International Institute for
Applied Systems Analysis
IIASA www.iiasa.ac.at

An Interactive Modeling Support System (IMSS)

Nakamori, Y., Ryobu, M., Fukawa, H. and Sawaragi, Y.

IIASA Working Paper

WP-85-077

November 1985

NOT FOR QUOTATION
WITHOUT PERMISSION
OF THE AUTHOR

AN INTERACTIVE MODELING SUPