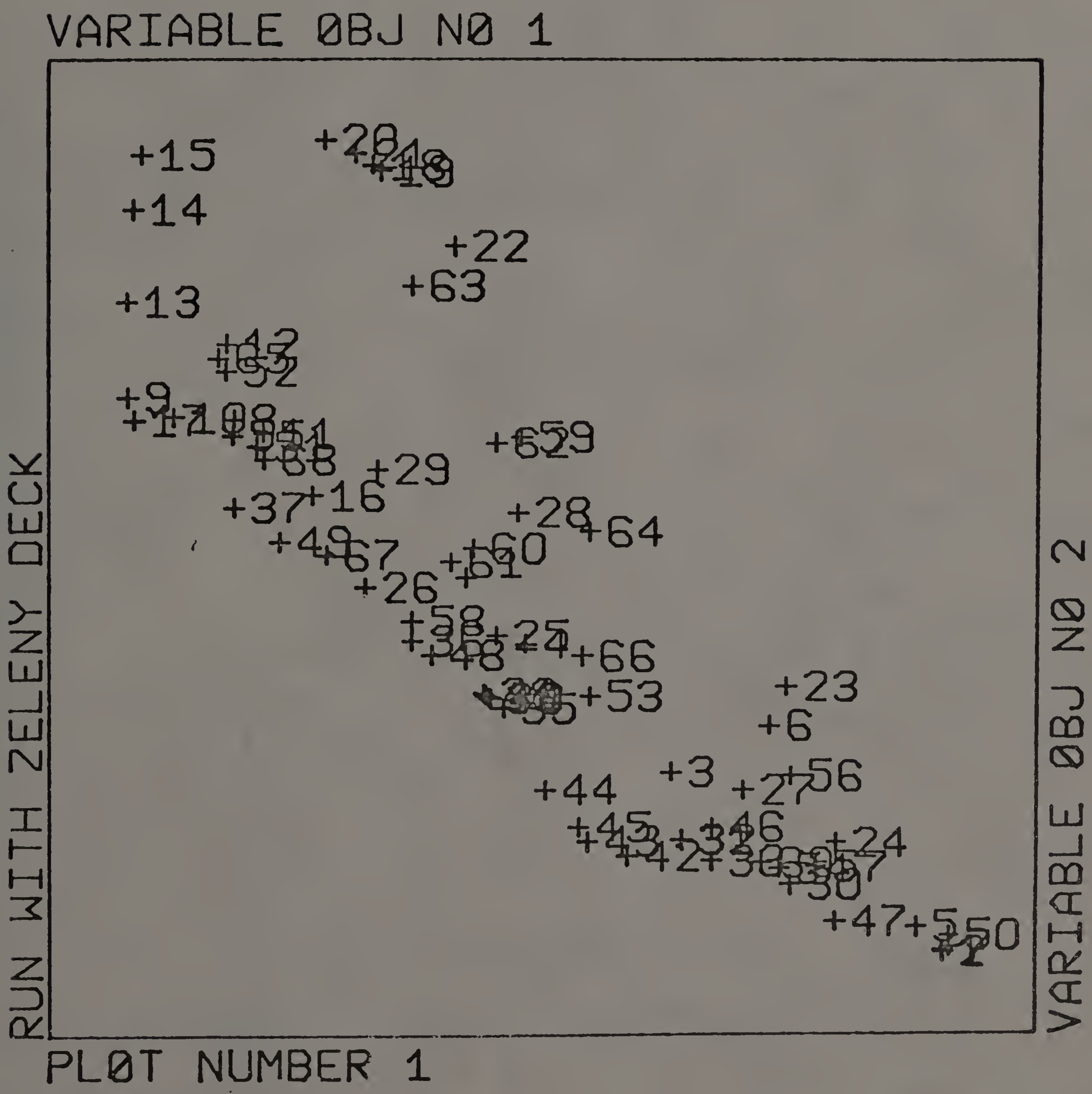
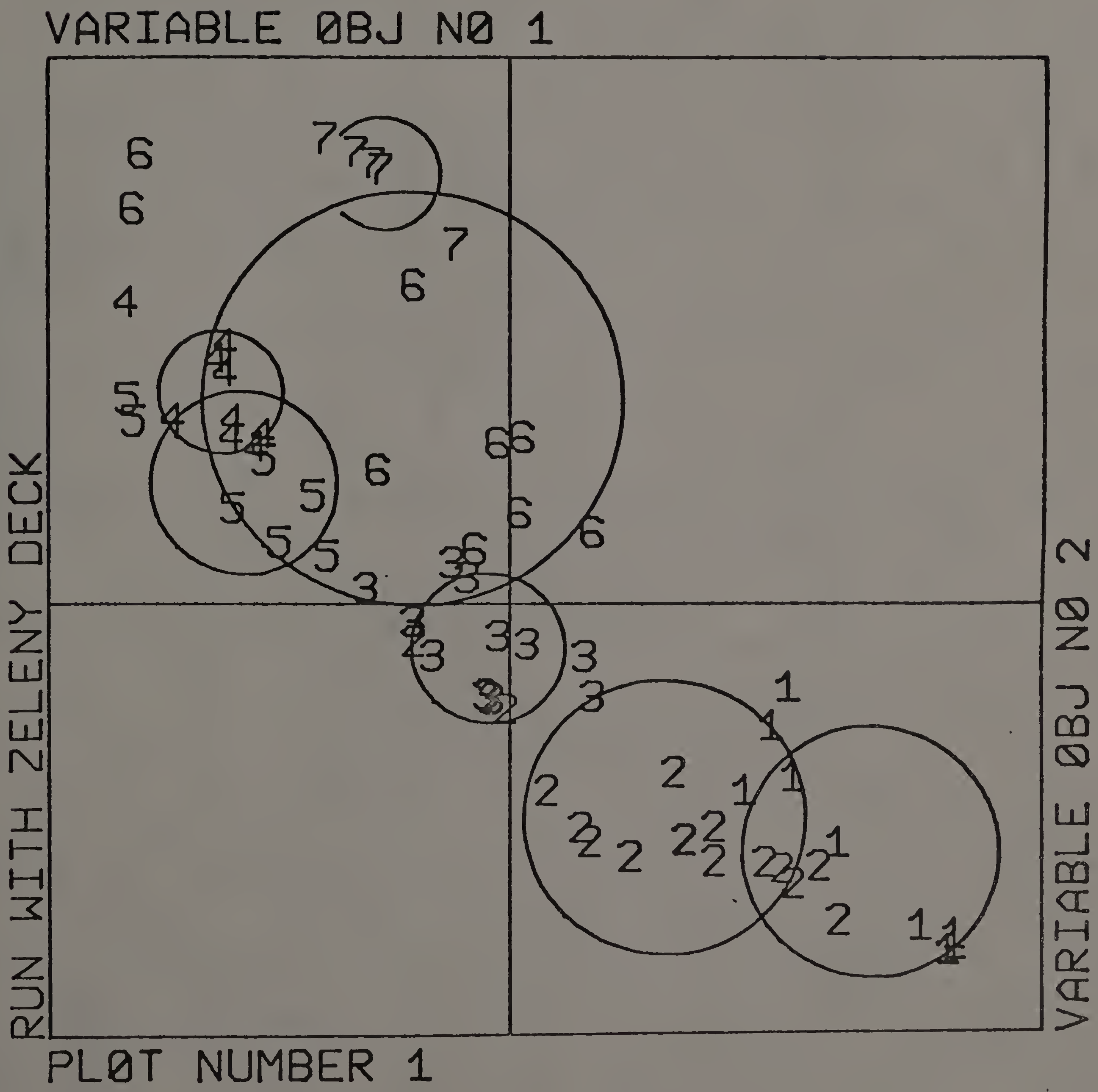


Figure 5.2

Cluster Diagram of Zeleny Problem



The most informative graphical aid is the dendrogram, as seen in Figure 5.3. The dendrogram draws the joining of two OTUs and indicates, on its vertical axis, the level of similarity at which this occurs. This is similar, but much clearer than the tree joining which was provided by Hartigan's direct clustering (see Appendices IV and V). Figure 5.3 shows both the clusters that are formed at varying levels of amalgamation and the order of that joining.

For the DM who prefers to think algebraically rather than graphically, some interesting data manipulations are available via the minimum spanning tree. This has appeared in graph theory and in operations research (for example, see Kruskal [1956]).

Consider this structure as it appears in what is called the traveling salesman problem. A group of cities is to be connected by a system of roadways. Various paris of cities will be linked, and over each link there will be a known fixed cost per mile. The problem is to find a route through the maze of cities, such that each one is visited and the cost is minimized.

The solution to this problem, in graph theory terms is the minimum spanning tree (MST). The DM can develop clusters directly from an examination of the MST. By deleting links whose length is greater than g, maximal single-linkage type clusters of diameter g are discovered. Table 5.6 gives the minimal spanning tree for the Zeleny problem.

Conceptually similar to the minimum spanning tree is the k-linkage list. This shows the k nearest neighbors for each OTU. It is rather like Steuer's "filter" method [1976(b), p. 11], which focuses the DM's attention on a particular set of N^{ex} points. Steuer

Figure 5.3

Dendrogram of Zeleny Problem

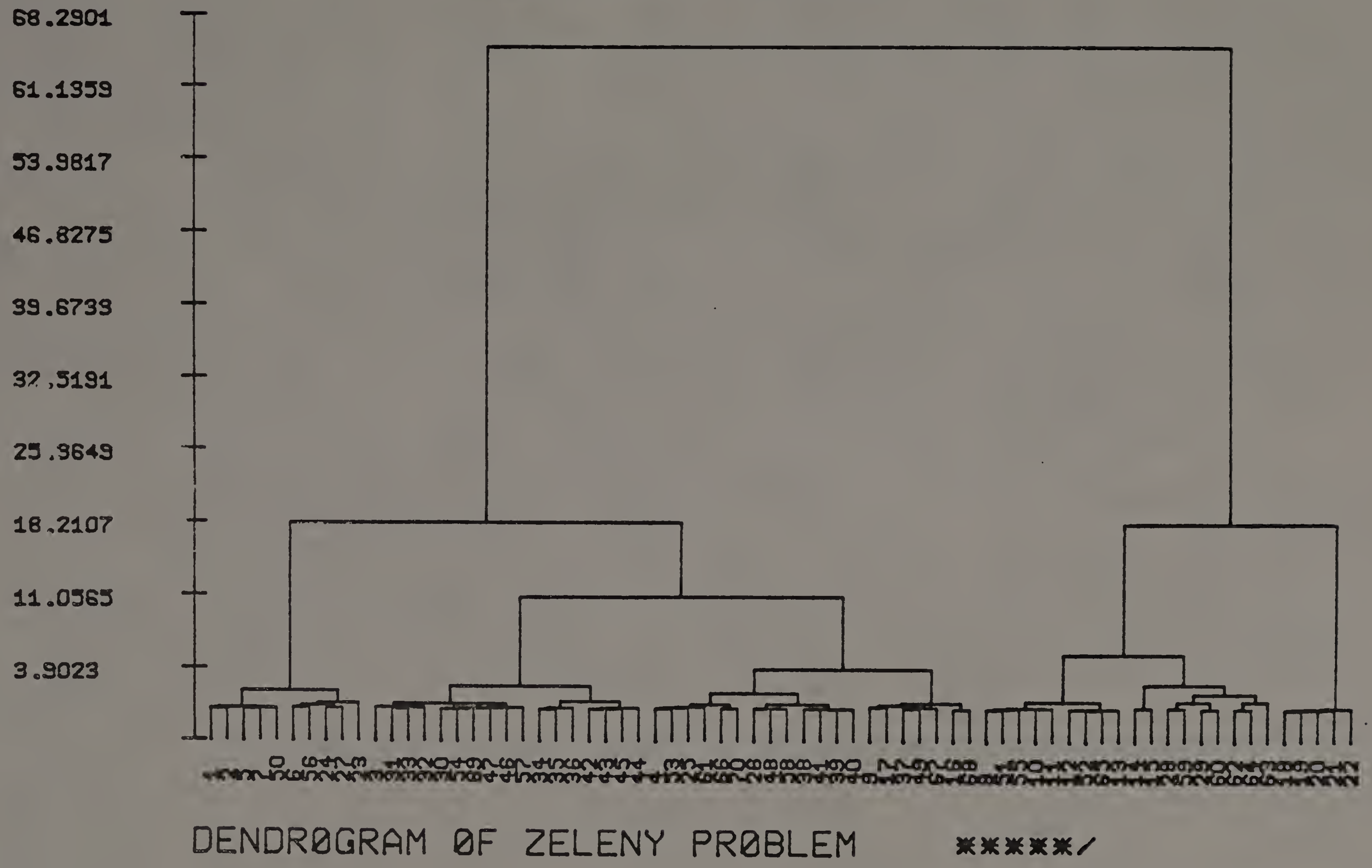


Table 5.6

The Minimum Spanning Tree of the Zeleny Problem*

Edge No.	First Vertex	Second Vertex	Edge Length	Edge No.	First Vertex	Second Vertex	Edge Length
1	1	7	0.005	36	36	38	0.111
2	2	1	0.001	37	37	49	0.042
3	3	53	0.093	38	38	41	0.000
4	4	53	0.032	39	39	41	0.000
5	5	7	0.007	40	40	39	0.000
6	6	56	0.099	41	41	48	0.044
7	7	50	0.007	42	42	43	0.008
8	8	55	0.031	43	43	45	0.001
9	9	55	0.092	44	44	35	0.082
10	10	51	0.062	45	45	44	0.012
				46	46	69	0.017
11	11	51	0.047	46	46	69	0.017
12	12	65	0.021	47	47	30	0.014
13	13	65	0.046	48	48	58	0.020
14	14	12	0.212	49	49	67	0.052
15	15	14	0.147	50	50	24	0.113
16	16	70	0.167	51	51	8	0.016
17	17	37	0.071	52	52	11	0.061
18	18	19	0.011	53	53	66	0.084
19	19	21	0.021	54	54	69	0.003
20	20	18	0.016	55	55	67	0.179
21	21	63	0.511	56	56	53	0.336
22	22	19	0.059	57	57	54	0.040
23	23	6	0.140	58	58	25	0.089
24	24	56	0.178	59	59	28	0.180
25	25	4	0.053	60	60	29	0.074
26	26	58	0.033	61	61	4	0.064
27	27	24	0.051	62	62	59	0.414
28	28	60	0.174	63	63	62	0.395
29	29	52	0.130	64	64	62	0.075
30	30	54	0.003	65	65	52	0.016
31	31	42	0.081	66	66	70	0.086
32	32	31	0.074	67	67	58	0.110
33	33	31	0.006	68	68	16	0.016
34	34	36	0.052	69	69	33	0.051
35	35	34	0.002				

* Edge length is Euclidean distance.

exhibits a subset of N^{ex} and asks, "Which point do you like best?" Then he prints the k nearest neighbors to that point.

Steuer's procedure is appealing but non-optimal. Certain attractive areas of the nondominated set may be prematurely discarded in this way. The DM should have the opportunity to examine a neighborhood around a tentatively desirable point and compare it to several other neighborhoods. This idea is clearer if the reader refers to Appendix VII. This appendix contains partial output from Ward's Hierarchical Clustering Aglorithm. The particular code used CLUSTAN 1 C) includes many features which help understand the structure of N. For example, one can see the order of agglomeration, a dendrofram table, k-linkage lists, and descriptive statistics for each of the final clusters.

V.2.3 Direct clustering versus hierarchical clustering: a summary. Direct clustering has two main advantages, computational speed and direct interpretation of the data. With modern computers the first is only relevant on extremely large problems. Hartigan [1975, p. 267] states that two-way splitting (another name for the type of direct clustering used in this chapter) draws "beautiful (data) pictures." But he cautions that the data must be rescaled so that an error of one unit along one variable dimension will be equal to a one-unit error along another dimension. This rescaling creates problems when the data are of several sorts such as binary, nominal, interval or ratio. Hierarchical clustering also has difficulty overcoming mixed data bases.

Shift phenomena are particularly well captured by direct clustering. In Hartigan [1972] presidential election data is analyzed. In South Carolina, Louisiana and Mississippi the percentage of votes cast for Republicans went up greatly from 1960 to 1964, and then down again in 1968. This shift over time would not be tracked well by hierarchical methods because states and years would not be simultaneously clustered.

From a managerial point of view direct clustering, as mentioned earlier, cannot specify the desired number of clusters. An allied problem is the meaning of a case cluster. Since these are seen in Appendix IV as forming along a subset of the objectives, the DM cannot tell whether the cases would cluster if all objectives (or criteria) were considered. There is an implicit 0-1 weighting of criteria; formation of the case clusters in this way does not interface well with the DM's intentions.

In the section on hierarchical clustering we saw how important the RESIDUE facility of CLUSTAN was. In direct clustering this feature could be copied by taking the small data blocks and listing them as additions to the list of generators $g \in G(N)$. But sometimes to keep the size of G(N) manageable, we need to assign outlying z vectors to existing clusters. This can be done with the RELOCATE option of CLUSTAN, but cannot be done in direct clustering.

The conclusion here is that hierarchical clustering methods show more promise than direct clustering for the purpose of pruning the non-dominated set N. However, the choice of clustering strategy is not as risky as it is in a static taxonomic study. In the dynamic decision context we are never quite certain of the effects of our strategies. As the problem evolves, new information overwhelms the effect of less than optimal clustering in the previous stage. We are not explicitly introducing multiperiod decision making.

V.3 Other Methods for Pruning the Nondominated Set

In the preceding sections of Chapter V experiments in pruning by clustering were conducted. The conclusion was that the task should be conducted by any hierarchical strategy which does not cause chaining on the particular realizations of $\{z\} \subseteq N^{ex} \subset N$. In addition the clustering technique must have the RESIDUE and RELOCATE capibilities.

The recommended pruning method emerged from a group of possible methods. These will now be described, along with reasons for their unsuitability. This section comprises an important part of the dissertation, because with the exception of the methods of Roy and Zeleny, these pruning methodologies are new to the MCDM literature. Although they are rejected here, it is possible that subsequent authors will successfully resurrect them.

V.3.1 Outranking relationships. Outranking relationships [Roy, 1971, 1973 and 1975(a)] are attempts to capture the information in certain binary relationships with a utility function that keeps many of the useful properties of the Von Neumann-Morgenstern utility function. Roy felt that it might not be necessary to force transitivity and completeness onto the multicriterion preference ordering. His outranking concepts "must not be viewed as an exact reflection of all the DM's preferences, but only as the expression of the part of his preferences that can be well accounted for by means of the available data" [from Roy, 1971, p. 252].

The set of criteria $I = \{1, 2, . . . , n\}$ is partitioned in this three-way design:

I^{+} (a', a"): set of criteria for which a' is preferred to a",

$I^{=}$ (a', a"): set of criteria for which a' is indifferent to a",

I^{-} (a', a"): set of criteria for which a" is preferred to a'.

(5.4.)

Additionally, let P^{+} (a', a"), $P^{=}$ (a', a"), and P^{-} (a', a") be three scalars that represent, as in weights, the differing importance of each of the subsets of I. For example, if the first and second criteria are two that favor alternative a' over a", the DM may choose to weight these criteria higher than the ones which fail to rank a' versus a". One decision making condition could be

$$C[P^{+} (a', a''), P^{=} (a', a''), P^{-} (a', a'')] \geqq c \qquad (5.5.)$$

such that a'Ra" cannot be true if the pair (a', a") does not satisfy (5.5.) where R is the relationship "is preferred to." The left-hand side of this expression shows how well the three subsets of criteria agree on the proposition that a' is preferred to a". The three criterial sets concur to various degrees; the term on the right, the c, is a level of concordance that is desired by the DM. It is a parameter which obviously affects the degree to which a preference relationship R approaches the completeness condition.[2]

---

[2]Completeness means that $\forall$ x, y ε X: xRy or yRx. In other words there are no indifferent pairs.

Other possible types of concordance functions are:

$$P^{+}(a', a'') + P^{=}(a', a'') \geqq c \tag{5.6.}$$

$$P^{+}(a', a'')/P^{-}(a', a'') \geqq c \tag{5.7.}$$

or

$$P^{*}(a', a'') = \sum_{i} P_{i} \geqq c \tag{5.8.}$$

where * represents all of the criteria, with their weights, as indicated by the equivalent summation expression in the middle of (5.8.).

Roy also lets the DM express restrictions on the preference ordering by what he calls discordance relationships. For the very lengthy associated mathematics see Roy's cited works or Duckstein and David [1975, p. .9 ff.]. I feel that this is an exceptionally cumbersome version of the separable additive forms that were described in Chapter II. It is nice to be free from the precise weighted sums format, but the cost is a tremendous interviewing task with the DM, and a large investment in algebraic graph theory. It should also be noted that the outranking relationship seems like a crude version of the subsemiorder which is described in Section VI.4.

In short, outranking suffers from many of the problems of weighting which were discussed in Chapters II, III, and IV. In addition, it is not a parsimonious treatment. It has been used mainly in Europe, where the author's influence is strong.

V.3.2 Multidimensional scaling and factor analysis. Multidimensional scaling (MDS) [Shepard, Romney, and Nerlove, 1972] and factor analysis (FA) [Tatsuoka, 1971] are treated here as dimensionally-reducing devices. As in the Principle of Occam's Razor, empirical researchers wish to represent the phenomenon being analyzed as parsimoniously as possible. Often the variables that are observed in nature can be transformed into a better picture of the underlying structure.

A potent form of MDS is nonmetric MDS. In nonmetric data sets the "distance" between objects is measured in less than ratio scales. Authors in this field believe that even from this kind of data a great deal of metric information can be extracted. It is not clear whether the z vectors in MOLP should be deemed metric or nonmetric. That would depend on the scales which were used to form the objective functions. Therefore we will describe a very general form of MDS.

For two objects i and j in a set of n objects, let a datum s_{ij} be the similarity, substitutability, affinity, association, interaction, correlation, or more generally "the proximity" between them. The objective is to discover a way to configure the points in a Euclidean space of lower dimension in such a way that the new interpoint distances d_{ij} are monotonically related to the original affinity measures as follows:

$$d_{ij} < d_{kl} \text{ iff } s_{ij} > s_{kl} \qquad (5.14.)$$

This becomes a nonlinear optimization problem. The extent to which the monotonicity condition is violated is tracked by a measure called "stress."

In MOLP the application of MDS would be to begin with a matrix of distances between every pair of z vectors. The output of MDS would be a representation of N in a new set of criterial coordinates which would capture most of the meaning of the original criteria. Normally MDS is able to reduce a data set down to three dimensions. This would allow the DM to wade through to a final choice with the help of decision supports such as computer graphic techniques.

There are several problems with the MDS approach. It is often difficult to know how much to lower the dimensionality of the data base. More important is the difficulty of naming, or interpreting the new dimensions. The similarity matrix which serves as input to MDS requires an s_{ij} measure for every possible pair (i,j) in the data. Technically we have this as

$$d_{ij} = [\sum_{t=1}^{r} (z_{it} - z_{jt})^2]^{\frac{1}{2}} \qquad (5.15.)$$

when z_i is the i^{th} nondominated vector, z_j is the j^{th} nondominated vector and the subscript t represents the summation over criteria.

But abandoning this technical level of thinking, how meaningful is a distance measure within N (or N^{ex})? Distance concepts were used to make a first pass through the problem, that is to find N (or N^{ex}). But the definition of N is that it is a subset of feasible solutions, which although preferable to the dominated set, is itself not ordered. Within the nondominated set the matrix $\|s_{ij}\|$ is probably not meaningful until new evaluative criteria have been superimposed.

Factor analysis reduces the number of dimensions of the original data by forming linear combinations of variables. These factors are found by various forms of linear transformation, which take the original data points and "name" them according to new coordinates. Algebraically new basis vectors are found to span the vector space. Through an eigenanalysis it is often clear that a p-variate probability distribution is degenerate; it can be represented by less than p coordinate axes (see a multivariate statistics book such as Tatsuoka [1971] for an elaboration).

FA, like MDS, is plagued by the problems of deciding how many factors there are, and how to interpret them. Using the capital budgeting example from Chapter I, are the criteria "percentage of business in unstable lands" and "foreign exchange risk" essentially one factor? FA will help us quantify this question by means of the factor loadings. But what would we call the new factor? These are classic managerial quandaries, and as such we know that similar problems are solved every day in the world of affairs.

The crucial drawback of using FA hers is that the factors are based on first-order linear correlations among the criteria. If the components of the z vectors interact in a higher order, or nonlinear, fashion, then the aggregation into variables would be an informationlosing process. For example, two highly correlated variables could interact with a third in very opposite ways. If FA hooked both variables onto a particular factor, the factor loading of that third variable would be close to zero. The interaction effect would be completely lost. But as we made clear in Chapter III and IV, the consciousness of preference interactions is one of the main points of this dissertation.

V.3.3 Conjoint measurement analysis. Conjoint measurement analysis (CMA) [Shepard, Romney and Nerlove, 1972] is a general theory of psychological modeling. Tversky (quoted in Shepard, Romney and Nerlove [1972]) said that a goal of scientists is to decompose "complex phenomena into sets of basic factors according to some specified rules of combination."

In econometric studies these factors (or independent variables) can be measured independently. Regression analysis, for example, provides coefficients which make a separable additive combination rule work to explain the joint effect of the variables. In some cases we cannot measure the basic factors, or variables. If a man had to choose between buying a Harris tweed sport coat or a pin-striped light weight wool and polyester suit, several non-measurable factors may be involved. To name a few, Harris tweed has status, durability, resistance to wrinkling, informality and warmth. The suit has formality, no problem in choosing what pants to wear with it, is more comfortable in overheated rooms and probably costs more.

This is a messy problem; the two choice items do not even have the same attributes. Nevertheless most men make decisions like this often in their lives. CMA requires only that the ordering implied by the joint effects of the intrinsic factors is known.

The conjoint measurement problem is to reduce these complex choice realities to a set of intrinsic factors and a combination rule which acts to coalesce them into a scale that recalls the known choice order. Thus the conjoint measurement problem is to find a theory of behavior.

The combination rule is a polynomial function which could be one of many forms, such as:

$$d_{ij} = [\sum_{t=1}^{n} |x_{it} + x_{jt}|^c]^b \qquad (5.16.)$$

$$d_{ij} = [\sum_{t=1}^{n} |x_{it} - x_{jt}|^c]^b \qquad (5.17.)$$

$$d_{ij} = [\sum_{t=1}^{n} (x_{it} + x_{jt})^c]^b \qquad (5.18.)$$

$$d_{ij} = [\ \overset{n}{\underset{t=1}{}}\ (x_{it} - x_{jt})^c]^b \qquad (5.19.)$$

$$d_{ij} = [\sum_{t=1}^{n} (x_{it}\, x_{jt})^c]^b \qquad (5.20.)$$

$$d_{ij} = [\sum_{t=1}^{n} |x_{it}\, x_{jt}|^c]^b \qquad (5.21.)$$

In these expressions b and c are abribrary exponents; particular choices of b and c and t can be shown to force attractive interval properties onto the scales which measure the t criterial dimensions. Equation (5.16.) is called the absolute additive combination rule, (5.17.) is the absolute difference combination rule, (5.18.) is the additive combination rule, (5.19.) is the difference combination rule, (5.20.) is the multiplicative combination rule, and (5.21.) is the absolute multiplicative combination rule.

To summarize, rank order data is processed using regression-type mathematics. What surfaces is a set of scales for the various factors, and a combination rule that minimizes a stress measure. It is clear now that CMA is a generalization of MDS. Its data requirements, though, are less stringent.

Two difficulties with CMA are the need for rank-ordered input data, and monotone utility functions on each of the criteria (in this section the word factors was used). Unfortunately, there is no evident rank ordering of N. Chapter VI will atempt to rank a subset of N, namely the generators $g \varepsilon G(N)$, in a much simpler fashion. Chapter III mentioned the lack of monotonicity that we often find in the individual utility functions. For these reasons it seems premature to apply CMA to prune N.

V.3.4 Zeleny's ADAM model. Ideal points are introduced in Zeleny [1974 (a), p. 171]. Subsequently [Zeleny, 1975(b) and 1976(c)] the author switched his focus to a theory of individual choice behavior which he called the attribute-dynamic model (ADAM).

He begins by citing an article which reviews what he calls the compensatory multi-attribute model. This is of the form:

$$A_j = \sum_{i=1}^{m} \lambda_i \, d_{ij} \qquad (5.22.)$$

where A_j is an individual's attitude toward brand j, m is the number of salient attributes, λ_i is a weight showing the importance of these attributes, and d_{ij} is the perception or score of the j^{th} brand along the i^{th} attribute. Of course this is identical, except for the terminology, to the separable additive form discussed in Chapter II. After translating this into the language of MCDM, Zeleny states three "failures" of this traditional model.

First he shows how several sets of weights lead to the same decision. This means that the weights in (5.22.) have not retrieved, in any unique way, the DM's mental procedures (cf. an allied result in Chapter III). Next we are shown how a good set of weights can lead to the wrong decision, when they are elicited independently of the relevant feasible set of alternatives. Finally, in what he calls the "fatal failure," the case of a convex approximation of a non-convex choice set is treated. The conclusion is that there are situations in which no set of weights would predict correctly.

Zeleny also reviews several mathematical programming models for the determination of the weights. He makes the interesting comment that these research methods

> "bootstrap" themselves into the attribute weights. That is the differential weights are not related to any intrinsic attribute properties in a given decision situation, but rather they are internally computed to satisfy the minimization of the Euclidean metric. Decision makers are presumed to use the weights so as to minimize a particular function. Consequently, any other distance metric used would imply a different set of weights under otherwise equal conditions. [Zeleny, 1976(c), p. 17]

For these reasons, combined with other thoughts on weighting from Chapters III and IV, we will review ADAM. Here Zeleny tries to reproduce the iterative dynamics and the context-dependency of individual choice behavior (cf. Tversky's EBA model which was partially motivated by "structural dependence of the alternatives").

Form a vector of attribute scores $d_j = (d_{ij}, \ldots, d_{mj})$ in which d_{ij} positions the j^{th} alternative with respect to the i^{th} attribute. A matrix $D = \|d_j\|$ is formed from these vectors. The set $\{d_j\}$ implies, for each attribute, the maximum attainable level. This anchor value is

$$d_i^* = \operatorname*{Max}_{j \varepsilon D} d_{ij} \qquad (5.23.)$$

and associated with any D is (overall) anchor

$$d^* = (d_1^*, \ldots, d_m^*) \qquad (5.24.)$$

If $\exists\ j' \ \varepsilon\ D \ni d_j = d^*$, then there is no decision problem. The ideal of all the attributes can be simultaneously achieved in the feasible region. Normally, though, the anchor point is not feasible.

Definition V.3.4.1 The axiom of choice. Alternatives that are closer to the anchor are preferred to those that are farther away. To be as close as possible to the perceived anchor point is the rationale of human choice.

Thus between any anchor point and any alternative d_j', there is a distance. Under various proposed metrics (not very different from those in Chapter V) the DM experiments with the "closeness" of several alternatives. The word experimentation is used for several reasons. First, there are variations of the distance functions which test the stability of the proximity to the anchor point. Second, some shuffling of the weights $\{\lambda_i\}$ is suggested.

Definition V.3.4.2 The attention level. An attention level λ_i, is assigned to the i^{th} attribute as a measure of its relative importance for a given decision situation, is directly related to the average intrinsic information generated by the given set of feasible alternatives through the i^{th} attribute, and, simultaneously, to the subjective assessment of its importance, reflecting the decision maker's cultural, psychological and environmental history.

Let us now use the symbol λ_i to be the attention level, which has two components. In other words, this will supercede, and hopefully refine, the concept of weights.

Zeleny decomposes these weights into two influences. There is one, noted w_i, which is quite steady because it reflects the importance of an attribute that stems from a person's make-up and background. The second component is an attention level, $\tilde{\lambda}_i$, defined as situationally dependent. It is a function of each problem's structure. Therefore, the $\tilde{\lambda}_i$ can change radically with changes in the average intrinsic information generated by D.

For example, in the capital budgeting example of Chapter I, say each capital investment item scored identically according to one of the criteria. Then that criterion would generate no ranking information; even if it were the most important attribute, the attention would naturally shift to another. We can interpret $\tilde{\lambda}_i$ as a measure of the <u>contrast intensity</u> of the i^{th} attribute, noted as $e(d_i)$.

Next define

$$D_i = \sum_{j=1}^{n} d_{ij}, \quad i = 1, 2, . . . , m \tag{5.25.}$$

and, for a finite set D, an entropy-like measure

$$e(d_i) = -k \sum_{j=1}^{n} (d_{ij}/D_i) \ln(d_{ij}/D_i) \tag{5.26.}$$

when $k > 0$ and $e(d_i) \geqq 0$. For the case where all d_{ij}'s are equal to each other for some i, $d_{ij}/D_i = 1/n$, and $e(d_i)$ is maximized at that point.

What comes from this is the possibility of moving the anchor point. Deleting an attribute which had low contrast intensity would change all the relative lengths of the $\{d_{ij}\}$; in like fashion dropping one brand which was the only outlier along some dimension would displace the anchor (also known as an ideal point).

It is precisely this dilation and contraction of the distances from the anchor, and the changing locus of the anchor itself, which Zeleny calls the decision dynamics. His claim is that multiattribute utility theory is not relevant because it is static and confuses the assessment of risk and uncertainty with the elicitation of preferences. As attributes gain or lose attention, "A computer-aided, dynamic, self-adjusting, interactive and iterative procedure, based on a man-machine interface, emerges" [Zeleny, 1976(c), p. 24].

The ADAM paradigm should be studied as a means of pruning N. It may suffer from the same cumbersomeness of Roy's outranking relationship. But some DMs may find that so much of the work is done by the computer that ADAM is stimulating to use. The main problem that I see is that each criterion (attribute) is examined essentially in isolation, although the entropy measure $e(d_i)$ could be considered a type of interaction effect. Although Zeleny normalizes the d_{ij}, I do not see how to incorporate incommensurability and compensatoriness in his model.[4]

<u>Conjecture V.3.4.1</u>. ADAM measures (transformed) distances from the anchor point; it neglects the ratios of distances over subsets of criteria.

---

[4] The particular sense in which the words incommensurability and compensatoriness are used can be found in Chapter VI.

To illustrate this point, consider the story of the Magic Mirror. By selling his soul to the Devil the basketball coach of the local university has obtained the use of the famous Magic Mirror. When you look into this mirror any (single) wish will be granted.

One of the ball players, John, has height, coordination and great reach. But he is myopic, so his shots, even the attempts at dunking the ball, fail. The coach gives John one trial with the mirror. John says, "Mirror, mirror, on the wall, cure my corneal curvature once and for all." The wish is granted and John becomes the team's leading scorer.

Previously Jack was the leading scorer, even though he is rather short for a ball player. Seeking to regain his prominence in scoring, Jack addresses the mirror, "Mirror, mirror, on the wall, give me the reach to dunk the ball." His wish was granted. Jack's arms were tripled in length, and he was now able to dunk the ball without even jumping. Unfortunately, his hands now dragged on the ground, which made his fellow teammates trip on them. It also made ready-to-wear clothing impossible to find; his adoring female fans now thought he was grotesque.

The moral of the allegory is that the ADAM model treats the distances from the anchor point criterion by criterion. But if attributes are misspecified, or if value interdependence is not recognized, Jack's disaster can occur. What he really wanted was a simultaneous change in several of his characteristics, such as his leg length, torso length and jumping ability, Certain ratios of characteristics, if properly specified, would have achieved the scoring ability he longed for.

V.3.5 Scaling with eigenvalue analysis. Thomas Saaty [1975 and 1977] begins with the premise that a basic problem of decision theory is to find out the importance of weights of a set of activities. In his view importance is usually judged according to several criteria. He constructs a matrix of pairwise comparisons of choice objects (activities). In each cell is a number from a scale which tells not only the direction of the preference, but also its intensity.

For n objects, $n(n-1)/2$ pairwise comparisons must be done. Using the scale 1 through 9 (a lengthy justification of this particular scale is in Saaty [1977, p. 244 ff.], this is done once for each criterion. Saaty's scale is as follows:

Table 5.7

Saaty's Priority Scale

Intensity of Importance	Definition
0	Not comparable
1	Equal importance
3	Weak importance of one over another
5	Essential or strong importance
7	Demonstrated importance
9	Absolute importance
2,4,6,8	Intermediate values between 1,3,5,7 and 9
Reciprocals of the above numbers	If activity i has a non-zero number that represents its dominance over ij, then j over i is assigned, by fiat, the reciprocal number.

The method uses the numbers of Table 5.7. Two kinds of comparisons are made--choice objects versus each other according to a criterion, or all the criteria compared with each other. The following mathematics relates to either type.

Let $A_1, A_2, \ldots, A_n$ be the choice objects and $w_1, w_2, \ldots, w_n$ be their weights as in Table 5. . The comparisons are collated in the matrix:

$$A = \begin{array}{c|cccc} & A_1 & A_2 & \cdots & A_n \\ \hline A_1 & w_1/w_1 & w_1/w_2 & \cdots & w_1/w_n \\ A_2 & w_2/w_1 & w_2/w_2 & \cdots & w_2/w_n \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ A_n & w_n/w_1 & w_n w_2 & \cdots & w_n/w_n \end{array} \qquad (5.27.)$$

A is called a reciprocal matrix, because all its entries are positive, and $a_{ji} = 1/a_{ij}$. A can be post-multiplied by the column vector w, which yields

$$Aw = nw \qquad (5.28.)$$

which is computable, but not very interesting if we know the weights in advance. But if we only knew A and wished to find w, the task would be an eigenvalue problem similar to the one in Chapter IV. We are now interested in the unknown vector w in

$$(A - nI)w = 0 \qquad (5.29.)$$

This has a positive solution if and only if n is one of the eigenvalues of the linear transformation A. Since every row of A is a multiple of the first row of A, A has only one linearly independent column. This means that the rank r(A) = 1; hence only one of the eigenvalues λ_i, i = 1, 2, . . . , n of A is non-zero. Reviewing a theorem from Chapter IV,

$$\sum_{i=1}^{n} \lambda_i = tr(A) = n \qquad (5.30.)$$

where tr means the trace of a matrix. Since all but one of the λ_i are zero, then the one positive one must be equal to n, and will be noted λ max.

The solution to (5.30.) is any column of A, all of which differ from each other only by a multiplicative constant. But any column of A, if normalized, will be a unique scale of weights on the choice objects recovered from the cells $\|a_{ij}\|$, the dominance ratios.

A has what can be called the <u>cardinal consistency</u> property, which is that

$$a_{ij}\, a_{jk} = a_{ik} \qquad (5.31.)$$

Any row of A will generate all the other rows by this property. Equation (5.31.) is associated with a preference ordering when a complete transitive ordering is evident. But because of finite human sensibility, transient tastes, and the arrival of new information, A matrices that are built up by interviewing the DM often lack the cardinal consistency property.[5] The DM is inconsistent if he states that $A_i > A_j$ and $A_j > A_k$ but $A_k > A_i$. Because of the element of randomness in human choice, and because of all the aspects of interdependency discussed in Chapters III and VI, inconsistency is a managerial reality. It makes allocative decisions more difficult, though, and we are motivated to study it in more detail.

If in $Aw = nw$ one perturbs (slightly) the a_{ij}, then there are small perturbations of λ_i as a consequence. (5.28.) becomes

$$A'w' = \lambda \max w' \qquad (5.32.)$$

The Perron-Frobenius theorem says that a matrix of positive entries has a real positive eigenvalue (of multiplicity 1) whose modulus exceeds those of all other eigenvalues. The eigenvector solution is nonnegative; some of the other eigenvalues are not necessarily real.

Does this mathematical apparatus help us track the inconsistency of the reciprocal matrix A? If the DM estimates w_1/w_2, for example, is there some difference between his expressed preference ordering and his intrinsic (perhaps never to be found) ordering? Using the developed notation, how far is λ max from n and w' from w? If they are not close we might ask the DM to allocate more of his/her time to the estimation of the a_{ij}. Of course we know that in the pruning

---

[5]This is related to a more thorough discussion of a subsemiorder in Chapter VI.

stage of MOLP a new criterion is endurance. We can only push the DM up to the point of information stress. It is hard to understand exactly what we would be achieving by this extra DM-analyst interface. The DM need not be consistent. All we can say is that the "sample" A is more likely to be a set of logically related entities than a random set of vectors. Obviously this is reminiscent of classical hypothesis testing.

Theorem V.3.5.1. A n X n matrix A with $a_{ji} = a_{ij}^{-1}$ is consistent iff $\lambda max = n$.

Proof. See Saaty [1977, pp. 239-240].

Because of this theorem we know that in the perfectly consistent case $\lambda max = n$. Otherwise, with inconsistency $\lambda max > n$.

Theorem V.3.5.2 Preservation of ordinal consistency. If $(O_1, O_2, \ldots, O_n)$ is an ordinal scale on the activities $C_1, C_2, \ldots, C_n$, where $O_i \geqq O_k$ implies $a_{ij} \geqq a_{kj}$, $j = 1, 2, \ldots, n$, then $O_i \geqq O_k$ implies $w_i \geqq w_k$.

Proof. It is true from $Aw = \lambda max\ w$ that

$$\lambda max\ w_i = \sum_{j=1}^{n} a_{ij}w_j \geqq \sum_{j=1}^{n} a_{kj}w_j = \lambda max\ w_k$$

with $w_i \geqq w_k$.

Because of this theorem, the expression $(\lambda\ max - n)/(n - 1)$ can be used as an index of the consistency of A. It can also be thought of as the reliability of the DM's estimate w_i/w_j. With human fallibility the estimate of w, namely w', will be $(w_i/w_j)(\varepsilon_{ij})$, with $\varepsilon_{ij} > 0$.

The particular generating function of ε_{ij} is dependent on the psychological model of the DM's thinking processes. Several of these were reviewed in Chapter III.

Saaty shows at length that the perturbations of the weight vector $dw = (dw_1, dw_2, \ldots, dw_n)$ are complicated nonlinear functions of the errors in the a_{ij}. He develops a single index of inconsistency, and a statistical test of its significance. Speaking managerially, inconsistency is not wrong; it is a manifestation of imperfect information processing ability. If all the problems of incommensurability, compensatoriness, nonmonotonicity and interdependence could be solved, the DM's preference order would be consistent. The measure of inconsistency tells us approximately the level of "damage" due to those problems.

As for the use of Saaty's method for pruning the nondominated set, we now state some reservations. The number of pairwise judgments that must be made, for each criterion or set of objectives, is $n(n-1)/2$. This, I feel, distracts the DM from the important tasks of forcing commensurability, compensatoriness and interdependency onto his evaluation of the choice set (see Chapter VI for an expansion of this thought). The sheer number of comparisons to be made tends to lower the quality of managerial thought. Chapter VI suggests a way to reduce the number of such comparisons.

Perhaps Saaty's search for ratio-scaled weights is wrong. His methods force the non-simultaneous consideration of criterial levels, ignoring interactions and local behavior in particular ranges of the individual preferences for each attribute.

On the plus side, it is very attractive to get at strength of preference in this quantitative way, and to have available an index of intransitivity. Philosophically, it may be wrong to assume that the zero representing incomparability is the weakest sort of information in the A matrix. It could be said that the indifference relationship expresses the most razor-sharp sort of preference comparison we can elicit. It represents a very precise statement, not merely an ordinal (fuzzy) assessment. A very intriguing statement appears on page 73 of Saaty [1975]. He states that incomparable activities may inherit a relationship via intermediate activities; the powers of the A matrix, as the limit is approached, generate information of this sort. But the mechanistic approach to decision support systems may infer too much from fuzzy input data such as the $\{a_{ij}\}$. Sophisticated mathematics sometimes makes us too confident in an analysis.

Summarizing Saaty's scaling method for priorities, from a matrix of paired comparisons of activities, for each objective, activity weights are found by solving an associated eigenvalue problem. Then from a pairwise comparison matrix of objectives, which also has an eigenvalue problem, Saaty obtains a vector of weights on the objectives. Now aggregation across the activity-by-criterion matrices is possible by using those objective weights; finally the composite priority vector emerges which ranks the set of activities.

Chapter V has treated the reduction of the nondominated set. The larger number of vectors in N (or N^{ex}) has been seen as a drawback to MOLP methods, since the DM may not have time or ability to sift through thousands of multiattributed items. A subset $G(N) \subseteq N^{ex} \subset N$ was identified by a data analysis technique known as cluster analysis. Many clustering algorithms were utilized; after analysis hierarchical clustering with RELOCATE and RESIDUE capabilities was recommended.

Five other methods of pruning the nondominated set were analyzed. Roy's outranking relation was considered cumbersome. Multidimensional scaling, factor analysis and conjoint measurement analysis were seen to have their assumptions violated. Saaty's scaling methods used eigenanalysis, which was discovered independently in Chapter IV. This pruning method warrants further study. Finally, Zeleny's ADAM model was treated as a promising theory of choice behavior. This dissertation decomposes the problem of pruning into two stages. In Chapter V, by mechanistic methods, the generator set G(N) is created. This list of possible solutions is now analyzed by interactive programming in Chapter VI.

CHAPTER VI

RANKING: CHOOSING ONE OF THE GENERATORS

VI.1 Introduction

At this stage the nondominated set has been calculated and pruned down to a subset called the generators. The action implied by decision making can commence as soon as one element of the set of generators is chosen. To accomplish this, ranking and weighting will be examined in the limited context of this final stage of multiobjective mathematical programming. Binary judgments will be the means the DM uses to search his preferences. That information will be collected and displayed by a linear programming formulation that is similar to the definition of nondominated vectors. A significant reduction in the number of paired comparisons needed to solve the second stage of the pruning algorithm is achieved.

It has been said that the multiargument utility function is transitory, dependent on the feasible set, and cannot immediately be expressed by the DM [Zeleny, 1974(a)]. If such a function could be abstracted, its form would be extremely complicated if any interdependence among criteria existed. The measurement error inherent in value assessment and evaluation is one type of randomness that we see. Another source is the stochastic nature of human response to cues. Thus we should view multicriterion decision making as a bi-stochastic process in which representations of the problem are to be traded back and forth between the human(s) and the computer-based model(s).

In Chapter IV, weights were characterized as naive or non-naive. It was shown that a priori ranking did not always drive the system in a desirable direction. The weighting and ranking in this chapter will be exploited gingerly. The particular setting of Chapter IV was a constrained problem in which goals were monitored as absolute deviations from targets. The final choice within the generator set G(N) is an unconstrained problem. Ranking is possible because we will use binary preference relationships between nondominated solutions instead of dealing directly with objectives.

A linear programming (LP) model is proposed in which weights and ranks will serve to identify a final choice and to build confidence in that choice. We will proceed from the ideas of Kornbluth [1977]. In his paper a ranking method is studied in the unconstrained case. Kornbluth also suggests that the methodology could be used as a first stage in a multicriterion decision process. He points out that in this manner one can develop the interval criterion weights that Steuer [1975] uses as an a priori way of pruning N (or N^{ex}). The Kornbluth model deserves to be expanded to the case of interdependency and noncompensatoriness, and to be applied as an appendage to the multicriterion simplex methods.

We have argued in Chapter IV that weights have no intrinsic meaning. Their effect depends on the type of solution methodology, the feasible set, and the experience of the DM. However, weights do affect the system in some way. Since that effect is, in general, not known with certainty, we are going to use weights as knobs. Imagine a person playing with unmarked knobs on a color television. He does not know the exact influence of each twirl of a knob, and he is even less sure of the total effect of several knobs turned at once. But by trial and error he learns a setting (or a set of settings) that is best for him at that time. Given the possibility of changes in taste, and different atmospheric and ambient lighting conditions, those settings may need to be modified periodically. The learning of the original knob turning process provides a reservoir that hastens the updating process.

Let us place the concept of knobs into the realm of mathematical programming. A set of weights corresponds to a set of consequences. These consequences, when dealing with the generator set, are really ranking implications, or restrictions, on the possible rank orders of generators. For example, imagine a college professor who is seeking another teaching job. He/she is negotiating with two schools. One offers a salary of \$20,000 and a nine-hour load. The other school offers a \$30,000 salary and a twelve-hour load. If the professor prefers the first school, he/she is expressing a feeling that increased salary does not compensate for increased teaching load. This is known as noncompensatoriness.

Imagine that the job applicant continues to interview at other institutions. If we model that person's choice process on a series of paired comparisons, the first preference ordering has implications. A likely one is that in the pair [(\$30,000, 9 hours), (\$40,000, 12 hours)] the first component would be chosen. The importance of developing restrictions on the possible preference orderings of objects in a choice set is that much less information must be gathered to form a complete order.

Since all we seek is a final choice from G(N), the generator set, why do we need a complete order? First, a complete order allows us to test all the implications of a first choice. Second, the process of examining the implied order of choice objects increases our confidence in the action that the whole analysis is leading to.

To summarize this section, think of a set of weights, $k \varepsilon K$, $k = \{k_1, k_2, \ldots, k_n\}$, and a set of orderings $o_i \varepsilon O$, $i \varepsilon I$. The orderings are simply ordered lists of the generators $g \varepsilon G(N)$. These ideas will be useful in the ranking algorithm of the following section.

VI.2 The Ranking Model

Within G(N) there are m vectors, each of which represents a cluster or comes from the residue set R(N). Each vector is written as $z^k = ||z_i||$, $k = 1, 2, \ldots, m$ where $i = 1, 2, \ldots, n$ represents the achievement along the i^{th} criterion for each z vector. Recall that z^k is the solution vector $x \varepsilon N$ as it is mapped into criterion space.

Let $k \varepsilon K = \{k_i, i = 1, 2, \ldots, n\}$ be the weights on criteria. Since we are using weights as knobs, the weights are initially unspecified, but almost any set of weights could serve as an initial position. Weights will imply that for z^p and z^q, and $k \varepsilon K$, $k(z^p - z^q) > 0$ implies that z^p is preferred to z^q.

Let O be the set of permutations of the numbers 1 through m, and let $o_i \varepsilon O$ be the first ordering of the z^k to be dealt with. For example, the initial ordering of $z^k \varepsilon G(N)$ could be obtained from any one of the n criteria, creating a simple rule to break ties if necessary.

If an ordering $o_i \varepsilon O$ is judged correct by the DM, then a set of weights $k \varepsilon K$ is implicitly accepted. Preference functions composed of such weights are not unique; often several can lead to the same ordering. This is similar to what Keeney and Raiffa [1976, pp. 88, 144] call <u>strategic equivalence</u>. If the DM chooses to think directly in terms of weights, the preferred weights generate the preferred ordering. For the proofs behind these ideas see Kornbluth [1974 and 1978] or Zeleny [1974(a)].

Kornbluth has developed a methodology for systematically moving from o_1 to a final order o_n. This is possible because any order $o_i \varepsilon O$ partitions the weight space $k \varepsilon K$ as follows:

$$K = \bigcup_{o} K(o)$$

$$K(o_i) \cap K(o_i) = \emptyset, \; i \neq j$$

where $K(o_i)$ means the strategically equivalent weights associated with an ordering o_i.

To discover $K(o_i)$ we must consider constraints formed by adjacent members of o_i. Non-adjacent pairs form constraints which are redundant. Let $z(i)$ be the z vector in the i^{th} position of the current order; then from the relations

$$k[z(i) - z(i+1)] > 0, \; i \varepsilon I, \; k \varepsilon K \tag{6.1.}$$

a linear constraint set can be developed. $K(o_i)$ is a function of a subset of those constraints, namely the tight ones for which (6.1.) could not hold as an equality. For example, if nondominated solutions z^p and z^q are adjacent in an order o_i, then $k[z(i) - z(i+1)] \geqq 0$, $i \in I$, $k \in K$ will be a tight constraint. This will be clearer as the algorithm is explained below.

If the DM switched his ranking from (z^p, z^q) to (z^q, z^p) he would be moving from a weight space $K(o_i)$ to an adjacent space $K(o_j)$. If o_i is feasible, o_j will also be feasible, since it is associated with an adjacent basis of a dual problem.

In summary, the ranking algorithm is:

(i) Find an initial feasible order $o_1 \in 0$.

(ii) Calculate the binding constraints of $K(o_1)$.

(iii) Give the DM the set of pairs in o_1 which imply the boundaries of $K(o_1)$. The form in which the pairs will be presented is the z vector format, i.e., the criterial achievement associated with each efficient solution in the generator set G(N).

(iv) If o_1 is considered correct by the DM, stop. If it is not correct, go to step (v).

(v) Instruct the DM to change the ranking of one of the pairs presented in step (iii) and go to step (ii).

Because of the property that adjacent orders remain feasible, there is no danger that any of the DM's pairwise choices will force cyclicity or intransitivity into the orders that unfold.

The ideas of dyadic and triadic judgments have been used for decades to elicit preference functions. The unusual property of the present method is the reduction in the number of paired comparisons to be made. The DM need only be concerned with the pairs of elements that are identified by the binding constraints of the form (6.1.). Pairs of efficient vectors that have to do with the slack constraints of $K(o_1)$ are properly ordered as a consequence of the ranking of the crucial parts.

In Chapter III the difficulties of human information processing were briefly stated. All the problems of transitory preferences, imperfect discrimination, fatigue and constrained managerial resources are exaggerated when $m(m-1)/2$ paired comparisons are made. Kornbluth's 1978 paper reports simulations of the ranking algorithm's performance. For example, for six criteria and twenty choice objects only seven paired comparisons were necessary to find the best order and the associated weight space.

VI.2.1 The mathematics of the ranking algorithm. Begin with a feasible order, sich as the lexicographic order that follows from ranking by any one of the criteria. We have notated this as $o_i \varepsilon O$, $i \varepsilon I$. Form the matrix A_o whose rows $A_o(i)$ are quantified by $z(i) - z(i+1)$. The closure, or the convex hull of the weight space $K(o_i)$ is the set $\{k\}$ such that:

$$A_o k \geqq \underline{0},\ \underline{0} \text{ a column vector of zeroes,} \qquad (6.2.)$$

$$\Sigma k_i = 1,\ i \varepsilon I \qquad (6.3.)$$

$$k_i \geqq 0, \ i \in I \tag{6.4.}$$

There are rows of (6.2.) that are satisfied as equalities for some $k \in K$. For these boundary rows of the convex hull of A_o, the value of an objective function, "minimize $A_o(i)k$ subject to (6.2.), (6.3.), and (6.4.)," is zero. This means that the optimum of the dual problem (formed as 6.5. below) is also zero. In general, m, the number of choice objects (z vectors in this context) will be greater than n, the number of attributes (criterion achievment). This is a reasonable assumption; therefore we consider the dual problem:

Maximize ρ

$$\text{Subject to} \quad \mu A_0^T + \underline{1}\rho \leqq c \tag{6.5.}$$

$$\mu \geqq 0, \ p \text{ unconstrained}$$

T indicates the matrix transpose, $\underline{1}$ is a column vector of ones and c is a particular row of A_o such as $A_o(i)$.

For the primal problem (6.2., 6.3., and 6.4.), if a row is satisfied as an equality, then that row used as the objective function yields an optimum of zero. Any time that a basic feasible solution of (6.5.) has $\rho = 0$, it must therefore also be a dual optimum. A strictly positive dual variable identifies the affiliated primal constraint $A_o(i)k \geqq 0$ as binding. In turn this means that the paired items z(i) and z(i+1) which formed that constraint are binding to the ordering $o_i \in O$.

For these reasons computational savings occur. To ascertain the crucial (binding) pairs in o_i, it is not necessary to solve all the related dual problems of the form (6.5.) with each row of A_o as the right-hand side in turn. If the k^{th} column is basic at an optimum of (6.5.) the pair of choice objects (k, k+1) will be binding. For the typical case of m much greater than n, (6.5.) has many more columns than rows, and hence is not too costly to solve.

The DM's binary (forced) choices among $g \in G(N)$ have led to restrictions on the orderings. The implied (through duality) weights $k \in K$ can be used to improve confidence in results of the ranking algorithm. They should not be used as a priori weights for a new problem, for the reasons explained in Chapters III and IV. They serve as a posteriori weights in the limited case of the generator set for a problem that has been attacked by multiobjective mathematical programming and the two-stage pruning algorithm of this dissertation.

VI.2.2 A numerical example. Kornbluth's example will be translated into the notation of this paper. The seven cho ce objects, each with three criterial dimensions (i.e., m = 9, n = 3) are as follows:

Item	The z vectors		
1	8	1	4
2	6	4	2
3	5	8	0
4	4	0	6
5	3	2	6
6	2	7	1
7	2	4	3

Start with a feasible order $o_1 = (1,2,3,4,5,6,7)$. Then there is an $(m \times n)$ A matrix:

$$A_{o_1} = \begin{bmatrix} 2 & -3 & 2 \\ 1 & -4 & 2 \\ 1 & 8 & -6 \\ 1 & -2 & 0 \\ 1 & -5 & 5 \\ 0 & 3 & -2 \end{bmatrix}$$

Observe that A_{o_1} is obtained by subtracting the values assumed by the z vector for item 2 from the z values of item 1 in the table above.

A series of LP problems of the form of (6.5.) must be solved to construct the convex hull of the weight space $K(o_i)$. Transposing A_{o_1} we get:

$$\begin{aligned} &\text{Maximize} && \rho \\ &\text{Subject to} && 2\mu_1 + \mu_2 + \mu_3 + \mu_4 + \mu_5 + \rho \leqq c_1 \\ &&& -3\mu_1 - 4\mu_2 + 8\mu_3 - 2\mu_4 - 5\mu_5 + 3\mu_6 + \rho \leqq c_2 \\ &&& 2\mu_1 + 2\mu_2 - 6\mu_3 - 0\mu_4 + 5\mu_5 - 2\mu_6 + \rho \leqq c_3 \\ &&& \mu_i \geqq 1,\ i = 1, 2, \ldots, 6 \\ &&& \rho \text{ unconstrained} \end{aligned} \tag{6.6.}$$

with $c = [c_1, c_2, c_3]$ set equal, in turn to the rows of

For example, if you take the first row of A_{o_1}, then $c = [2, -3, 2]$ and we find that at the optimum $\rho > 0$. At this point μ_2 and μ_4 are in the basis. Since they are also set at a non-zero level, we know that the pairs of items (2,3) and (4,5) are crucial to the present ordering. Since μ_1 is not basic and strictly positive, the pair (1,2) is not crucial, and the first constraint of A_{o_1} is not binding.

Here is where we start to see the computational savings. Since we know that (2,3) is a crucial pair, there is no need to solve the problem associated with the second row of A_{o_1} which in the dual problem would be c = [-1, -4, 2].

For the dual problem with c = [-1, 8, -6] the optimum ρ is again strictly positive. An analysis of the basic variables shows that the crucial pairs are (2,3), (4,5), (5,6), and (6,7). Using P to denote a binary preference relation, we have

$$o_1 \in O = \{2P3,\ 4P5,\ 5P6,\ 6P7\}.$$

With these four binary preferences labeled as crucial, the relations 1P2 and 3P4 are also valid. But these relations, which can be seen by inspecting the first feasible order o_1, follow as a consequence of the four parts. The DM, with his limited time and attention, should decide only whether or not he can live with the four binding pairs.

Suppose that the DM finds he does not agree that item 2 is preferred to item 3. This means that o_2 = (1, 3, 2, 4, 5, 6, 7) which is a move across constraint (2,3) to the adjacent weight space $K(o_2)$. In this space the binding pairwise preferences are (3,2), (4,5), and (5,6).

Geometrically, in this three-dimensional example, each weight space associated with an order can be marked off on the triagle $K = \{k \mid \sum_{i=1}^{n} k_i = 1,\ k_i \geqq 0\}$. Algebraically, the dual variables of (6.6.) solve the primal problem that had the weights k ε K as its decision variables. The problem of the form (6.2.) through (6.4.)

implies the dual problem. This dual (6.6.) must be solved between one and (m-1) times, using rows of A_{o_1} as the RHS. The process traces out the convex hull of the weight space $K(o_i)$. The DM can see not only the crucial pairs, but can determine if he likes the implicit range of weights.

We know from Chapter IV that weights often behave counter-intuitively in a goal programming context. But in the relatively restricted domain of discourse of Chapter VI we work with binary preferences on the nondominated solutions in G(N). There is no criterion matrix and hence no troubling intercorrelations.

As the summary of the ranking algorithm at the end of VI.2 states, at this point the DM can make further pairwise switches, one at a time, or he can stop. At the stopping point, a maximal element in G(N) has been identified. This was the goal of the present chapter--namely choosing a final solution from the generators of the nondominated set.

VI.3 Further Justification of the Ranking Algorithm

Four issues will be mentioned again: incommensurability, compensatoriness, interdependence and the introduction of new evaluative criteria.

Section VI.1 defined compensatoriness. Chapters I and II mentioned incommensurability. We stated that decision making forces, or emulates, commensurability. We cannot add apples and oranges; but even the addition of apples to apples requires the fiction of fungible apples.

Chapter II also described forcing commensuarbility by means of fines in the case of the polluting firm.

Similarly, noncompensatoriness,[1] although very real, is essentially overcome by the act of picking a strategy. The phenomenon of noncompensatoriness arises primarily from the attempt to describe it with mathematical functions. In heuristic, intuitive or imitative decision making, the DM focuses completely on the attributes for which there are no compensations. But within certain ranges of the attribute variables the attention shifts (cf. the state of mind of Tversky's elimination-by-aspects model as described in Chapter III). In the example of the various salary-teaching load combinations, it is likely that the DM we described would be willing to teach a twelve-hour load for a salary of $500,000 per annum. There seem to be domains of noncompensatoriness, separated by a discontinuity which leads abruptly (and stochastically) into a domain of a gross sort of compensatoriness.

It is the discontinuity which we find hard to model explicitly. This need not preclude the scientific study of decision making. Just as the scales created by multidimensional scaling recapture much of the ranking that was apparent in the input data, there are separable additive utility functions that can represent certain noncompensatory structures. Keeney and Raiffa [1976, pp. 88, 144] call this <u>strategic equivalence</u>.

---

[1] Noncompensatoriness is defined in the example of the college teacher in VI.1.

The third point is interdependence of the multiple criteria. This was described in Chapters III and IV. In particular the works of Peter Farquhar [1974(a), 1974(b), 1975, and 1977] were referenced. These papers show that to consider interaction effects is, at best, very complicated.

The final point to be reviewed is the introduction of new evaluative criteria into the problem. Consider another example of a professor seeking a teaching position. He/she is a marine biologist, and therefore begins by ranking the schools in descending order of their proximity to the ocean. Soon it is apparent that some unknown coastal community college outranks, under this decision system, the University of Chicago. Hoping to overcome this distortion the biologist adds a new criterion--quality. This process discloses and searches for new information.

These four concepts have been recalled to help justify the ranking algorithm of Chapter VI. In early chapters of this dissertation ranking (called prioritization in Chapter IV) of objectives was shown to be inappropriate in the goal programming context. I conjecture that similar problems would be found if eigenanalysis of the criterion matrix in the multiobjective linear programming were performed. This is a problem with the *a priori* reduction of N^{ex} that Ralph Steuer talks about in his paper on interval criterion weights programming [Steuer, 1975].

The ranking algorithm of Chapter VI is an *a posteriori* method of reducing G(N) (or even N or N^{ex} in the case of small problems). Laboring over the limited number of binary preference decisions seems to be one of the best ways to treat four of the major problems of multi-criterion choice that I reviewed a few paragraphs above. In the non-global situation of ranking two elements in G(N), the human brain is very subtle.

To illustrate my contention that the four named problems can at least be approached in this non-formal way, we will use the capital budgeting example from Chapter I.

The choices C > A and D > B indicate that at any cost the firm wishes to avoid operating in South Africa. No other attribute of the set of projects compensates for that.

Commensurability is lacking in this example. The four criteria of choice appear in the form of dollars, percentages, and two different nominal scales. But the DM can nevertheless make paired comparisons. As for interdependency, the percentage of business in unstable lands is not value independent of foreign exchange risk. Again this can be handled in the forced choice representation of Chapter IV. Of course programming formulations of this problem are also possible (see Bernardo and Lanser [1977] and Morse [1978]). This example does not deal with the discovery of new evaluative criteria, but we recognized that phenomenon in the situation of the marine biologist.

VI.4 Summary and Cautionary Remarks

The ranking algorithm may force a complete ordering onto G(N). Two reasons suggest that the order may be less than complete. It could, for instance, be a semiorder or a subsemiorder [Ng, 1977]. These terms are now defined.

<u>Definition VI.4.1</u>. A binary relation R over the set X is called a semiorder of X if the following are satisfied:

(i) Reflexivity: $(\forall x)\ (xRx)$

(ii) Completeness: $(\forall x,\ y{:}x \neq y)\ (xRy \vee yRx)$

(iii) Weak transitivity: $(\forall x,y,z)\ (rPxIyPz \vee rIxPyPz \Rightarrow rPz)$

In these statements a binary relation R over a set X is an ordering if it is complete, reflexive and transitive. R is the relationship "is preferred to or indifferent from," I is "indifference," and P means "is preferred to."

<u>Definition VI.4.2</u>. A binary relation R on a set X is a subsemiorder of X iff it is reflexive and complete and satisfies:

(iv) Mild transitivity: $(\forall x,y{:}xDy)\ (sPrRyPs \vee sRrPyPs \Rightarrow xPs;$
$sPxRrPy \vee sPsPrRy \Rightarrow sPy)$.

Here D means "does not differ in more than one dimension."

These concepts were developed to explain intransitivity of preference. In the capital budgeting example of Chapter I, the firm might prefer C over A, and D over A, but be indifferent between C and D, because along the first two dimensions (net present value and percentage in unstable lands) there might be a "noticeable difference." This kind of imperfect discrimination exists when we rank the elements of G(N).

A related phenomenon is finite human sensibility. Most people cannot tell the difference between Coca-Cola and Pepsi-Cola, no matter how hard they try. Another difficulty in preference comparisons is lack of data or less than perfect anticipatory ability. I may be indifferent between car A and car B because at this moment of choice I do not really know how each will suit me.

Because of these two sources of imperfect discrimination, and because of measurement error in the criterial readings (the z vectors), intransitivity of preference may exist in G(N). Therefore, it is probably wrong to conclude as Kornbluth [1978] does, that the weights discovered through the ranking algorithm can be used as inputs to an expanded problem based on multiobjective linear programming.

Although Ng [1977, p. 57] feels that the intrinsic underlying preference of a subsemiorder is an ordering, we have no need here for this much information. The ranking algorithm relies on the well-ordering condition.

Definition VI.4.3. The binary linear relation > on a set X satisfies the well-ordering condition if each non-empty subset of X has a maximal element. This condition lets us find the maximal element in G(N).

Choice with G(N) as the domain is probably a subsemiorder process. Forcing an ordering onto it is acceptable only because we do not seek to explain decision-making behavior or to construct a preference function. We do not have here a theory of choice; we have not captured the process as a mathematical psychologist strives to do.

Our plan was to bring the abilities of the computer to bear where appropriate, and to force the human mind to concentrate on the kinds of reasoning he does best. The binary choice by the DM and the linear programming computer analysis have achieved the stated goal of Chapter

VI, which was to choose a single response from the set of generators $G(N) \subseteqq N^{ex} \subset N$. This was accomplished by using a complete preference order to force introspection and to build confidence in the final solution. The ordering was accomplished by forming constraints from the differences between the z vectors, finding the convex hull of these constraints, and identifying the pairs of z vectors that are crucial to an ordering. The DM drives the algorithm by stating agreement or disagreement with the order of these pairs. Incommensurability, noncompensatoriness, new evaluative criteria, finite human sensibility and imperfect anticipatory ability were treated in this framework of sequential binary choices of nondominated solution vectors.

CHAPTER VII

SUMMARY, POSSIBLE EXTENSIONS AND CONCLUSIONS

London, Sept. 9 1772

Dear Sir,

In the affair of so much importance to you, wherein you ask my advice, I cannot, for want of sufficient premises, advise you what to determine, but if you please I will tell you how. When those difficult cases occur, they are difficult, chiefly because while we have them under consideration, all the reasons pro and con are not present to the mind at the same time; but sometimes one set present themselves, and at other times another, the first being out of sight. Hence the various purposes or inclinations that alternately prevail, and the uncertainty that perplexes us. To get over this, my way is to divide half a sheet of paper by a line into two columns; writing over the one Pro, and over the other Con. Then during three or four days consideration, I put down under the different heads short hints of the different motives, that at different times occur to me, for or against the measure. When I have thus got them all together in one view, I endeavor to estimate their respective weights; and where I find two, one on each side, that seem equal, I strike them both out. If I find a reason pro equal to some two reasons con, I strike out the three. If I judge some two reasons con, equal to some three reasons Pro, I strike out the five; and thus proceeding I find at length where the balance lies; and if, after a day or two of further consideration, nothing new that is of importance occurs on either side, I come to a determination accordingly. And, though the weight of reasons cannot be taken with the precision of algebraic quantities, yet when each is thus considered, separately and comparatively, and the whole lies before me, I think I can judge better, and am less liable to make a rash step, and in fact I have found great advantage from this kind of equation, in what may be called moral or prudential algebra.

Wishing sincerely that you may determine for the best, I am ever, my dear friend, yours most affectionately.

B. Franklin

[From B. Franklin, "A Letter to Joseph Priestly (1772), reprinted in The Benjamin Franklin Sampler, New York, Fawcett, 1956.]

VII.1 Summary

This dissertation has attempted to treat decision making in the multiple objective situation. Earlier literature has typically concentrated on a particular segment of the decision process, leaving several gaps in the body of techniques called multiple criteria decision making. The main contribution of the present work has been to create and justify a two-stage pruning process for treating the nondominated set which results from multiobjective mathematical programming (linear or nonlinear).

Drawing on both the behavioral and the mathematical sciences, it has been possible to address two important questions:

a) Should the decision maker use goal programming or multiobjective linear programming?

b) What can be done to aid the decision maker in choosing one final solution from a very large set of nondominated solutions?

The area stressed is the multiobjective linear programming model of the individual decision maker. To answer the first question, I show behavioral and mathematical support for the rejection of <u>a priori</u> ranking methods. The implication of this is that goal programming may be less applicable than multiobjective linear programming, since in using the latter technique it is not necessary to prioritize the objectives. For the second question a two-stage pruning algorithm is proposed; this aids the decision maker to process a large unwieldy nondominated set.

In Chapter I various conceptualizations of the firm introduced the multiple objective situation. These included the agency approach, the disaggregation approach, and the autopoietic approach. Additional brief remarks on non-profit organizations led naturally to the group decision environment. However, the body of this dissertation was solely concerned with the individual decision maker.

The second chapter searched the literature of mathematical programming. I reviewed the origins of the field in genera., and then discussed the optimization approach to the special problem of decision making with multiple objective functions. Some of the major methods included in the review were the surrogate worth trade-off method of Haimes, Hall, and Freedman [1975], the interactive programming of Dyer et al. [1972], the general parametric programming model, the multicriterion simplex methods of Zeleny [1974(a)] and Evans and Steuer [1973], and the goal programming of Lee [1973] and Ignizio [1976].

Since models that are not foreign to the mind's perceptual and cognitive habits are more likely to be successfully implemented, Chapter III reviewed the individual choice theories of mathematical psychology. Particularly interesting were Thurstone's [1927(a), 1927(b)] two laws of comparative and categorical judgment, Luce's [1959] study of individual choice behavior, Coombs's [1964] theory of data, and Tversky's [1972] elimination-by-aspects model.

The natural limits on man's ability to process information were also viewed as significant in the choice of a multicriterion model. The Brunswik Lens Model [Rappoport and Summers, 1973, p. 16 ff.] provided the framework for describing judgment processes. In the Lens model, criteria implied cues, which in turn caused judgments. The famous Miller [1957] article on how many objects can be accommodated in short-term memory led to other work on human limitations. Examples of this included fixation, quality, primacy, averaging, and partialization.

The result of Chapter III is a pessimistic assessment of our ability to set a priori weights on multiple goals. This does not mean, however, that MCDM is not practical. It indicates that goal programming (GP) may be far less applicable than multiobjective linear programming (MOLP), because MOLP does not require advance setting of weights on goals. Whereas GP pre-empts the decision by prematurely discarding most of the nondominated set, MOLP relinquishes control to the DM in the pruning stage. This is precisely the point where man's creative reasoning ability should take over from the computer's combinatorial sort of reasoning. Future research should be cognizant of the fact that the dominated set might contain the most desirable solution if the DM has reformulated his objectives and trade-offs more explicitly.

Chapter IV drives in the same direction as the previous chapter; this time the reasoning is mathematical, rather than behavioral. The way in which the priority weights in GP drive a linear equation is extremely hard to characterize. The effects of a change in the weight vector can be counterintuitive. This was illustrated by a discussion of naive weights. Naive weights are obtained when the DM associates the highest weights with the most important goals. The unsatisfactory nature of this weighting procedure was analyzed by woeking with the eigenvalues and eigenvectors of the criterion matrices of the goal programming proglem.

In Chapters V and VI we face the problem of the often overwhelming size of the nondominated set (N). Although integer and other nonlinear cases in multicriterion programming are still under development, they do share with the linear case the concept of nondomination. We attack the general problem of reducing the size of this nondominated set of solutions. This effort has been called pruning; here it has been treated as a two-stage process. In the first stage most of the work is done by the computer. In the second, much more of the task is completed by the decision maker.

The fifth chapter treats the first stage of pruning by using the data analytic technique called cluster analysis. The idea is to portray N by a representative subset. Cluster analysis partitions N into groups that are relatively homogeneous within their boundaries. Essentially a very general evaluative criterion has been added: minimum redundancy. Since there is a threshold of resolution which hinders the DM in perceiving the difference between two very similar solution vectors, there is little point in making him wast time in processing all of N as he searches for a final solution.

Throughout the dissertation, weights (on objectives or attributes) are treated as "knobs." Chapters III and IV showed some problems with a priori weighting; if we go to a television set it is nearly impossible to set the knobs properly on the first twirl. First the set must be switched on, a channel must be picked, and then we can tune with the controls. Think of solving the MOLP problem

$$\begin{aligned} &\text{Maximize} && Cx = \{z\} \\ &\text{Subject to} && Ax \leqq b \end{aligned}$$

as turning on the set. Looking at the static and blur of the TV's image as it warms up is like the DM facing the set N^{ex} (or N). In Chapter V clustering is used to "pick a channel." Creating the generator set $\{g\} \varepsilon G(N)$ is analogous to the existence of channels; once one is chosen we use the tuning knobs to improve the picture.

Technically, Chapter V experimented with the nondominated set of two problems in the literature (Zeleny [1974(a)] and Steuer [1975]). Two forms of cluster analysis were tested--direct clustering versus hierarchical clustering. Within the group of hierarchical methods eight were tried. In the present application the two worst things that could happen were clusters that "chained" and outlying vectors (the residue set) that were obscured. Under these two criteria, on the particular data used, three algorithms worked best--Ward's Method, the Group Average Method, and the Centroid Method. The hierarchical methods are recommended over direct clustering. However, the strength of this statement is weakened because some similarity between direct and hierarchical clustering was discovered. In the two-stage pruning process, this clustering serves as a first stage to minimize redundancy, and thereby reduces the chance that the selection of a final solution will stress the DM beyond his information endurance.

Chapter VI retrieves the concept of weights once the generators $\{g\} \varepsilon G(N)$ are available. In this second stage of the pruning process, weights are used for fine tuning, just as we turn the knobs once a channel has been chosen. By trial and error we use the feedback from our eyes to move toward the best picture (knowing that there might be a better setting that is not easily discoverable). The ranking algorithm of this chapter asks the DM to make binary comparisons between selected $g \in G(N)$. He does not deal directly with criterion weights, just as the person with a television set does not need calibrated knobs. The TV viewer reacts to "better" or "worse" pictures, and moves the knobs as a consequence. The DM in the ranking algorithm responds to z vectors as better or worse; the weights merely fall out as a side effect of the linear programming approach.

VII.2 Possible Extensions of the Research

This dissertation has dealt with mathematical, statistical, psychological and financial ideas; however, additional applications come to mind. Of paramount interest would be the application of the MOLP model with two-stage pruning to a real decision situation. Another worthwhile effort would be to compare several competing methods of multiple objective decision analysis in a laboratory or field setting. Several more theoretical possibilities will be described below.

VII.2.1 The ranking algorithm versus Saaty's method. The ranking algorithm of Chapter VI deserves to be carefully compared with Saaty's [1975, 1977] method of priority scaling. Section V.3.5 describes the strengths and weaknesses of this eigenanalysis method. The attraction of the approach is that strength of preference is obtained in a quantitative mode, and an index of the DM's inconsistency is available.

It would be interesting to try to reduce the number of judgments that the DM must make. The DM should also be encouraged to deal more explicitly with incommensurability, compensatoriness, interdependent preferences and local behavior of individual preferences.

VII.3.2 Ideal points and second choices. Chapter III mentions Coombs's [1964] comment that little work has been done on second choices. Zeleny [1976(c)] also mentions this. The idea is that our models are fairly successful in the situation of a consumer who must decide whether to buy a Porsche, a Lincoln Continental, a Jeep or a Honda Civic. But what if this person has just picked one of the cars and we say, "You can keep the first car you chose, and we will let you choose a second one?"

Preferences are different for the second choice. The research questions here are "portfolio" considerations--the attributes of the second car should balance and complement those of the first car. Perhaps the lucky DM would like to have a Lincoln and a Jeep.

I suggest two approaches to this problem. Zeleny's [1976(c)] nondominated pocket could be one way of freeing the DM from the somewhat narrow concept of the nondominated set. Or vectors could be constructed for contingent second choices. Various combinations of first and second choices form a set of vectors: those could be pruned by the two-stage algorithm of this dissertation.

VII.2.3 The agency theory of the firm. Jensen and Meckling [1976] describe the agency theory of the firm (see also Chapter I). Since the firm as they conceive it consists of multiple agents with multiple criteria, the techniques of multicriterion mathematical programming fit very well both as prescriptive and descriptive models. Research of this nature is already in progress at the Wharton School (see, for example, Banker [1977] for a very large MOLP model of a Philadelphia bank).

VII.2.4 Financial planning models. As I mentioned in Chapter I, optimization modeling in finance is strongly rooted in the single-criterion (profit) maximizing tradition. Since the linkage between activities x_j and the objective function is often time-staged and non-linear, some efforts to disaggregate the firm's objectives have appeared in the literature (see, for example, Bernardo and Lanser [1977] and Morse [1978]). If the operating budget and the capital budget were represented as two right-hand sides in an MOLP framework,[1] then we might be able to expand on the traditional concepts of capital and money markets. To the normal external markets I propose adding an internal market, where the various functional areas (such as production and marketing) compete for funds. In this way the nascent duality theory of multiobjective linear programming (see Duesing [1976] and Isermann [1977]) might improve our knowledge of one of the most perplexing dilemmas of finance, the cost of capital.

VII.2.5 Strategic capital budgeting. Strategic capital budgeting is a term used to signify the addition of high-level business policy and strategy to the purely financial capital budgeting criteria such as net present value and internal rate of return. An example of this appears in Chapter I. Until recently this has not been possible, since capital investment projects on not continuous variables. Last year, though, a Brazilian researcher appears to have perfected an

---

[1]Kornbluth [1974] shows how the dual of a MOLP problem has multiple right-hand sides.

integer variant of multiobjective linear programming (see Bitran [1977]).

VII.2.6 Zero-base budgeting. Zero-base budgeting is currently very popular both in the corporate world (see Newsweek Magazine Publishing Company [1977]) and in the Carter administration. The output of zero-base budgeting is a set of vectors called decision packages. This set, like the nondominated set in multicriterion mathematical programming, is usually too large to be easily pruned by the DM. I plan to try to use the two-stage pruning algorithm on decision packages.

VII.2.7 Decomposition analysis of financial ratios. Traditional financial ratio analysis is a univariate analytical technique. Raw ratios fail to recognize trends. They are also not as full of information as they are purported to be; for example, does a current ratio of 3 mean that the firm is well-managed, or that its investment opportunity set is nearly empty?

One improvement in the use of financial ratios is to treat balance sheet items as percentages, and to look at changes over time by using natural logarithm transformations and entropy measures. Decomposition analysis creates multivariate financial ratios; a set of firms becomes a set of vectors. If this number of vectors is very large, the two-stage pruning algorithm may be applicable as a decision support. For example, securities analysts might want to reduce the firms in an industry to a representative subset. The ranking algorithm of Chapter VI could then be applied to formulate portfolio recommendations. Preliminary work on 155 companies in the electric utility industry is available (see Dise and Morse [1978]).

VII.3 Conclusions

I conclude that individual and group decisions can be described and improved by the set of tools that Zeleny [1977] has called multiple criteria decision making (MCDM). When decision alternatives must be created and compared (this is what Peter Fishburn [1977] refers to as multicriterion choice theory) two methodologies have achieved prominence in the management science literature, namely goal programming (GP) and multiobjective linear programming (MOLP). The first contribution of this dissertation is to demonstrate the more general applicability of MOLP. Finite human sensibility and information processing constraints indicate that the main drawback to MOLP methods is the arduous job of choosing a final solution from the nondominated set.

The second main result of this dissertation is a two-stage pruning algorithm that automates a large percentage of that choice process. The simplicity of the two-stage pruning algorithm offers the hope that decision makers can implement MCDM techniques. The generality of the two-stage paradigm seems to unify many situations that are essentially multidimensional choice processes.

BIBLIOGRAPHY

Agarwal, S. K. "Optimizing Techniques for the Interactive Design of Transportation Networks Under Multiple Objectives." Ph.D. dissertation, Northwestern University, 1973.

Akyilmaz, M. O. Ph.D. dissertation, University of Pennsylvania, 1976.

Anderberg, Michael R. Cluster Analysis for Applications. New York: Academic Press, 1973.

Arbib, Michael. The Metaphorical Brain. New York: Wiley-Interscience, 1972.

Arrow, K. J. Social Choice and Individual Values. New York: John Wiley and Sons, 1951.

Awerbuch, Shimon, and Wallace, William K. "Alternative Multi-Criteria Programming Models for Community Development Planning." Paper presented at the TIMS-ORSA National Meetings, Chicago, May 1974.

Awerbuch, Shimon, and Wallace, William K. "A Goal-Setting and Evaluation Model for Community Development." Working paper, School of Management, Rensselaer Polytechnic Institute, Troy, N.Y., March 1976.

Baker, Frank B. "Stability of Two Hierarchical Grouping Techniques Case 1: Sensitivity to Data Errors." JASA 69 (June 1974): 440-445.

Baker, Frank B., and Hubert, Lawrence J. "Measuring the Power of Hierarchical Cluster Analysis." JASA 70 (March 1975): 31-37.

Balintfy, Joseph L.; Duffy, William J.; and Sinha, Prabhakant. "Modeling Food Preferences Over Time." Operations Research 22 (July-August 1974): 711-727.

Banker, Robert L., and Gupta, Shiv K. "Multiple Objective Resolution for Stakeholders in a Complex Organization." Working paper, Wharton School, 1977.

Baron, Robert J. "A Model for Cortical Memory." Journal of Mathematical Psychology 20 (1970): 37-59.

Bartoszynski, Robert. "A Metric Structure Derived from Subjective Judgments: Scaling under Perfect and Imperfect Discrimination." Econometrica 42 (January 1974).

Bass, Frank M.; Lehmann, Donald R.; and Pessemier, Edgar A. "An Experimental Study of Relationships between Attitudes, Brand Preference, and Choice." Behavioral Science 17 (November 1972): 532-541.

Bass, Frank M.; Pessemier, Edgar A.; Teach, Richard; and Talarzyk, W. Wayne. "Preference Measurement in Consumer Market Research." 1969 Business and Economic Statistics Section Proceedings of the American Statistical Association (1969), pp. 88-95.

Becker, Ernest. The Denial of Death. New York: The Free Press, 1973.

Beedles, William L. "A Micro-econometric Investigation of Multi-objective Firms." Journal of Finance 32 (September 1977): 1217-1234.

Belenson, Sheldon M., and Kapur, Kailish C. "An Algorithm for Solving Multicriterion Linear Programming Problems with Examples." Mathematical Programming 24 (1973): 65-77.

Benayoun, R.; Montgolfier, J. de; Tergny, J.; and Laritchev, O. "Linear Programming with Multiple Objective Functions: STEP Method (STEM)." Mathematical Programming 1 (March 1971): 366-375.

Bennett Robert A. "Big Apple's Banking Lure." New York Times, 26 September 1976, Section 3, p. 1.

Benson, H. P., and Morin, T. L. "The Vector Maximization Problem: Proper Efficiency and Stability." SIAM Journal on Applied Mathematics 32 (January 1977): 64-72.

Bergstresser, Kenneth A., Jr. Ph.D. dissertation, University of Rochester, 1976.

Bergstresser, K.; Charnes, A.; and Yu, P. L. "Generalization of Domination Structures and Nondominated Solutions in Multi-criteria Decision Making." Research report CS-185, Center for Cybernetic Studies, University of Texas at Austin, 1974. (Forthcoming in the Journal of Optimization Theory and Applications.)

Bernardo, John J., and Lanser, Howard P. "A Capital Budgeting Decision Model with Subjective Side Criteria." Journal of Financial and Quantitative Analysis 12 (1977): 261-275.

Bijnen, E. J. _Cluster Analysis: Survey and Evaluation of Techniques_. Holland: Tilburg University Press, 1973.

Biomedical Computer Programs, BMD P Series. Los Angeles: University of California Press, 1975.

Bitran, Gabriel R. "Linear Multiple Objective Programs with Zero-One Variables." _Mathematical Programming_ 13 (1977): 121-129.

Black, Duncan. "On the Rationale of Group Decision Making." _Journal of Political Economy_ 56 (1948)(a): 23-25.

Black, Duncan. "The Decisions of a Committee Using a Special Majority." _Econometrica_ 16 (1948)(b): 245-261.

Bod, P. "Linear Optimization with Several Simultaneously Given Objective Functions" (in Hungarian). _Publications of the Mathematical Institute of the Hungarian Academy of Sciences_ 8 (1963): B, Fase 4.

Boise Cascade Corporation. "The Story of Our Future--We Trust," public relations handout, 1974.

Borges, Jorge Luis. _Labyrinths; Selected Stories and Other Writings_. Edited by Donald A. Yates and James E. Irby. New York: New Directions, 1962.

Boring Edwin G. _A History of Experimental Psychology_. New York: Appleton, Century and Cross, 1950.

Brown, Murray, and Heien, Dale. "The S-Branch Utility Tree: A Generalization of the Linear Expenditure System." _Econometrica_ 40 (July 1972): 737-747.

Brunswik E. _The Conceptual Framework of Psychology_. Chicago: University of Chicago Press, 1952.

________. "Representative Design and Probabilistic Theory in a Functional Psychology." _Psychological Review_ 62 (1955): 193-217.

Calantone, Roger; Schewe, Charles D.; and Wiek, James L. "Marketing Information System Usage: An Application of Benefit Bundle Segmentation." Working paper #74-78, Center for Business and Economic Research, School of Business Administration, University of Massachusetts, November 1974.

Caplin, D. A. _International Journal of Management Science_ 3 (August 1975): 423-441.

Carmichael, James; George, J.; and Julius, R. "Finding Natural Clusters." Systematic Zoology 17 (June 1968): 144-150.

Chang, S. S. L. "General Theory of Optimal Processes." SIAM Journal on Control Theory 4 (1966): 46-55.

Charnes, A., and Cooper, W. W. "Goal Programming and Constrained Regression--A Comment." OMEGA 3 (1975): 403-409.

Charnes A., and Cooper, W. W. Management Models and Industrial Applications of Linear Programming. New York: John Wiley and Sons, 1961.

Charnes, A.; Cooper, W. W.; and Ferguson, R. "Optimal Estimation of Executive Compensation by Linear Programming." Management Science 1 (January 1955): 138-151.

Clark, Lindley H., Jr. "Zero-base Budgeting, Advocated by Carter, Used by Many Firms." Wall Street Journal, 14 March 1977, p. 1.

Contini, Bruno. "A Stochastic Approach to Goal Programming." Operations Research 16: 576-583.

Coombs, Clyde H. "The Measurement and Analysis of Family Composition Preferences." Notes from a lecture at the University of Massachusetts, March 1975.

________. A Theory of Data. New York: John Wiley and Sons, 1964.

Cyert, Richard M., and DeBroot, Morris H. "Adaptive Utility." In Adaptive Economic Models. Edited by Richard H. Day and Theodore Groves. New York: Academic Press, 1975.

Da Cunha, N. O., and Polak, E. "Constrained Minimization Under Vector-Valued Criteria in Finite Dimensional Spaces." Journal of Mathematical Analysis and Applications 19 (1967): 103-124.

Day, Richard H., and Groves, Theodore, eds. Adaptive Economic Models. New York: Academic Press, 1975.

Dantzig, G. B. Linear Programming and Extensions. Princeton: Princeton University Press, 1963.

Dearing, P. M.; Francis, R. L.; and Lowe, T. J. "Convex Location Problems on Tree Networks." Operations Research 24 (1976).

Debreu, G. Theory of Value, An Axiomatic Analysis of Economic Equilibrium. New Haven: Yale University Press, 1959.

Dee, Norbert, et al. "An Environmental Evaluation System for Water Resource Planning." Water Resources Research 9 (1973): 523-535.

Delft, A. van, and Nijkamp, P. "A Multi-Objective Decision Model for Regional Development, Environmental Quality Control and Industrial Land Use." Research Memorandum No. 31, Department of Economics, Vrije Universiteit, Amsterdam, 1974.

Dise, Galen, and Morse, J. N. "Decomposition Analysis: A Survey and Application to the Electric Utility Industry." Working paper, University of Delaware, 1978.

Driver, Michael J.; Mock, Theodore J.; and Vasarhelyi, Miklos A. "Some Experimental Results in MIS, Human Information Processing and Tailored Information Systems." Working paper, Graduate School of Business Administration, University of Southern California, Los Angeles, 1975.

Duckstein, Lucien, and David, Laszlo. "Multi-Criterion Ranking of Long-Range Water Resource Development Plans." Working paper, Department of Systems and Industrial Engineering and Hydrology and Water Resources, University of Arizona, Tucson, 1975.

Duckstein, Lucien; Monarchi, David E.; and Weber, Jean R. "An Interactive Multiple Objective Decision Making Aid Using Nonlinear Goal Programming." Working paper, University of Colorado, 1975.

Duesing, Erick C. "A Dual Algorithm for Linear Multiple Objective Programming." Working paper, Department of Economics, Western Michigan University, Kalamazoo, 1976.

Duran, Benjamin S., and Odell, Patrick L. Cluster Analysis: A Survey. Lecture Notes in Economics and Mathematical Systems, No. 100. New York: Springer-Verlag, 1974.

Dyer, James S. "Interactive Goal Programming." Management Science 19 (September 1972): 62-70.

Dyer, James S.; Farrell, William; and Bradley, Paul. "Utility Functions for Test Performance." Management Science 20 (December 1973): Part I, pp. 507-519.

Dyer, James S., and Sarin, Rekeih K. "A Guide to the Choice and Estimation of Multi-Attribute Utility Models." Working paper, Graduate School of Management, University of California, Los Angeles, 1975.

Easton, Allan. Complex Managerial Decisions Involving Multiple Objectives. New York: John Wiley and Sons, 1973.

Ecker, J. G., and Kouada, I. A. "Finding Efficient Points for Linear Multiple Objective Programs." Mathematical Programming 8 (June 1975): 375-377.

Eilon, Samuel. "Note on Many-Sided Shadow Prices." OMEGA 2 (December 1974): 821-824.

Eto, Hajime. "An Evaluation of the Duality Gap and its Application to a Technology Assessment of Commuting Transportation Systems." Working paper, Systems Development Laboratory, Hitachi, Ltd., Yokohama, Japan, 1975.

Evans, J. P., and Steuer, R. E. "A Revised Simplex Method for Linear Multiple Objective Programs." Mathematical Programming 5 (1973): 54-72.

Farquhar, Peter H. "A Fractional Hypercube Decomposition Theorem for Multiattribute Utility Functions." Rand Paper Series, P-5322. Rand Corporation, Santa Monica, November 1974 (a).

________. "Interdependent Criteria in Decision Making." Working paper, Department of Industrial Engineering and Management Science, The Technological Institute, Northwestern University, Evanston, Ill., 1975.

________. "Pyramid and Semicube Decompositions of Multi-attribute Utility Functions." Rand Paper Series, P-5323. Rand Corporation, Santa Monica, November 1974 (b).

________. "A Survey of Multiattribute Utility Theory and Applications." Management Science (1977).

Farquhar, Peter H., and Rao, Vithala R. "A Balance Model for Evaluating Subsets of Multiattributed Items." Management Science 22 (January 1976): 528-539.

Ferrel, William R., and Sheridan, Thomas P. Man-Machine Systems: Information, Control, and Decision Models of Human Performance. Cambridge, Mass.: MIT Press, 1974.

Findlay, M. Chapman, and Whitmore, G. A. "Beyond Shareholder Wealth Maximization." Financial Management 3 (Winter 1974): 25-35.

Fishburn, Peter C. "Lexicographic Orders, Utility, and Decison Rules: A Survey." Management Science 20 (July 1974)(b): 1442-1471.

Fishburn, Peter C. "Single-Peaked Preferences and Probabilities of Cyclical Majorities." Behavioral Science 19 (January 1974) (a): 21-27.

________. "A Study of Lexicographic Expected Utility." Management Science 17 (July 1971): 672-678.

________. "A Survey of Multiattribute/Multicriterion Evaluation Theories." Paper presented at the Conference on Multiple Criteria Problem Solving, SUNY at Buffalo, August 1977.

________. "Utility Theory." Management Science 14 (January 1968): 335-378.

________. Utility Theory for Decision Making. New York: John Wiley and Sons, 1970.

________. "Voter Concordance, Simple Majorities and Group Decision Methods." Behavioral Science 18 (September 1973): 364-376.

Fishburn, Peter C., and Keeney, Ralph L. "Generalized Utility Independence and Some Implications." Operations Research 23 (September-October 1975): 928-940.

Gal, Tomas, and Nedoma, Josef. "Multiparametric Linear Programming." Management Science 18 (March 1972): 406-422.

Gehrlein, William V., and Fishburn, Peter C. "Information Overload in Mechanical Processes." Management Science 23 (December 1976): 391-398.

Gembicki, F. "Vector Optimization for Control with Performance and Parameters Sensitivity Indices." Ph.D. dissertation, Case Western Reserve University, 1973.

Geoffrion, Arthur M. "Proper Efficiency and the Theory of Vector Maximization." Journal of Mathematical Analysis and Applications 22 (1968): 618-630.

Geoffrion, A. M.; Dyer, J. S.; and Feinberg, A. "An Interactive Approach for Multi-criterion Optimization, with an Application to the Operation of an Academic Department." Management Science 15 (December 1972): Part I, pp. 357-369.

Gomes, L. F. A. M. Ph.D. dissertation, University of California, Berkeley, 1976.

Green, Paul E., and Rao, Vithala R. "A Note on Proximity Measures and Cluster Analysis." Journal of Marketing Research 6 (August 1969): 359-364.

Green, Paul E., and Tull, Donald S. Research for Marketing Decisions. Englewood Cliffs, N.J.: Prentice-Hall, 1975.

Grossman, S. J., and Stiglitz, J. E. "On Value Maximization and Alternative Objectives of the Firm." Journal of Finance 32 (May 1977): 389-402.

Groves, David L., and Kahalas, Harvey. "An Empirical Analysis of Personal Value Information." 1975 Southeastern AIDS Proceedings. Edited by T. F. Anthony and A. B. Carroll. pp. 112-113.

Haimes, Y. Y. "The Integration of System Identification and System Optimization." Ph.D. dissertation, University of California, Los Angeles, 1970.

Haimes, Yacov Y., and Hall, Warren A. "Multiobjectives in Water Resource Systems Analysis: The Surrogate Worth Trade Off Method." Water Resources Research 10 (August 1974): 615-624.

Haimes, Yacov Y.; Hall, Warren A.; and Freedman, Herbert T. Multiobjective Optimization in Water Resources Systems. New York: American Elsevier Publishing Co., 1975.

Hanson, Pierre. "Bicriterion [sic] Cluster Analysis as an Exploratory Tool." Proceedings of the Conference on Multiple Criteria Problem Solving. Buffalo, August 1977.

Harmon, Harry H. Modern Factor Analysis. Chicago: University of Chicago Press, 1960, revised 1964.

Harrald, Jack; Loetta, Joseph; Loo, Ernest; Wallace, William A.; and Wendell Richard. "The Use of Goal Programming in Evaluating Managerial Standards for a Marine Environmental Protection Program." Working paper, Marine Environmental Protection Program, U.S. Coast Guard Headquarters, Washington, D.C., 1975.

Hartigan, J. A. Clustering Algorithms. New York: John Wiley and Sons, 1975.

________. "Direct Clustering of a Data Matrix." JASA 67 (March 1972): 123-129.

Hendrix, G. G., and Stedry, A. C. "The Elementary Redundancy--Optimization Problem: A Case Study in Probabilistic Multiple-Goal Programming." Operations Research 22 (May-June 1974): 639-653.

Ignizio, James P. Goal Programming and Extensions. Lexington, Mass.: Lexington Books, D. C. Heath, 1976.

Ijiri, Yuji. Management Goals and Accounting in Control. Chicago: Rand McNally, 1965.

Isermann, Heinz. "Duality in Multiple Objective Linear Programming - A Review of Four Duality Concepts." Paper presented at the Conference on Multiple Criteria Problem Solving, Buffalo, August 1977.

Jensen, Michael C., and Meckline, William H. "Theory of the Firm: Managerial Behavior, Agency Costs, and Ownership Structure." Journal of Financial Economics 3 (October 1976): 305-360.

Johnson, Erik. Studies in Multiobjective Decision Models. Monograph No. 1. Lund, Denmark: Economic Research Center, 1968.

Johnson, Stephen C. "Hierarchical Clustering Schemes." Psychometrika 36 (September 1967): 241-254.

Jones, Wesley H., and Pessemier, Edgar A. "Joint-Space Analysis of the Structure of Affect Using Single-Subject Discriminant Configurations, Part I." Paper No. 435, Krannert Graduate School of Industrial Administration, Purdue University, West Lafayette, Ind., February 1974.

Karlin, S. Mathematical Methods and Theory in Games, Programming and Economics, Vol. 1. Boston: Addison-Wesley, 1959.

Keefer, D. L. Ph.D. dissertation, University of Michigan, 1976.

Keeney, Ralph L. "A Decision Analysis with Multiple Objectives: The Mexico City Airport." The Bell Journal of Economics and Management Sciences 4 (Spring 1973): 101-117.

Keeney, Ralph L., and Raiffa, Howard. Decisions with Multiple Objectives: Preferences and Value Tradeoffs. New York: John Wiley and Sons, 1976.

Khumawala, Basheer M. "An Efficient Algorithm for the p-Median Problem with Maximum Distance Constraints." Geographical Analysis 5 (October 1973): 309-321.

Kierulff, Herbert E. "Return on Investment and the Fatal Flaw." California Management Review 19 (Winter 1976): 61-70.

Knight, Frank H. The Ethics of Competition. New York: Harper, 1935.

Koopmans, T. C. "Analysis of Production as an Efficient Combination of Activities." In Activity Analysis of Production, Cowles Commission Monograph 12. Edited by T. C. Koopmans. New York: John Wiley and Sons, 1951.

Kornbluth, J. S. H. "Duality, Indifference and Sensitivity Analysis in Multiple Objective Linear Programming." Operational Research Quarterly 25 (1974): 599-614.

________. "Ranking with Multiple Objectives." Proceedings of the Conference on Multiple Criteria Decision Making. Edited by Stanley Zionts. Buffalo: Springer-Verlag, 1978.

Kornbluth, J. S. H., and Caplin, D. A. "Multiobjective Investment Planning Under Uncertainty." OMEGA 3 (August 1975): 423-441.

Kraft, Arthur, and Kraft, John. "V-Branch: A Generalized Utility Function." Memorandum, University of Nebraska, Lincoln, 1975.

Krajewski, Leroy J., and Ritzman, Larry P. "Multiple Objectives in Linear Programming - An Example in Scheduling Postal Resources." Decision Sciences 4 (July 1973): 364-378.

Kruskal, J. B. "On the Shortest Spanning Subtree of a Graph, and a Traveling Salesman Problem." Proceedings of the American Mathematical Society 7 (1956): 48-50.

Kuhn, H. W., and Tucker, A. W. "Nonlinear Programming." Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability. Berkeley: University of California Press, 1951. pp. 481-492.

Lance, G. N., and Williams, W. T. "A General Theory of Classificatory Sorting Strategies: Hierarchical Systems." Computer Journal 9 (1967): 373-380.

Lawrence, Kenneth D., and Koch, Howard B. "A Multiple Objective Linear Programming Model for the Transshipment of Goods Problem." Working paper, Management Information Services Department, Hoffman La-Roche, Nutley, N.J., 1975.

Lee, Sang M. Goal Programming for Decision Analysis. Philadelphia: Auerbach, 1972.

Lindsay, B. E. Ph.D. dissertation, University of Massachusetts, 1976.

Ling, Robert F. "A Probability Theory of Cluster Analysis." JASA 68 (March 1973): 159-164.

Loasby, Brian J. Choice, Complexity and Ignorance. Cambridge: Cambridge University Press, 1976.

Luce, R. Duncan. Individual Choice Behavior: A Theoretical Analysis. New York: John Wiley and Sons, 1959.

Luce, R. Duncan, and Raiffa, Howard. Games and Decisions. New York: John Wiley and Sons, 1957.

MacCrimmon, Kenneth R., and Siu, John K. "Making Trade-Offs." Decision Sciences 5 (October 1974): 680-704.

Major, D. C. "Benefit-Cost Ratios for Projects in Multiple Objective Investment Programs." Water Resources Research 5 (1969): 1174-1178.

Mason, Richard O., and Moskowitz, Herbert. "Construction in Information Processing: Implications for Management Information Systems." Decision Sciences 3 (October 1972): 35-54.

McCann, John M., and Wilkie, William L. "The Halo Effect and Related Issues in Multi-Attribute Attitude Models - An Experiment." Working paper, Krannert Graduate School of Industrial Administration, Purdue University, Lafayette, Ind., October 1972.

McGrew, D., and Haimes, Y. Y. "A Parametric Approach to the Integrated System Identification and Optimization Problems." Journal of Optimization Theory and Applications 13 (1974); 582-605.

Miller, George A. "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information." Psychological Review (March 1956): 81-97.

________. "Information and Memory." Scientific American (August 1965): 2-6.

________. The Psychology of Communication: Seven Essays. New York: Basic Books, 1967.

Monarchi, David E.; Weber, Jean E.; and Duckstein, Lucien. "An Interactive Multiple Objective Decision Making Aid Using Nonlinear Goal Programming." Working paper, University of Colorado, 1975.

Morse, J. N. "Strategic Capital Budgeting: A Response to the Bernardo-Lanser Model." Working paper, University of Delaware, 1978.

Morse, Joel N., and Clark, Rolf. "Goal Programming in Transportation Planning: The Problem of Setting Weights." Northeast Regional Science Review 5 (1975): 140-147.

Moskowitz, Herbert. "Some Observations on Theories of Collective Decisions." Paper No. 428, Krannert Graduate School of Industrial Administration, Purdue University.

Newsweek Magazine Publishing Company. "Zero-based Media Planning." Corporate brochure, 1977.

Nerlove, Sara B.; Romney, A. Kimball; and Shepard, Roger No. Multidimensional Scaling: Theory and Applications in the Behavioral Sciences, Vol. 1 (Theory). New York: Seminar Press, 1972.

Ng, Yew-Kwang. "Sub-semiorder: A Model of Multidimensional Choice with Preference Transitivity." Journal of Mathematical Psychology 16 (1977): 51-59.

Pareto, V. Cours d'Economie Politique. Lausanne: Rouge, 1896.

Pessemier, Edgar A. "A Measurement and Composition Model for Individual Choice Among Social Alternatives." Working paper No. 348, Krannert Graduate School of Industrial Administration, Purdue University, Lafayette, Ind., April 1972.

Pessemier, Edgar A., and Wilkie, William L. "Issues in Marketing's Use of Multi-Attribute Models." Working paper No. 365, Krannert Graduate School of Industrial Administration, Purdue University, Lafayette, Ind., August 1972.

Pessemier, Edgar A., and Wilkie, William L. "Multi-attribute Choice Theory - A Review and Analysis." Working paper No. 372, Krannert Graduate School of Industrial Administration, Purdue University, Lafayette, Ind., September 1972.

Philip, Johan. "Algorithms for the Vector Maximization Problem." Mathematical Programming 2 (1972): 207-229.

Quinn, J. B. "Strategic Goals: Process and Politics." Sloan Management Review 19 (Fall 1977): 21-37.

Raiffa, Howard. Decision Analysis. Boston: Addison-Wesley, 1970.

________. "Preferences for Multiattributed Alternatives." Rand Corporation Memorandum RM-5868-DOT/RC, April 1969.

Rao, M. R. "Cluster Analysis and Mathematical Programming." JASA 66 (September 1971): 622-626.

Rapoport, Anatol, and Horvath, William J. Information Processing in Neurones and Small Nets. Mental Health Research Institute, University of Michigan, 1960.

Rappoport, Leon, and Summers, David A., eds. Human Judgment and Social Interaction. New York: Holt, Rinehart, and Winston, 1973.

Reardon, Kevin J. "A Multi-Stage Model for Capital Budgeting with Uncertain Future Investment Opportunities." Technical Report SOL 74-9, Systems Optimization Laboratory, Stanford University, August 1974.

Reid, R. W., and Vemuri, V. "On the Noninferior Index Approach to Large-scale Multi-criteria Systems." Journal of the Franklin Institute 291 (1971): 241-254.

Richard, Scott F. "Multivariate Risk Aversion, Utility Independence and Separable Utility Functions." Management Science 22 (September 1975): 12-21.

Ritzman, Larry P. "Decision Analysis with Multiple Objectives." Proceedings of the American Institute of Decision Sciences, pp. 339-343. Atlanta, Georgia, November 1974.

Rivett, P. "The Model and the Objective Function." Operational Research Quarterly 21 (1970).

Roy, Bernard. "From Optimization on a Fixed Set to Multi-Criteria Decision Aid." Paper presented at the meetings of the Institute of Management Sciences, Kyoto, Japan, July 1975.

________. "How Outranking Relation Helps Multiple Criteria Decision Making." Theorie de la Decision, Actes du Seminaire de Beaulieu-Sainte-Assise, December 1973.

________. "Outranking and Fuzzy Outranking: A Concept Ranking Operational Partial Order Analysis." In Decision Making With Multiple Conflicting Objectives. Edited by Raiffa and Kenney. Vienna: IIASA, 1976.

________. "Problems and Methods with Multiple Objective Functions." Mathematical Programming 1 (November 1971): 239-266.

Ryan, L. J. Ph.D. dissertation, North Texas State University, 1976.

Saaty, Thomas L. "Hierarchies and Priorities - Eigenvalue Analysis." Rough draft of part of a book, Wharton School, 1975.

Saaty, Thomas L. "A Scaling Method for Priorities in Hierarchical Structures." Journal of Mathematical Psychology 15 (1977): 234-281.

Sahin, Kenan E. "Predicting Business Failure with Mathematical Programming." Working paper, School of Business Administration, University of Massachusetts, Amherst, 1976.

Scheerer, Martin. "Problem-Solving." Scientific American (April 1963): 2-9.

Scott, Jerome E., and Wright, Peter. "Modeling an Organizational Buyer's Product Evaluation Strategy: Validity and Procedural Considerations." Journal of Marketing Research 13 (August 1976): 211-224.

Sfeir-Younis, A. Ph.D. dissertation, University of Wisconsin, 1976.

Shepard, R. N. "On Subjectively Optimum Selection Among Multiattribute Alternatives." In Human Judgments and Optimality. Edited by M. W. Shelly and G. L. Bryan. New York: John Wiley and Sons, 1964.

Shepard, Robert N.; Romney, A. Kimball; and Nerlove, Sara Beth. Multidimensional Scaling: Theory and Application in the Behavioral Sciences, Vol. 1 (Theory). New York: Seminar Press, 1972.

Sheridan, Thomas B., and Ferrell, William R. Man-Machine Systems: Information, Control, and Decision Models of Human Performance. Cambridge, Mass.: The MIT Press, 1974.

Shim, Jae K., and Siegel, Joel. "Quadratic Preferences and Goal Programming." Decision Sciences 6 (October 1975): 662-669.

Shipley, Frederic B. II. "Convergence of Adaptive Decisions." In Adaptive Economics. Edited by Richard H. Day and Theodore Groves. New York: Academic Press, 1975.

Silverman, J., and Hatfield, G. B. "A Theoretical Approach to Multi-Objective Problems." Paper presented at the TIMS 20th International Meeting, Tel Aviv, Israel, June 1973.

Simmonard, Michel. Linear Programming. Englewood Cliffs, N.J.: Prentice-Hall, 1966.

Simon, Herbert A. The Sciences of the Artificial. Cambridge, Mass.: MIT Press, 1969.

Sneath, Peter H. A., and Sokal, Robert R. Numerical Taxonomy; The Principles and Practice of Numerical Classification. San Francisco: W. H. Freeman, 1973.

Sokal, Robert R., and Sneath, Peter H. A. Principles of Numerical Taxonomy. San Francisco: W. H. Freeman, 1963.

Starr, Martin K. "Message from the President." OR/MS Today 2 (May 1975): 4 ff.

Starr, Martin K., and Greenwood, Len H. "Normative Generation of Strategies with Multiple Criteria Evaluation." Working paper, Columbia University, January 1976.

Steuer, Ralph E. ADBASE, Operating Manual, An Adjacent Efficient Basis Algorithm for Vector-Maximum and Interval Weighted-Sums Linear Programming Problems. College of Business and Economics, University of Kentucky, August 1974 (b).

________. "A Five Phase Procedure for Implementing a Vector-Maximum Algorithm for Multiple Objective Linear Programming Problems." Working paper, University of Kentucky, 1976 (a).

________. "An Interactive Linear Multiple Objective Programming Procedure Employing an Algorithm for the Vector-Maximum Problem." Working paper, University of Kentucky, 1976 (b).

________. "Interval Criterion Weights Programming." Working paper, University of Kentucky, 1975.

________. "Interval Criterion Weights Programming: A Portfolio Selection Example, Gradient Cone Modification, and Computational Experience." Proceedings, Tenth Annual Southeastern Regional TIMS Meeting. Miami, 1974 (a).

________. "Linear Multiple Objective Programming: Theory and Computational Experience." Ph.D. dissertation, University of North Carolina, 1973.

Tatsuoka, Maurice. Multivariate Analysis. New York: John Wiley and Sons, 1971.

Thurstone, L. L. "A Law of Comparative Judgment." Psychological Review 34 (1927)(b): 273-286.

________. "Psychophysical Analysis." American Journal of Psychology 38 (1927)(a): 368-389.

Troutman, C. Michael, and Shanteau, James. "Do Consumers Evaluate Products by Adding or Averaging Attribute Information?" Journal of Consumer Research 3 (June 1976).

Tversky, Amos. "Choice by Elimination." Journal of Mathematical Psychology 9 (1972)(b): 341-367.

________. "Elimination by Aspects: A Theory of Choice." Psychological Review 79 (July 1972)(a): 281-299.

________. "Regret and Decision." Working paper, Department of Psychology, Hebrew University, Jerusalem, 1975.

Vatcu, M. "A Method of Solving a Linear Programming Problem with Several Objective Functions." Economic Computation and Economic Cybernetics Studies and Research (Rumania) 2 (1975): 97-106.

Wallenius, Jyriki. "Comparative Evaluation of Some Interactive Approaches to Multicriterion Optimization." Management Science 21 (August 1975): 1387-1396.

Wallenius, J., and Zionts, S. "Some Tests of an Interactive Programming Method for Multicriterion Optimization and an Attempt at Implementation." Working paper 75-3, European Institute for Advanced Studies in Management, Brussels, January 1975.

Wallenius, Jyriki, and Zionts, Stanley. "Decision-Making with Multiple Objectives - Some Analytic Approaches." Working paper, SUNY at Buffalo, March 1976.

Webster's New World Dictionary. Cleveland: World Publishing Co., 1956.

Weingartner, H. Martin. Mathematical Programming and the Analysis of Capital Budgeting Problems. Englewood Cliffs, N.J.: Prentice-Hall, 1963.

White, D. J. "Kernels of Preference Structures." Econometrica (forthcoming).

Wickens, Thomas D. "Attribute Elimination Strategies for Concept Identification with Practical Subjects." Journal of Mathematical Psychology 8 (1971): 453-480.

Widhelm, William B., and Doyle, Thomas H. "Scaling and Norming Considerations in Linear Goal Programming Models." Working paper, University of Maryland, 1976.

Wilde, Douglass J., and Beightler, Charles S. Foundations of Optimization. Englewood Cliffs, N.J.: Prentice-Hall, 1967.

Wilkie, William L., and McCann, John M. "The Halo Effect and Related Issues in Multi-Attribute Attitude Models--An Experiment." Working paper No. 377, Krannert School of Industrial Administration, Purdue University, October 1972.

Wilkie, William L., and Weinreich, Rolf P. "Effects of the Number and Type of Attributes Included in an Attitude Model: More is Not Better." Working paper No. 385, Krannert School of Industrial Administration, Purdue University, January 1973.

Winter, Sidney G. "Optimization and Evolution in the Theory of the Firm." In Adaptive Economics. Edited by Richard H. Day and Theodore Groves. New York: Academic Press, 1975.

Wishart, David. "CLUSTAN 1C." London: Computer Center, University College, July 1975.

Wright, Peter. "The Harassed Decision Maker: Time Pressures, Distractions, and the Use of Evidence." Journal of Applied Psychology 59 (November 1974): 555-561.

Yu, P. L. "A Class of Solutions for Group Decision Problems." Management Science 19 (April 1973): 936-946.

________. "Cone Convexity, Cone Extreme Points, and Nondominated Solutions in Decision Problems with Multiobjectives." Journal of Optimization and Applications 14 (September 1974): 319-377.

________. "Decision Dynamics, Persuasion and Negotiation." Working paper, Graduate School of Business, University of Texas, Austin, 1975.

Yu, P. L., and Leitman, G. "Compromise Solutions, Domination Structures and Salukvadze's Solution." Problems of Control and Information Theory 2 (1973): 183-197.

Yu, P. L., and Zeleny, Milan. "The Set of All Nondominated Solutions in Linear Cases and a Multicriteria Simplex Method." Journal of Mathematical Analysis and Applications 49 (February 1975): 430-468.

Zadeh, Lofti A. "Application of the Linguistic Approach to Decision-Making Under Multiple Criteria." Working paper, College of Engineering, University of California, Berkeley, 1975.

________. "Optimality and Nonscalar-valued Performance Criteria." IEEE Transactions on Automatic Control, AC-8. (1963): 59-60.

Zeleny, Milan. "The Attribute-Dynamics Model." Management Science 23 (September 1976)(c): 12-26.

________. "A Concept of Compromise Solutions and the Method of the Displaced Ideal." Computing and Operations Research 1 (1974) (b): 479-496.

________. "Conflict Dissolution." Paper presented at the 1976 TIMS-ORSA Meetings, Philadelphia, April 1976 (a).

________. Linear Multiobjective Programming. Lecture Notes in Economics and Mathematical Systems, No. 95. New York: Springer-Verlag, 1974 (a).

________. "Managers Without Management Science." Interfaces 5 (August 1975)(c): 35-42.

________. "MCDM - The State and Future of the Art." Management Science (1977).

________. "Self-Organization of Living Systems." International Journal of General Systems 4 (1977): 13-28.

________. "The Theory of the Displaced Ideal." Multiple Criteria Decision Making: Kyoto, 1975. New York: Springer-Verlag, 1976 (b): 151-205.

Zionts, Stanley, and Wallenius, Jyriki. "An Interactive Programming Method for Solving the Multiple Criteria Problem." Management Science 22 (February 1976): 652-663.

APPENDIX I

EXPLANATION OF HARTIGAN'S DIRECT CLUSTERING

(from Biomedical Computer Programs, BMD P Series, "Block Clustering, Program BMDP3M," adopted from computer output)

Purpose

The program represents the data completely with relatively few printed symbols. After permutation of rows and columns, a block diagram is printed out, and certain blocks, submatrices of contiguous values, are outlined. All values within a block may be recovered from the value in the upper left-hand corner, so only these calues are printed. The row margins of each block form a row cluster, the column margins form a column cluster, and the block itself is a cluster of equivalent data values.

Method

First code all variables into the Range 1,9,A,Z. The number of times each variable takes each value is computed. The matrix is computed so that frequently appearing values are pushed to the upper left corner, to reduce dependence of the final clusters on input order. A leader structure on rows and columns is next computed on the first pass through the data, the first row is a leader, and all subsequent rows which are not within threshold distance of a previously defined leader. If row I is not a leader, but is closest to Row J among previously defined leaders, say that J leads I at level 1 on all later passes, only the row leaders are used. On the second pass, column leaders are defined similarly, and only these columns are used in later passes. The passes continue, alternating on rows and columns, until finally, only the first row and column remain.

Rows and columns are now permuted, so that if I is a leader at pass K, no J greater than I has a leader less than I at pass k. A leader structure is next defined on the data values themselves. If II leads row I at level K and JJ leads column J at level L, then (II,J) leads (I,J) if K is less than L, and (I,JJ) leads (I,J) if K is greater than L. If the value at (I,J) may be predicted from the value for its leader, the value at (I,J) is not printed on the block diagram.

Thresholds

Thresholds specify the minimum distance between leaders at each pass. Most computation time is spent in passes with small thresholds. With more passes fewer blocks are necessary to represent the data. More than twice the number of variables is rarely necessary.

Codings

Each variable is coded onto the range 1 . . . 9, A . . . Z, according to one of the following options. N6 is the number of intervals for a variable, then (1) if N6 is greater than zero, the range is divided into N6 intervals of equal length. Values in the various intervals are coded 1, 2, . . . , N6. (2) if N6 equals zero, but the variable takes integer values between 1 and 35, the values are coded directly onto 1, . . . Z. (3) if N6 equals zero, but the variable takes values outside the range (1, . . . , 35) the first 35 different values are ordered and coded 1, . . . , Z, respectively.

Counts

The number of times each variable takes each of its coded values.

Leaders

Row and column leaders, after permutation of rows and columns, are printed out in tree form. To find the leader of a given row, go furthest east, north, west. The position of the north segment of this path is the pass number. Each T-intersection defines the cluster of all rows which reach the intersection by east or north movements.

Blocks

Using the permutation of rows and columns developed to display the leader structure, blocks are outlined and values are printed in the upper left corner of each block. From the prediction table specifying relations between values in different columns. The first row of the block may be predicted. All other rows in the block are identical to this one.

Any entry in the data matrix may be recovered from the block value of the smallest block containing it.

Missing values are represented by asterisks. A good clustering is printed/total values = 0.1. A mediocre reduction is printed/total values = 0.5.

Prediction

The prediction table is computed during the construction of column leaders. If column J leads column I, a prediction rule is given specifying a value in column I for each value in column J, the value of I most frequently associated with the given value of J.

The prediction rules are used for recovering data values in the block diagram. Thus, if column 6 takes value 3 in the upper left corner of a block, find the values of other columns, on that row in the prediction table where variable 6 takes value 3. Fill in the first row of the block from these values.

APPENDIX II

THE ZELENY PROBLEM

(from Zeleny [1974, pp. 117 ff.])

Maximize

$$3x_1 - 7x_2 + 4x_3 + x_4 - x_6 - x_7 + 8x_8$$

$$2x_1 + 5x_2 + x_3 - x_4 + 6x_5 + 8x_6 + 3x_7 - 2x_8$$

$$5x_1 - 2x_2 + 5x_3 + 6x_5 + 7x_6 + 2x_7 + 6x_8$$

$$4x_2 - x_3 - x_4 - 3x_5 \qquad x_8$$

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8$$

Subject to:

$$x_1 + 3x_2 - 4x_3 + x_4 - x_5 + x_6 + 2x_7 + 4x_8 \leqq 40$$

$$5x_1 + 2x_2 + 4x_3 - x_4 - 3x_5 + 7x_6 + 2x_7 + 7x_8 \leqq 84$$

$$4x_2 - x_3 - x_4 - 3x_5 + x_8 \leqq 18$$

$$-3x_1 - 4x_2 + 8x_3 + 2x_4 + 3x_5 - 4x_6 + 5x_7 - x_8 \leqq 100$$

$$12x_1 + 8x_2 - x_3 + 4x_4 + x_6 + x_7 \leqq 40$$

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 \geqq 12$$

$$8x_1 - 12x_2 - 3x_3 + 4x_4 - x_5 \leqq 30$$

$$-5x_1 - 6x_2 + 12x_3 + x_4 - x_7 + x_8 \leqq 100$$

$$x_i \geqq 0, \quad i = 1, \ldots, 8$$

Nondominated Solutions to the Zeleny Problem

Extreme Point	z_1	z_2	z_3	z_4	z_5
1	115.93	-28.75	87.18	-3.18	26.13
2	116.08	-29.07	87.00	-3.00	25.55
3	64.39	16.74	81.13	2.87	17.65
4	37.20	49.33	86.54	-2.54	22.46
5	110.84	-22.72	88.12	-4.12	27.12
6	82.72	28.52	111.24	-27.24	28.99
7	$\boxed{117.25}$	-27.75	89.50	-5.50	27.00
8	-17.73	106.18	88.45	-4.45	26.64
9	-37.52	111.59	74.07	9.92	22.63
10	-29.00	106.55	77.55	6.45	29.44
11	-12.09	102.56	90.46	-6.47	31.66
12	-19.09	125.21	106.13	-22.13	33.43
13	-37.72	135.90	98.17	-14.17	31.78
14	-36.53	159.20	122.66	-38.66	33.24
15	-35.00	173.00	138.00	-54.00	29.66
16	-2.78	86.78	84.00	0.00	12.70
17	-36.37	105.85	69.48	14.52	14.59
18	8.51	170.55	$\boxed{179.06}$	-95.06	$\boxed{39.35}$
19	10.00	168.94	178.94	-94.94	37.49
20	-0.50	$\boxed{176.83}$	176.33	-92.33	38.61
21	5.33	173.33	178.66	-94.66	35.11
22	24.00	150.00	174.00	-90.00	39.00
23	85.84	38.35	124.18	-40.18	33.56
24	95.40	-1.38	94.01	-10.01	31.08
25	31.75	51.42	83.17	0.83	26.22
26	7.05	63.63	70.68	13.31	28.11
27	77.84	12.10	89.94	-5.94	31.30
28	35.74	82.51	118.25	-34.25	34.60
29	9.31	93.48	102.80	-18.80	29.03
30	86.73	-11.89	74.84	9.16	13.08

	z_1	z_2	z_3	z_4	z_5
31	66.34	-0.34	66.00	18.00	15.00
32	66.56	-0.56	66.00	18.00	19.84
33	72.00	-6.00	66.00	18.00	14.62
34	30.40	35.60	66.00	18.00	14.00
35	33.12	32.88	66.00	18.00	14.35
36	15.89	50.11	66.00	18.00	16.30
37	-17.73	83.73	66.00	18.00	14.80
38	29.34	36.65	66.00	18.00	21.37
39	30.43	35.57	66.00	18.00	21.36
40	30.62	35.38	66.00	18.00	21.42
41	29.72	36.28	66.00	18.00	21.50
42	56.30	-4.70	51.61	18.00	12.00
43	48.85	-0.77	49.63	18.00	12.00
44	40.80	12.20	53.00	18.00	12.00
45	47.13	2.71	49.84	18.00	12.00
46	72.00	2.67	74.67	9.33	12.00
47	95.27	-21.27	74.00	10.00	12.55
48	19.55	46.46	66.00	18.00	24.44
49	-9.06	75.06	66.00	18.00	17.88
50	117.19	-24.95	92.24	-8.24	28.09
51	-13.34	99.19	85.85	-1.85	28.22
52	-19.18	118.22	99.00	-15.04	29.75
53	49.25	35.93	85.18	-1.18	21.20
54	84.66	-8.00	76.60	7.33	13.30
55	-18.00	102.00	84.00	0.00	24.00
56	86.80	15.60	102.40	-18.40	24.80
57	91.60	-7.50	84.00	0.00	15.30
58	15.88	55.14	71.03	12.96	25.57
59	36.34	101.60	138.00	-54.00	34.34
60	27.40	73.40	100.80	-16.80	31.13
61	23.13	69.95	93.00	-9.07	22.24
62	31.75	100.21	131.90	-47.96	23.12
63	16.00	140.00	156.00	-72.00	26.66
64	49.28	77.60	126.80	-42.80	22.50
65	-20.72	121.58	100.86	-16.86	31.80

	z_1	z_2	z_3	z_4	z_5
66	47.80	46.12	93.94	-9.94	17.36
67	-0.17	71.74	71.57	12.42	21.00
68	-12.00	96.00	84.00	0.00	12.00
69	81.37	-6.40	75.00	9.00	12.94
70	26.00	66.23	92.33	-8.33	15.57

There are 70 different N^{ex} points. Individual maxima of all objectives are boxed in to simplify the review of data.

APPENDIX III

THE STEUER PROBLEM

(from Steuer [1975, pp. 18, 26, and 27])

Maximize

$$
\begin{aligned}
&-4x_1 - 2x_2 + x_3 + 2x_4 - 4x_5 - 3x_6 + 2x_7 \\
&\qquad - 3x_2 - 4x_3 + 4x_4 + 5x_5 - 2x_6 + x_7 \\
&5x_1 + 5x_2 - 2x_4 + 3x_5 + 5x_7 \\
&3x_1 - 3x_2 + 5x_3 + 2x_5 + 2x_6 - 4x_7
\end{aligned}
$$

Subject to:

$$
\begin{aligned}
7x_1 + 6x_3 + 2x_4 + 5x_5 &\leqq 100 \\
4x_1 + 7x_6 + 9x_7 &\leqq 100 \\
5x_1 + 6x_4 &\leqq 100 \\
9x_3 + 4x_7 &\leqq 100 \\
2x_1 + 8x_2 + 5x_3 &\leqq 100 \\
x_1 + 2x_2 + 2x_3 + 6x_4 + 5x_5 + 8x_6 + 5x_7 &\leqq 100 \\
5x_3 + 3x_5 + 7x_6 &\leqq 100
\end{aligned}
$$

$$x_i \geqq 0, \quad i = 1, 2, \ldots, 7$$

Nondominated Solutions of the Steuer Problem

Extreme Point	z_1	z_2	z_3	z_4
1	-80.	100.	60.	40.
2	-96.92	38.46	120.76	6.15
3	-30	71.42	31.42	32.85
4	-87.02	8.10	68.64	58.37
5	-52.97	60.81	105.94	8.91
6	33.33	66.66	-33.33	0.
7	-13.33	55.55	82.22	-26.66
8	-99.87	-30.32	116.83	28.64
9	-71.59	11.78	153.37	-16.75
10	24.44	11.11	-15.55	60.
11	-15.55	-11.11	20.	68.88
12	-44.89	10.20	30.61	69.38
13	-19.34	24.59	63.27	35.08
14	19.23	23.07	-7.69	50.
15	-21.98	48.07	75.38	10.19
16	-58.10	-41.41	43.81	77.19
17	-15.61	17.12	86.02	20.13
18	37.03	7.40	-25.92	55.55
19	37.03	40.74	40.74	-44.44
20	-18.33	-6.94	129.72	-74.16
21	2.71	18.51	74.81	-0.74
22	-76.53	-55.71	87.84	48.97
23	-93.36	-39.03	116.07	28.31
24	-26.41	-17.48	116.80	3.49
25	-65.47	-22.02	139.88	-2.97
26	-16.54	-13.	10.16	74.46
27	32.22	6.66	-20.	58.88
28	-26.03	-18.09	20.	75.87
29	-7.93	-44.44	23.80	69.84
30	-27.89	-6.48	34.54	64.90
31	-2.97	44.04	61.90	-17.85
32	-5.70	31.17	65.43	14.35
33	-57.86	-49.63	48.42	74.57
34	-55.44	-46.	42.37	78.20
35	-29.76	-47.52	44.91	68.26

Extreme Point	z_1	z_2	z_3	z_4
36	-7.69	-23.07	76.92	30.76
37	11.81	3.59	60.31	22.03
38	39.09	7.81	44.85	-13.58
39	3.70	-13.42	111.57	-81.94
40	-2.77	-26.38	118.05	-81.94
41	-0.74	-24.69	107.65	-33.58
42	-54.21	-66.88	61.49	63.85
43	-34.91	-60.04	76.41	45.88
44	-6.73	-23.39	96.39	6.36
45	-19.23	-40.38	105.76	13.46
46	-14.76	-41.58	10.47	79.20
47	-26.98	-57.14	23.80	82.53
48	-31.70	-54.12	48.79	66.68
49	-29.62	-51.00	43.58	69.80
50	16.04	-29.62	93.82	-39.50
51	11.11	-39.50	98.76	-39.50
52	-38.05	-67.32	59.89	60.33

APPENDIX IV

DIRECT CLUSTERING OF ZELENY PROBLEM

(partial output)

Code for Each Variable

FIRST	REAL	[1] -38.00 [2]	40.00
SECOND	REAL	-30.00	73.50
THIRD	REAL	49.00	114.50
FOURTH	REAL	-96.00	-38.50
FIFTH	REAL	11.00	25.50

ROW PASS	THRESHOLD	COL PASS	THRESHOLD
1	.125	2	.125
3	.250	4	.250
5	.375	6	.375
7	.500	8	.500
9	.625	10	.625
11	.750	12	.750
13	.875	14	.875
15	1.000	16	1.000

Code Frequencies

VARIABLE	MISSING	1	2
FIRST	0	43	27
SECOND	0	41	29
THIRD	0	57	13
FOURTH	0	12	58
FIFTH	0	38	32

```
T R E E   G I V I N G   J O I N I N G   S E Q U E N C E
PASS                 1111111111222222222
NUMBER   0123456789012345678901234567891
------   -------------------------------
FOURTH   -,-----,-------,--
THIRD    -/     I       I
SECOND   -------I       I
FIRST    -------/       I
FIFTH    ---------------/
```

```
T R E E   G I V I N G   J O I N I N G   S E Q U E N C E
PASS                 1111111111
NUMBER   01234567890123456789
------   --------------------
       4 --,---,---,---,--
      70 --I   I   I   I
      67 --I   I   I   I
      61 --I   I   I   I
      48 --I   I   I   I
      41 --I   I   I   I
      40 --I   I   I   I
      39 --I   I   I   I
      38 --I   I   I   I
      36 --I   I   I   I
      35 --I   I   I   I
      34 --/   I   I   I
      25 --,---I   I   I
      58 --I   I   I   I
      26 --/   I   I   I
       9 --,---I   I   I
      68 --I   I   I   I
      55 --I   I   I   I
      49 --I   I   I   I
      37 --I   I   I   I
      17 --I   I   I   I
      16 --/   I   I   I
       3 --,---/   I   I
      69 --I       I   I
      66 --I       I   I
      57 --I       I   I
      56 --I       I   I
      54 --I       I   I
      53 --I       I   I
      47 --I       I   I
      46 --I       I   I
      45 --I       I   I
      44 --I       I   I
      43 --I       I   I
      42 --I       I   I
      33 --I       I   I
      32 --I       I   I
      31 --I       I   I
      30 --/       I   I
       1 --,---,-,-/   I
      50 --I   I I     I
      27 --I   I I     I
      24 --I   I I     I
       7 --I   I I     I
       6 --I   I I     I
       5 --I   I I     I
       2 --/   I I     I

      23 ------/ I     I
       8 --,---,-/     I
      28 --I   I       I
      65 --I   I       I
      60 --I   I       I
      52 --I   I       I
      51 --I   I       I
      29 --I   I       I
      13 --I   I       I
      12 --I   I       I
      11 --I   I       I
      10 --/   I       I

      33 --/       I   I
       1 --,---,-,-/   I
      50 --I   I I     I
      27 --I   I I     I
      24 --I   I I     I
       7 --I   I I     I
       6 --I   I I     I
       5 --I   I I     I
       2 --/   I I     I
      23 ------/ I     I
       8 --,---,-/     I
      28 --I   I       I
      65 --I   I       I
      60 --I   I       I
      52 --I   I       I
      51 --I   I       I
      29 --I   I       I
      13 --I   I       I
      12 --I   I       I
      11 --I   I       I
      10 --/   I       I
      14 --,---/       I
      63 --I           I
      59 --I           I
      22 --I           I
      21 --I           I
      20 --I           I
      19 --I           I
      18 --I           I
      15 --/           I
      62 ------,-------/
      64 ------/
```

```
     B L O C K   D I A G R A M
NUMBER OF DATA POINTS.....  =   350
NUMBER OF BLOCKS..........  =    12
NUMBER OF BLOCKS,NOT SINGLE =    10

       FT  S  F  F
       OH  E  I  I
       UI  C  R  F
       FR  O  S  T
       TD  N  T  H
       H   D

   ,-------------,
  4( 2           )
 70(             )
 67(             )
 61(             )
 48(             )
 41(             )
 40(             )
 39(             )
 38(             )
 36(             )
 35(             )
 34(             )
   (          ,-,)
 25(          (2))
 58(          ( ))
 26(          ( ))
   (          "-")
   (    ,-,      )
  9(    (2)      )
 68(    ( )      )
 55(    ( )      )
 40(    ( )      )
 37(    ( )      )
 17(    ( )      )
 16(    ( )      )
   (    "-"      )
   (       ,-,   )
  3(       (2)   )
 69(       ( )   )
 66(       ( )   )
 57(       ( )   )
 56(       ( )   )
 54(       ( )   )
 53(       ( )   )
 47(       ( )   )
 46(       ( )   )
 45(       ( )   )
 44(       ( )   )
 43(       ( )   )
 42(       ( )   )
 33(       ( )   )
 32(       ( )   )
 31(       ( )   )
 30(       ( )   )
   (       "-"   )
   (       ,-,,-,)
  1(       (2)(2))
 50(       ( )( ))
 27(       ( )( ))
 24(       ( )( ))
  7(       ( )( ))
  6(       ( )( ))
  5(       ( )( ))
  2(       ( )( ))
   (,--,   ( )( ))
 23((1 )   ( )( ))
   ("--"   "-"( ))
   (    ,-,   ( ))
  8(    (2)   ( ))
 28( 2  ( )   ( ))
 65(    ( )   ( ))
 60(    ( )   ( ))
 52(    ( )   ( ))
 51(    ( )   ( ))
 29(    ( )   ( ))
 13(    ( )   ( ))
 12(    ( )   ( ))
 11(    ( )   ( ))
 10(    ( )   ( ))
   (,--,( )   ( ))
 14((1 )( )   ( ))
 63((  )( )   ( ))
 59((  )( )   ( ))
 22((  )( )   ( ))
 21((  )( )   ( ))
 20((  )( )   ( ))
 19((  )( )   ( ))
 18((  )( )   ( ))
 15((  )( )   ( ))
   ("--""-"   "-")
   (,--------,   )
 62((1       )   )
 64((       2)   )
   ("--------"   )
   "-------------"
```

```
B L O C K   D I A G R A M
NUMBER OF DATA POINTS.....   =    350
NUMBER OF BLOCKS..........   =     12
NUMBER OF BLOCKS,NOT SINGLE  =     10

          FT   S   F   F
          OH   E   I   I
          UI   C   R   F
          TR   O   S   T
          TD   N   T   H
          H    D

      ,--------------,
     4( 2            )
    70(              )
    67(              )
    61(              )
    48(              )
    41(              )
    40(              )
    39(              )
    38(              )
    36(              )
    35(              )
    34(              )
      (          ,-,)
    25(          (2))
    58(          ( ))
    26(          ( ))
      (          "-")
      (    ,-,      )
     9(    (2)      )
    68(    ( )      )
    55(    ( )      )
    40(    ( )      )
    37(    ( )      )
    17(    ( )      )
    16(    ( )      )
      (    "-"      )
      (       ,-,   )
     7(       (2)   )
    69(       ( )   )
    66(       ( )   )
    57(       ( )   )
    56(       ( )   )
    54(       ( )   )
    53(       ( )   )
    47(       ( )   )
    46(       ( )   )
    45(       ( )   )
    44(       ( )   )
    43(       ( )   )
    42(       ( )   )
    33(       ( )   )
    32(       ( )   )
    31(       ( )   )
    30(       ( )   )
      (       "-"   )
      (       ,-,,-,)
     1(       (2)(2))
    50(       ( )( ))
    27(       ( )( ))

    24(          ( )( ))
     7(          ( )( ))
     6(          ( )( ))
     5(          ( )( ))
     2(          ( )( ))
      (,--,      ( )( ))
    23((1 )      ( )( ))
      ("--"      "-"( ))
      (      ,-,    ( ))
     8(      (2)    ( ))
    28(   2  ( )    ( ))
    65(      ( )    ( ))
    60(      ( )    ( ))
    52(      ( )    ( ))
    51(      ( )    ( ))
    29(      ( )    ( ))
    13(      ( )    ( ))
    12(      ( )    ( ))
    11(      ( )    ( ))
    10(      ( )    ( ))
      (,--,  ( )    ( ))
    14((1 )  ( )    ( ))
    63((  )  ( )    ( ))
    59((  )  ( )    ( ))
    22((  )  ( )    ( ))
    21((  )  ( )    ( ))
    20((  )  ( )    ( ))
    19((  )  ( )    ( ))
    18((  )  ( )    ( ))
    15((  )  ( )    ( ))
      ("--""-"      "-")
      (,--------,      )
    62((1       )      )
    64((       2)      )
      ("--------"      )
      "-----------------"
```

P R E D I C T I O N   T A B L E

```
F T S F F
O H E I I
U I C R F
R R O S T
T D N T H
H   D

1 2 2 1 -
2 1 1   .
```

APPENDIX V

DIRECT CLUSTERING OF THE STEUER PROBLEM

```
          PROGRAM REVISED FEBRUARY 7, 1975
          MANUAL DATE  --  1975


 BMDP3M - BLOCK CLUSTERING
 HEALTH SCIENCES COMPUTING FACILITY
 UNIVERSITY OF CALIFORNIA, LOS ANGELES

 IN THIS VERSION OF BMDP3M
     -- THE SAVE FILE IS UNAVAILABLE FOR INPUT OR OUTPUT


   PROGRAM CONTROL INFORMATION

           PROBL
                   TITLE IS "CLUSTER ANALYSIS OF STEUR DATA"./
           INPUT     VARIABLES ARE 4.
                    FORMAT IS"(6X,4F6.2)".
                    CASE=52./
           VARIABLE   NAMES ARE FIRST,SECOND,THIRD,FOURTH.
                 MAXIMUMS ARE (1)40,(2)100.0,(3)154,(4)83.
                      MINIMUMS ARE (1)-100,(2)-68,(3)-34,(4)-82.
                  INTERVAL=5,5,5,5 /
           PASS
               TOTAL=16./
           END/

   PROBLEM TITLE . . . . . . .CLUSTER ANALYSIS OF STEUR DATA

   NUMBER OF VARIABLES TO READ IN. . . . . . . . . .           4
   NUMBER OF VARIABLES ADDED BY TRANSFORMATIONS. .             0
   TOTAL NUMBER OF VARIABLES . . . . . . . . . . . .           4
   NUMBER OF CASES TO READ IN. . . . . . . . . . . .          52
   CASE LABELING VARIABLES . . . . . . . . . . . . .
   LIMITS AND MISSING VALUE CHECKED BEFORE TRANSFORMATIONS
   INPUT UNIT NUMBER . . . . . . . . . . . . . . . .           5
   REWIND INPUT UNIT PRIOR TO READING . DATA. . .             NO

   INPUT FORMAT
         (6X,4F6.2)

   VARIABLES TO BE USED
               1  FIRST          2  SECOND          3  THIRD          4  FOURTH

  MAXIMUM NUMBER OF CASES THAT CAN BE CLUSTERED(ASSUMING 250 BLOCKS)=  454

   NUMBER OF CASES READ. . . . . . . . . . . . . . .          52
```

C O D E   F O R   E A C H   V A R I A B L E

		1	2	3	4	5
FIRST	REAL	-100.00	-72.00	-44.00	-16.00	12 00
SECOND	REAL	-68.00	-34.40	-.80	32.80	66.40
THIRD	REAL	-34.00	3.60	41.20	78.80	116.40
FOURTH	REAL	-82.00	-49.00	-16.00	17.00	50.00

ROW PASS	THRESHOLD	COL PASS	THRESHOLD
1	.125	2	.125
3	.250	4	.250
5	.375	6	.375
7	.500	8	.500
9	.625	10	.625
11	.750	12	.750
13	.875	14	.875
15	1.000	16	1.000

T R E E   G I V I N G   J O I N I N G   S E Q U E N C E

```
PASS                      1111111111
NUMBER        01234567890123456789
------        --------------------
          35  --,---,-,-,-,-,--
          52  --I   I I I I I
          49  --I   I I I I I
          48  --/   I I I I I
          47  ------I I I I I
          43  ------I I I I I
          16  --,---/ I I I I
          42  --I     I I I I
          34  --I     I I I I
          33  --/     I I I I
          45  --------I I I I
          29  --,-----I I I I
          46  --/     I I I I
          15  --------I I I I
          13  ------,-I I I I
          37  ------/ I I I I
           4  --------/ I I I
          26  --,---,-,-/ I I
          30  --I   I I   I I
          28  --/   I I   I I
          11  ------/ I   I I
          12  --------I   I I
           3  --------/   I I
          51  ------,---,-I I
          41  ------/   I I I
          50  ----------/ I I
          20  ------,---,-I I
          40  ------I   I I I
          24  ------/   I I I
          39  ------,---I I I
          44  ------/   I I I
          25  ----------/ I I
           1  --------,-,-/ I
          22  ------,-/ I   I
          23  ------/   I   I
           8  ----------I   I
          36  ----------/   I
           9  --------------I
          21  --,---,-,---,-I
          32  --/   I I   I I
          38  ------/ I   I I
          17  --------I   I I
          31  ------,-/   I I
           7  ------/     I I
           2  --------,---/ I
           5  --------/     I
          10  --,-----,---,-/
          27  --I     I   I
          18  --I     I   I
          14  --/     I   I
           6  --------/   I
          19  ------------/
```

C O D E   F R E Q U E N C I E S

VARIABLE	MISSING	1	2	3	4	5
FIRST	0	6	8	15	15	8
SECOND	0	15	14	13	7	3
THIRD	0	5	10	19	11	7
FOURTH	0	3	7	11	10	21

T R E E   G I V I N G   J O I N I N G   S E Q U E N C E

```
PASS              1111111111222222222
NUMBER    0123456789012345678901234567890
------    ------------------------------
FOURTH    --------------,-,--
FIRST     --------------I I
THIRD     --------------/ I
SECOND    ----------------/
```

```
 B L O C K   D I A G R A M
NUMBER OF DATA POINTS..... =   208
NUMBER OF BLOCKS.......... =    69
NUMBER OF BLOCKS,NOT SINGLE =   28

       F F   T   S
       O I   H   E
       U R   I   C
       R S   R   O
       T T   D   N
       H         D

  ,-------------------,
  (       ,-----,     )
35(  5    (  3  )     )
52(       (     )     )
49(       (     )     )
48(       (     )     )
47(       (  2  )     )
43(  4    (     )     )
  (    ,-,(     )     )
16(    (2)(     )     )
42(    ( )(     )     )
34(    ( )(     )     )
33(    ( )(     )     )
  (    "-"(     )     )
45(  3    (  4  )     )
  (    ,-,( ,-, )     )
29(    (4)( (2) )     )
46(    ( )( ( ) )     )
  (    "-"( "-" )     )
15(  3    (     )  4  )
  ( ,-,   (     ) ,-, )
13( (4)   (     ) (3) )
37( ( ) 4 (     ) ( ) )
  ( "-"   (     ) "-" )
 4(     1 (     )  3  )
  (       ( ,-, ) ,-, )
26(       ( (2) ) (2) )
30(       ( ( ) ) ( ) )
28(       ( ( ) ) ( ) )
11(     4 ( ( ) ) ( ) )
  (       ( ( ) ) "-" )
12(     2 ( ( ) )  3  )
 3(  4    ( ( ) )  5  )
  (       ( "-" )     )
  ( ,-,,-,( ,-, )     )
51( (2)(4)( (4) )     )
41( ( )( )( ( ) )  2  )
  ( ( )"-"( ( ) )     )
50( ( ) 5 ( ( ) )  2  )
  ( "-"   ( "-" )     )
  ( ,-,   (,---,) ,-, )
20( (1)   (( 5 )) (2) )
40( ( ) 4 ((   )) ( ) )
24( (3)   ((   )) ( ) )
  ( ( ),-,((,-,)) ( ) )
39( ( )(4)(((4))) ( ) )
44( (3)( )((( ))) ( ) )
  ( ( )"-"(("-")) ( ) )
25( (3) 2 ((   )) ( ) )
  ( "-"   ("---") "-" )
```

```
  ( ,-,,-,(     )     )
 1( (4)(1)(     )  5  )
  ( ( )( )( ,-, )     )
22( ( )( )( (4) )     )
23( ( )( )( ( ) )  4  )
  ( ( )( )( "-" )     )
 8( ( )( )(  5  )  2  )
36( ( )(4)(     )  2  )
  ( "-""-""-----"     )
  ( ,---------,       )
 9( (2        )       )
  ( "---------"       )
  (,----------,  ,---,)
21(( 3        )  ( 3 ))
32((          )  (   ))
38((    5     )  (   ))
17(( 4       4)  (   ))
  ((,-,       )  (,-,))
31(((2)       )  ((4)))
 7((( )      4)  (( )))
  (("-"       )  ("-"))
  ((          )  "---")
  ((          )   ,-, )
 2((    1    5)   (4) )
 5((    2    4)   ( ) )
  ("----------"   "-" )
  (    ,-,        ,-, )
10(    (5)        (3) )
27(    ( )        ( ) )
18(    ( )        ( ) )
14(    ( )        ( ) )
  (    ( )        "-" )
 6(  3 ( )         5  )
19(  2 ( )    2    4  )
  (    "-"            )
  "-------------------"
```

```
P R E D I C T I O N   T A B L E

F F T S
O I H E
U R I C
R S R O
T T D N
H     D

2 2 5 -
3 4 3 3
5 3 1 1
```

APPENDIX VI

A COMPARISON OF THREE HIERARCHICAL CLUSTERING ALGORITHMS

ON THE ZELENY DATA

	Ward's Method		Group Average Method		Centroid Method	
(with 10 clusters specified)						
#1	1		1	50	1	
	2		2	56	2	
	5		5		5	
	7		6		6	
	50		7		7	
			24		24	
			27		27	
					50	
					56	
#2	3	46	3	44	3	42
	30	47	30	45	4	43
	31	54	31	46	25	44
	32	57	32	47	26	45
	33	69	33	54	30	46
			34	57	31	47
			35	59	32	48
			36		33	53
			42		34	54
			43		35	57
					38	61
					39	66
					40	69
					41	70
#3	4	48	4	48	8	51
	25	53	25	53	9	52
	26	58	26	58	10	55
	38	61	38	61	11	60
	39	66	39	66	12	65
	40	70	40	70	13	
	41		41		29	

#4	6		8	55	14
	23		9	60	15
	24		10	65	
	27		11		
	56		12		
			13		
			29		
			51		
			52		
#5	8	51	14		16
	10	52	15		17
	11	55			37
	12	65			49
	13				68
#6	9	68	16		18
	16		17		19
	17		37		20
	37		49		21
	49		67		22
	67		68		
#7	14		18		23
	15		19		
			20		
			21		
			22		
#8	18		23		28
	19		28		59
	20		59		
	21				
	22				
#9	28	63	62		62
	29	64	64		64
	59				
	60				
	62				
#10	34		63		63
	35				
	36				
	42				
	43				
	44				
	45				

(with 9 clusters specified)

#1	1	24		1			1	24	
	2	27		2			2	27	
	5	50		5			5	50	
	6	56		6			6	56	
	7			7			7		
	23			24					
				27					
				50					
				56					
#2	3	47		3	36		3	41	
	30	54		30	42		4	42	
	31	57		31	43		25	43	
	32	59		32	44		26	44	
	33			33	45		30	45	
	46			34	46		31	46	
				35	47		32	47	
					54		33	48	
					57		34	53	
					59		35	54	
							36	57	
							38	58	
							39	61	
							40	66	
								69	
								70	
#3	4	40	61	4	40	61	8	13	60
	25	41	66	25	41	66	9	29	65
	26	48	70	26	48	70	10	51	
	38	53		38	53		11	52	
	39	58		39	58		12	55	
#4	8	51		8	13	60	14		
	10	52		9	29	65	15		
	11	55		10	51				
	12	65		11	52				
	13			12	55				
#5	9	67		14			16	68	
	16	68		15			17		
	17						37		
	37						49		
	49						67		

#6	14 15	16 17 37 49 67 68	18 19 20 21 22
#7	18 19 20 21 22	18 19 20 21 22	23
#8	28 29 59 60 62 63 64	23 28 59	28 59 62 64
#9	34 35 36 42 43 44 45	62 63 64	63

(with 8 clusters specified)

#1	1 2 5 6 7 23 24 27 50 56	1 2 5 6 7 24 27 50 56	1 2 5 6 7 24 27 50 56
#2	3 30 31 32 33 34 35 36 42 43 44 45 46 47 54 57 69	3 4 25 26 30 31 32 33 34 35 36 38 39 40 41 42 43 44 45 46 58 61 66 69 70	(Same as cluster #2 when 9 clusters were specified)
#3	4 25 26 38 39 40 41 48 53 58 61 66 70	8 9 10 11 12 13 29 51 52 55 60 65	8 9 10 11 12 13 29 51 52 55 60 65
#4	8 10 11 12 13 51 52 55 65	14 15	14 15

#5	9 67 16 68 17 37 49	16 68 17 37 49 67	16 68 17 37 49 67
#6	14 15	18 19 20 21 22	18 19 20 21 22
#7	18 19 20 21 22	23 28 59	23 28 59 62 64
#8	28 63 29 64 59 60 62	62 63 64	63

(with 7 clusters specified)

#1	1 23 2 24 5 27 6 50 7 56	1 24 2 27 5 50 6 56 7	23 28 59 62 64
#2	3 34 44 57 30 35 45 69 31 36 46 32 42 47 33 43 54	(Same as cluster #2 when 8 clusters were specified)	(Same as cluster #2 when 8 clusters were specified, plus 16 17 37 49 67 68)
#3	4 40 61 25 41 66 26 48 70 38 53 39 58	8 13 49 65 9 16 51 67 10 17 52 68 11 29 55 12 37 60	8 13 60 9 29 65 10 51 11 52 12 55

#4	8 51 10 52 11 55 12 65 13	14 15	14 15
#5	9 67 16 68 17 37 49	18 19 20 21 22	(Same as group average method)
#6	14 60 15 62 28 63 29 64 59	23 28 29	23 28 29 59 62 64
#7	18 19 20 21 22	62 63 64	63

(with 6 clusters specified)

#1	1 23 2 24 5 27 6 50 7 56	1 24 2 27 5 50 6 56 7	1 24 2 27 5 50 6 56 7
#2	(Same as when 7 clusters were specified)	(Same as when 7 clusters were specified)	(Same as when 7 clusters were specified)
#3	4 26 41 61 9 37 48 66 16 38 49 67 17 39 53 68 25 40 58 70	8 13 49 65 9 16 51 67 10 17 52 68 11 29 55 12 37 60	8 13 60 9 29 65 10 51 11 52 12 55
#4	8 51 10 52 11 56 12 63 13	14 15	14 15

#5	14 60 15 62 28 63 29 64 59	18 19 20 21 22	18 19 20 21 22
#6	18 19 20 21 22	23 64 28 59 62 63	23 64 28 59 62 63

(with 5 clusters specified)

#1	1 23 2 24 5 27 6 50 7 56	1 27 40 53 2 30 41 54 3 31 42 56 4 32 43 57 5 33 44 58 6 34 45 61 7 35 46 66 24 36 47 69 25 38 48 70 26 39 50	(Same as when 6 clusters were specified)
#2	(Same as when 6 clusters were specified)	8 13 49 65 9 16 51 67 10 17 52 68 11 29 55 12 37 60	(Same as when 6 clusters were specified)
#3	4 26 41 61 9 37 48 66 16 38 49 67 17 39 53 68 25 40 58 70	14 15	8 13 52 9 14 55 10 15 60 11 29 65 12 51
#4	8 14 52 63 10 15 55 64 11 28 59 65 12 29 60 13 51 62	18 19 20 21 22	18 19 20 21 22
#5	18 19 20 21 22	23 64 28 59 62 63	23 64 28 59 62 63

APPENDIX VII

WARD'S METHOD ON ZELENY PROBLEM

(partial output)

```
 CCCCC  LL       U     UU  SSSSS  TTTTTT  AAAAA  NN     N         II    CCCCC
CC    C LL       U     UU SS     S   TT   A     AA NNN   N        III   CC     C
CC      LL       U     UU SSS        TT   A     AA N NN  N       IIII   CC
CC      LL       U     UU  SSSSS     TT   A     AA N  NN N        II    CC
CC      LL       U     UU      SSS   TT   AAAAAAA N   NNN         II    CC
CC    C LL       U     UU S     SS   TT   A     AA N    NN        II    CC     C
 CCCCC  LLLLLLL  UUUUU     SSSSS     TT   A     AA N     N       IIIIII  CCCCC

=============================================================================

RELEASE 1 OF VERSION 1C.
JULY 1975

=============================================================================

COMPUTER CENTRE
UNIVERSITY COLLEGE LONDON
19 GORDON STREET
LONDON WC1H 0AH

=============================================================================



=============================================================================

CONVERTED TO UNIVERSITY OF DELAWARE          B6700 MARCH 1977
BY UDCC RESEARCH SERVICES

=============================================================================


PROCEDURE FILE

RUN WITH ZELENY DECK


NUMBER OF CASES =    70
NUMBER OF BINARY VARIABLES =     0
NUMBER OF NUMERIC VARIABLES =     5

FILE DATA MATRIX


READ AND FILE RAW NUMERIC DATA
REPLACE NUMERIC DATA BY STANDARD SCORES
PRINCIPAL COMPONENTS SOLUTION SELECTED FOR NUMERIC DATA
INPUT DATA FORMAT(6X,5F6.0)
RAW DATA FILED
NUMERIC MEANS FILED
NUMERIC VARIANCES FILED
FILE EIGENVALUES
STANDARD SCORES FILED
  5 COMPONENT SCORES FILED
FILE COMPLETE
JOB ENDS
```

```
WARDS METHOD

OUTPUT CLASSIFICATIONS FOR    5 TO  10 CLUSTERS

CYCLE   1 NOW FUSE POINTS  39  40 AT COEFFICIENT    0.000 -  69 CLUSTERS AND NEW CLUSTER CODE IS  39
CYCLE   2 NOW FUSE POINTS  38  41 AT COEFFICIENT    0.000 -  68 CLUSTERS AND NEW CLUSTER CODE IS  38
CYCLE   3 NOW FUSE POINTS  38  39 AT COEFFICIENT    0.000 -  67 CLUSTERS AND NEW CLUSTER CODE IS  38
CYCLE   4 NOW FUSE POINTS  43  45 AT COEFFICIENT    0.001 -  66 CLUSTERS AND NEW CLUSTER CODE IS  43
CYCLE   5 NOW FUSE POINTS   1   2 AT COEFFICIENT    0.001 -  65 CLUSTERS AND NEW CLUSTER CODE IS   1
CYCLE   6 NOW FUSE POINTS  34  35 AT COEFFICIENT    0.002 -  64 CLUSTERS AND NEW CLUSTER CODE IS  34
CYCLE   7 NOW FUSE POINTS  54  69 AT COEFFICIENT    0.003 -  63 CLUSTERS AND NEW CLUSTER CODE IS  54
CYCLE   8 NOW FUSE POINTS  30  54 AT COEFFICIENT    0.004 -  62 CLUSTERS AND NEW CLUSTER CODE IS  30
CYCLE   9 NOW FUSE POINTS  31  33 AT COEFFICIENT    0.006 -  61 CLUSTERS AND NEW CLUSTER CODE IS  31
CYCLE  10 NOW FUSE POINTS   5   7 AT COEFFICIENT    0.007 -  60 CLUSTERS AND NEW CLUSTER CODE IS   5
CYCLE  11 NOW FUSE POINTS  18  19 AT COEFFICIENT    0.011 -  59 CLUSTERS AND NEW CLUSTER CODE IS  18
CYCLE  12 NOW FUSE POINTS   5  50 AT COEFFICIENT    0.012 -  58 CLUSTERS AND NEW CLUSTER CODE IS   5
CYCLE  13 NOW FUSE POINTS  42  43 AT COEFFICIENT    0.013 -  57 CLUSTERS AND NEW CLUSTER CODE IS  42
CYCLE  14 NOW FUSE POINTS  52  65 AT COEFFICIENT    0.016 -  56 CLUSTERS AND NEW CLUSTER CODE IS  52
CYCLE  15 NOW FUSE POINTS   8  51 AT COEFFICIENT    0.016 -  55 CLUSTERS AND NEW CLUSTER CODE IS   8
CYCLE  16 NOW FUSE POINTS  16  68 AT COEFFICIENT    0.016 -  54 CLUSTERS AND NEW CLUSTER CODE IS  16
CYCLE  17 NOW FUSE POINTS  48  58 AT COEFFICIENT    0.020 -  53 CLUSTERS AND NEW CLUSTER CODE IS  48
CYCLE  18 NOW FUSE POINTS  18  20 AT COEFFICIENT    0.021 -  52 CLUSTERS AND NEW CLUSTER CODE IS  18
CYCLE  19 NOW FUSE POINTS   1   5 AT COEFFICIENT    0.029 -  51 CLUSTERS AND NEW CLUSTER CODE IS   1
CYCLE  20 NOW FUSE POINTS   4  53 AT COEFFICIENT    0.032 -  50 CLUSTERS AND NEW CLUSTER CODE IS   4
CYCLE  21 NOW FUSE POINTS  42  44 AT COEFFICIENT    0.033 -  49 CLUSTERS AND NEW CLUSTER CODE IS  42
CYCLE  22 NOW FUSE POINTS  30  47 AT COEFFICIENT    0.036 -  48 CLUSTERS AND NEW CLUSTER CODE IS  30
CYCLE  23 NOW FUSE POINTS  37  49 AT COEFFICIENT    0.042 -  47 CLUSTERS AND NEW CLUSTER CODE IS  37
CYCLE  24 NOW FUSE POINTS  24  27 AT COEFFICIENT    0.051 -  46 CLUSTERS AND NEW CLUSTER CODE IS  24
CYCLE  25 NOW FUSE POINTS  12  52 AT COEFFICIENT    0.053 -  45 CLUSTERS AND NEW CLUSTER CODE IS  12
CYCLE  26 NOW FUSE POINTS  18  21 AT COEFFICIENT    0.054 -  44 CLUSTERS AND NEW CLUSTER CODE IS  18
CYCLE  27 NOW FUSE POINTS   8  55 AT COEFFICIENT    0.056 -  43 CLUSTERS AND NEW CLUSTER CODE IS   8
CYCLE  28 NOW FUSE POINTS  30  46 AT COEFFICIENT    0.063 -  42 CLUSTERS AND NEW CLUSTER CODE IS  30
CYCLE  29 NOW FUSE POINTS  26  48 AT COEFFICIENT    0.072 -  41 CLUSTERS AND NEW CLUSTER CODE IS  26
CYCLE  30 NOW FUSE POINTS  29  60 AT COEFFICIENT    0.074 -  40 CLUSTERS AND NEW CLUSTER CODE IS  29
CYCLE  31 NOW FUSE POINTS  34  36 AT COEFFICIENT    0.075 -  39 CLUSTERS AND NEW CLUSTER CODE IS  34
CYCLE  32 NOW FUSE POINTS  62  64 AT COEFFICIENT    0.075 -  38 CLUSTERS AND NEW CLUSTER CODE IS  62
CYCLE  33 NOW FUSE POINTS  12  13 AT COEFFICIENT    0.079 -  37 CLUSTERS AND NEW CLUSTER CODE IS  12
CYCLE  34 NOW FUSE POINTS  66  70 AT COEFFICIENT    0.086 -  36 CLUSTERS AND NEW CLUSTER CODE IS  66
CYCLE  35 NOW FUSE POINTS  30  57 AT COEFFICIENT    0.098 -  35 CLUSTERS AND NEW CLUSTER CODE IS  30
CYCLE  36 NOW FUSE POINTS   6  56 AT COEFFICIENT    0.099 -  34 CLUSTERS AND NEW CLUSTER CODE IS   6
CYCLE  37 NOW FUSE POINTS  31  32 AT COEFFICIENT    0.108 -  33 CLUSTERS AND NEW CLUSTER CODE IS  31
CYCLE  38 NOW FUSE POINTS   4  25 AT COEFFICIENT    0.110 -  32 CLUSTERS AND NEW CLUSTER CODE IS   4
CYCLE  39 NOW FUSE POINTS  10  11 AT COEFFICIENT    0.114 -  31 CLUSTERS AND NEW CLUSTER CODE IS  10
CYCLE  40 NOW FUSE POINTS  18  22 AT COEFFICIENT    0.125 -  30 CLUSTERS AND NEW CLUSTER CODE IS  18
CYCLE  41 NOW FUSE POINTS  37  67 AT COEFFICIENT    0.138 -  29 CLUSTERS AND NEW CLUSTER CODE IS  37
CYCLE  42 NOW FUSE POINTS   4  61 AT COEFFICIENT    0.140 -  28 CLUSTERS AND NEW CLUSTER CODE IS   4
CYCLE  43 NOW FUSE POINTS   8  10 AT COEFFICIENT    0.147 -  27 CLUSTERS AND NEW CLUSTER CODE IS   8
CYCLE  44 NOW FUSE POINTS  14  15 AT COEFFICIENT    0.147 -  26 CLUSTERS AND NEW CLUSTER CODE IS  14
CYCLE  45 NOW FUSE POINTS  28  59 AT COEFFICIENT    0.180 -  25 CLUSTERS AND NEW CLUSTER CODE IS  28
CYCLE  46 NOW FUSE POINTS   3  31 AT COEFFICIENT    0.180 -  24 CLUSTERS AND NEW CLUSTER CODE IS   3
```

```
CYCLE 47 NOW FUSE POINTS   9  17 AT COEFFICIENT   0,214 =  23 CLUSTERS AND NEW CLUSTER CODE IS   9
CYCLE 48 NOW FUSE POINTS   6  24 AT COEFFICIENT   0,329 =  22 CLUSTERS AND NEW CLUSTER CODE IS   6
CYCLE 49 NOW FUSE POINTS   9  37 AT COEFFICIENT   0,364 =  21 CLUSTERS AND NEW CLUSTER CODE IS   9
CYCLE 50 NOW FUSE POINTS   4  66 AT COEFFICIENT   0,407 =  20 CLUSTERS AND NEW CLUSTER CODE IS   4
CYCLE 51 NOW FUSE POINTS  26  38 AT COEFFICIENT   0,412 =  19 CLUSTERS AND NEW CLUSTER CODE IS  26
CYCLE 52 NOW FUSE POINTS   3  30 AT COEFFICIENT   0,483 =  18 CLUSTERS AND NEW CLUSTER CODE IS   3
CYCLE 53 NOW FUSE POINTS   6  23 AT COEFFICIENT   0,559 =  17 CLUSTERS AND NEW CLUSTER CODE IS   6
54 NOW FUSE POINTS   9  16 AT COEFFICIENT   0,602 -  16 CLUSTERS AND NEW CLUSTER CODE IS   9
55 NOW FUSE POINTS  34  42 AT COEFFICIENT   0,696 -  15 CLUSTERS AND NEW CLUSTER CODE IS  34
56 NOW FUSE POINTS  62  63 AT COEFFICIENT   0,738 -  14 CLUSTERS AND NEW CLUSTER CODE IS  62
57 NOW FUSE POINTS  28  29 AT COEFFICIENT   0,758 -  13 CLUSTERS AND NEW CLUSTER CODE IS  28
58 NOW FUSE POINTS   8  12 AT COEFFICIENT   0,785 -  12 CLUSTERS AND NEW CLUSTER CODE IS   8
59 NOW FUSE POINTS   4  26 AT COEFFICIENT   1,500 -  11 CLUSTERS AND NEW CLUSTER CODE IS   4
METHOD GROUP  60 FUSE POINTS  28  62 AT COEF   1,553  10 CLUSTERS
  1   3   4   1   6   1   8   9   8   8   8   8  14  14   9   9  18  18  18
 18   6   6   4   4   6  28  28   3   3   3   3  34  34  34   9   4   4   4
 34  34  34  34   3   3   4   9   1   8   8   4   3   8   6   3   4  28  28
 28  28  28   8   4   9   9   3   4
METHOD GROUP  61 FUSE POINTS   1   6 AT COEF   1,729   9 CLUSTERS
  1   3   4   1   1   1   8   9   8   8   8   8  14  14   9   9  18  18  18
 18   1   1   4   4   1  28  28   3   3   3   3  34  34  34   9   4   4   4
 34  34  34  34   3   3   4   9   1   8   8   4   3   8   1   3   4  28  28
 28  28  28   8   4   9   9   3   4
METHOD GROUP  62 FUSE POINTS   3  34 AT COEF   2,184   8 CLUSTERS
  1   3   4   1   1   1   8   9   8   8   8   8  14  14   9   9  18  18  18
 18   1   1   4   4   1  28  28   3   3   3   3   3   3   3   9   4   4   4
  3   3   3   3   3   3   4   9   1   8   8   4   3   8   1   3   4  28  28
 28  28  28   8   4   9   9   3   4
METHOD GROUP  63 FUSE POINTS  14  28 AT COEF   2,435   7 CLUSTERS
  1   3   4   1   1   1   8   9   8   8   8   8  14  14   9   9  18  18  18
 18   1   1   4   4   1  14  14   3   3   3   3   3   3   3   9   4   4   4
  3   3   3   3   3   3   4   9   1   8   8   4   3   8   1   3   4  14  14
 14  14  14   8   4   9   9   3   4
METHOD GROUP  64 FUSE POINTS   4   9 AT COEF   3,944   6 CLUSTERS
  1   3   4   1   1   1   8   4   8   8   8   8  14  14   4   4  18  18  18
 18   1   1   4   4   1  14  14   3   3   3   3   3   3   3   4   4   4   4
  3   3   3   3   3   3   4   4   1   8   8   4   3   8   1   3   4  14  14
 14  14  14   8   4   4   4   3   4
METHOD GROUP  65 FUSE POINTS   8  14 AT COEF   5,403   5 CLUSTERS
  1   3   4   1   1   1   8   4   8   8   8   8   8   8   4   4  18  18  18
 18   1   1   4   4   1   8   8   3   3   3   3   3   3   3   4   4   4   4
  3   3   3   3   3   3   4   4   1   8   8   4   3   8   1   3   4   8   8
  8   8   8   8   4   4   4   3   4
66 NOW FUSE POINTS   3   4 AT COEFFICIENT  10,947 -   4 CLUSTERS AND NEW CLUSTER CODE IS   3
67 NOW FUSE POINTS   8  18 AT COEFFICIENT  18,063 -   3 CLUSTERS AND NEW CLUSTER CODE IS   8
68 NOW FUSE POINTS   1   3 AT COEFFICIENT  18,212 -   2 CLUSTERS AND NEW CLUSTER CODE IS   1
69 NOW FUSE POINTS   1   8 AT COEFFICIENT  65,038 -   1 CLUSTERS AND NEW CLUSTER CODE IS   1
```

DENDROGRAM TABLE

1
2
5
7
50
6
56
24
27
23
3
31
33
32
30
54
69
47
46
57
34
35
36
42
43
45
44
4
53
25
61
66
70
26
48
58
38
41
39
40
9
17
37
49
67
16
68
8
51
55
10
11

FUSION SUMMARY

39	40	0.000
38	41	0.000
38	39	0.000
43	45	0.001
1	2	0.001
34	35	0.002
54	69	0.003
30	54	0.004
31	33	0.006
5	7	0.007
18	19	0.011
5	50	0.012
42	43	0.013
52	65	0.016
8	51	0.016
16	68	0.016
48	58	0.020
18	20	0.021
1	5	0.029
4	53	0.032
42	44	0.033
30	47	0.036
37	49	0.042
24	27	0.051
12	52	0.053
18	21	0.054
8	55	0.056
30	46	0.063
25	48	0.072
29	60	0.074
34	36	0.075
62	64	0.075
12	13	0.079
66	70	0.086
30	57	0.098
6	56	0.099
31	32	0.108
4	25	0.110
10	11	0.114
18	22	0.125
37	67	0.138
4	61	0.140
8	10	0.147
14	15	0.147
28	59	0.180
3	31	0.180
9	17	0.214
6	24	0.329
9	37	0.364
4	66	0.407
26	38	0.412
3	30	0.483

12	6	23	0.559
52	9	16	0.602
65	34	42	0.696
13	62	63	0.738
14	28	29	0.758
15	8	12	0.785
28	4	26	1.500
59	28	62	1.553
29	1	6	1.729
60	3	34	2.180
62	14	28	2.435
64	4	9	3.944
63	8	14	5.403
18	3	4	10.947
19	8	18	18.063
20	1	3	18.212
21	1	8	65.038
22			

JOB ENDS

PROCEDURE RESULT

RUN WITH ZELENY DECK

NUMBER OF CASES = 70
NUMBER OF BINARY VARIABLES = 0
NUMBER OF NUMERIC VARIABLES = 5

STANDARD SCORES FILED
5 COMPONENT SCORES FILED
NUMERIC COEFFICIENT CALCULATED
NO MASK USED
COEFFICIENT NUMBER 1 CALCULATED
5 K-LINKAGE LISTS FILED

NUMERIC MEANS AND STANDARD DEVIATIONS

1	33.0879	43.3465
2	59.4616	57.9650
3	92.5684	31.9073
4	-9.4286	30.9667
5	23.5361	7.9654

CORRELATIONS NOT FILED

5 K-LINKAGE LISTS - (NEAREST NEIGHBOURS)

S 1	0.001	2	0.005	7	0.008	5	0.024	50	0.186	24
S 2	0.001	1	0.009	7	0.014	5	0.033	50	0.207	24
S 3	0.093	53	0.126	32	0.133	31	0.135	57	0.135	46
S 4	0.032	53	0.053	25	0.064	61	0.117	66	0.178	58
S 5	0.007	7	0.008	1	0.014	2	0.014	50	0.116	24
S 6	0.099	56	0.140	23	0.204	24	0.219	27	0.445	50
S 7	0.005	1	0.007	5	0.007	50	0.009	2	0.153	24
S 8	0.016	51	0.031	55	0.085	52	0.085	11	0.086	10
S 9	0.092	55	0.160	10	0.178	8	0.214	17	0.226	51
S 10	0.062	51	0.086	8	0.114	11	0.124	55	0.160	9
S 11	0.047	51	0.061	52	0.073	65	0.085	8	0.114	10
S 12	0.021	65	0.066	52	0.078	13	0.145	11	0.211	29
S 13	0.046	65	0.068	52	0.078	12	0.160	11	0.217	8
S 14	0.147	15	0.212	12	0.282	13	0.310	65	0.397	52
S 15	0.147	14	0.501	63	0.619	12	0.739	13	0.752	65
S 16	0.016	68	0.167	70	0.169	37	0.228	49	0.238	17
S 17	0.071	37	0.175	49	0.176	68	0.214	9	0.238	16
S 18	0.011	19	0.016	20	0.058	21	0.061	22	0.785	63
S 19	0.011	18	0.021	21	0.022	20	0.059	22	0.637	63
S 20	0.016	18	0.022	19	0.045	21	0.109	22	0.727	63
S 21	0.021	19	0.045	20	0.058	18	0.126	22	0.511	63
S 22	0.059	19	0.061	18	0.109	20	0.126	21	0.624	63
S 23	0.140	6	0.401	28	0.465	56	0.492	24	0.539	27
S 24	0.051	27	0.113	50	0.116	5	0.153	7	0.178	56
S 25	0.053	4	0.089	58	0.118	61	0.128	53	0.147	48

S 26	0.033	58	0.086	48	0.148	25	0.169	67	0.213	51
S 27	0.051	24	0.205	56	0.219	6	0.245	5	0.281	50
S 28	0.174	60	0.180	59	0.276	29	0.401	23	0.492	12
S 29	0.074	60	0.130	52	0.137	11	0.169	65	0.175	51
S 30	0.003	54	0.005	69	0.014	47	0.039	46	0.053	57
S 31	0.006	33	0.063	46	0.072	69	0.074	32	0.081	42
S 32	0.074	31	0.091	33	0.126	3	0.208	69	0.219	54
S 33	0.006	31	0.051	69	0.057	46	0.064	30	0.069	54
S 34	0.002	35	0.052	36	0.090	44	0.158	45	0.171	39
S 35	0.002	34	0.061	36	0.082	44	0.144	45	0.156	39
S 36	0.052	34	0.061	35	0.111	38	0.111	49	0.116	39
S 37	0.042	49	0.071	17	0.168	68	0.169	16	0.175	67
S 38	0.000	41	0.000	39	0.000	40	0.046	48	0.106	58
S 39	0.000	40	0.000	41	0.000	38	0.050	48	0.111	58
S 40	0.000	39	0.000	41	0.000	38	0.049	48	0.111	58
S 41	0.000	38	0.000	39	0.000	40	0.044	48	0.104	58
S 42	0.008	43	0.013	45	0.043	44	0.084	31	0.089	33
S 43	0.001	45	0.008	42	0.019	44	0.114	31	0.133	33
S 44	0.012	45	0.019	43	0.043	42	0.082	35	0.090	34
S 45	0.001	43	0.012	44	0.013	42	0.120	31	0.143	33
S 46	0.017	69	0.031	54	0.039	30	0.057	33	0.063	31
S 47	0.014	30	0.027	54	0.035	69	0.077	57	0.093	46
S 48	0.020	58	0.044	41	0.046	38	0.049	40	0.050	39
S 49	0.042	37	0.052	67	0.111	36	0.175	17	0.228	16
S 50	0.007	7	0.014	5	0.024	1	0.033	2	0.113	24
S 51	0.016	8	0.047	11	0.060	55	0.062	10	0.103	52
S 52	0.016	65	0.061	11	0.066	12	0.068	13	0.085	8
S 53	0.032	4	0.084	66	0.093	3	0.128	25	0.170	61
S 54	0.003	69	0.003	30	0.027	47	0.031	45	0.040	57
S 55	0.031	8	0.060	51	0.092	9	0.124	10	0.179	67
S 56	0.099	6	0.178	24	0.205	27	0.248	5	0.272	50
S 57	0.049	54	0.053	30	0.062	69	0.077	47	0.117	46
S 58	0.020	48	0.033	26	0.089	25	0.104	41	0.106	38
S 59	0.180	28	0.414	62	0.449	63	0.545	64	0.578	23
S 60	0.074	29	0.174	28	0.233	25	0.261	11	0.276	61
S 61	0.064	4	0.118	25	0.142	70	0.170	53	0.174	66
S 62	0.075	64	0.395	63	0.414	59	0.510	29	0.512	28
S 63	0.395	62	0.449	59	0.501	15	0.511	21	0.624	22
S 64	0.075	62	0.493	6	0.512	28	0.538	61	0.545	59
S 65	0.016	52	0.021	12	0.046	13	0.073	11	0.161	8
S 66	0.084	53	0.086	70	0.117	4	0.147	3	0.174	61
S 67	0.052	49	0.110	58	0.120	48	0.138	36	0.169	26
S 68	0.016	16	0.168	37	0.176	17	0.267	49	0.275	70
S 69	0.003	54	0.005	30	0.017	46	0.035	47	0.051	33
S 70	0.086	66	0.142	61	0.167	16	0.194	4	0.233	53

RUN WITH ZELENY DECK
WARDS METHOD GROUP 60 FUSE POINTS 28 62 AT COEF 1,553 10 CLUSTERS

CLASSIFICATION ARRAY

```
 1  1  2  3  1  4  1  5  6  5  5  5  5  7  7  6  6  8  8  8  8  8  4  4  3  3  4  9  9  2
 2  2  2 10 10 10  6  3  3  3  3 10 10 10 10  2  2  3  6  1  5  5  3  2  5  4  2  3  9  9
 3  9  9  9  5  3  6  6  2  3
```

CLUSTER 1 NUMBER OF CASES = 5

CASE NUMBERS
1 2 5 7 50

CLUSTER DIAGNOSIS OF MEANS, STANDARD DEVIATIONS AND F-RATIO

VAR	F-RATIO	T	MN-ORIG	STD-ORIG	VAR	F-RATIO	T	MN-ORIG	STD-ORIG
2	0,0018	-1,4855	-26,6480	2,4418	1	0,0030	1,9003	115,4580	2,3726
3	0,0037	-0,1179	88,8080	1,9313	4	0,0039	0,1492	-4,8080	1,9313
5	0,0120	0,4070	26,7780	0,8736					

CLUSTER 2 NUMBER OF CASES = 10

CASE NUMBERS
3 30 31 32 33 46 47 54 57 69

CLUSTER DIAGNOSIS OF MEANS, STANDARD DEVIATIONS AND F-RATIO

VAR	F-RATIO	T	MN-ORIG	STD-ORIG	VAR	F-RATIO	T	MN-ORIG	STD-ORIG
2	0,0267	-1,0992	-4,2550	9,4749	3	0,0345	-0,5875	73,8240	5,9283
4	0,0367	0,6329	10,1690	5,9316	1	0,0606	1,0382	78,0920	10,6729
5	0,0869	-1,1184	14,6280	2,3478					

CLUSTER 3 NUMBER OF CASES = 13

CASE NUMBERS
4 25 26 38 39 40 41 48 53 58 61 66 70

CLUSTER DIAGNOSIS OF MEANS, STANDARD DEVIATIONS AND F-RATIO

VAR	F-RATIO	T	MN-ORIG	STD-ORIG	VAR	F-RATIO	T	MN-ORIG	STD-ORIG
2	0.0424	-0.1923	48.3146	11.9394	1	0.0667	-0.0930	29.0554	11.1984
3	0.1249	-0.4762	77.3746	11.2753	4	0.1327	0.5182	6.6185	11.2819
5	0.1678	-0.1656	22.2169	3.2629					

CLUSTER 4 NUMBER OF CASES = 5

CASE NUMBERS
6 23 24 27 56

CLUSTER DIAGNOSIS OF MEANS, STANDARD DEVIATIONS AND F-RATIO

VAR	F-RATIO	T	MN-ORIG	STD-ORIG	VAR	F-RATIO	T	MN-ORIG	STD-ORIG
1	0.0177	1.2142	85.7200	5.7616	2	0.0559	-0.7043	18.6380	13.7034
5	0.1374	0.8047	29.9460	2.9521	3	0.1491	0.3694	104.3540	12.3213
4	0.1583	-0.3528	-20.3540	12.3213					

CLUSTER 5 NUMBER OF CASES = 9

CASE NUMBERS
8 10 11 12 13 51 52 55 65

CLUSTER DIAGNOSIS OF MEANS, STANDARD DEVIATIONS AND F-RATIO

VAR	F-RATIO	T	MN-ORIG	STD-ORIG	VAR	F-RATIO	T	MN-ORIG	STD-ORIG
1	0.0300	-1.2423	-20.7633	7.5118	2	0.0424	0.9244	113.0433	11.9404
3	0.0753	-0.0092	92.2744	8.7569	4	0.0800	0.0371	-8.2800	8.7601
5	0.1236	0.7657	29.6356	2.8002					

CLUSTER 6 NUMBER OF CASES = 7

CASE NUMBERS
9 16 17 37 49 67 68

CLUSTER DIAGNOSIS OF MEANS, STANDARD DEVIATIONS AND F-RATIO

VAR	F-RATIO	T	MN-ORIG	STD-ORIG	VAR	F-RATIO	T	MN-ORIG	STD-ORIG
3	0.0496	-0.5948	73.5886	7.1029	4	0.0526	0.6406	10.4086	7.1026
2	0.0578	0.5287	90.1071	13.9365	1	0.1041	-1.1444	-16.5186	13.9830
5	0.2273	-0.8815	16.5143	3.7977					

CLUSTER 7 NUMBER OF CASES = 2

CASE NUMBERS
14 15

CLUSTER DIAGNOSIS OF MEANS, STANDARD DEVIATIONS AND F-RATIO

VAR	F-RATIO	T	MN-ORIG	STD-ORIG	VAR	F-RATIO	T	MN-ORIG	STD-ORIG
1	0.0003	-1.5884	-35.7650	0.7650	2	0.0142	1.8397	166.1000	6.9000
5	0.0505	0.9935	31.4500	1.7900	3	0.0578	1.1835	130.3300	7.6700
4	0.0613	-1.1916	-46.3300	7.6700					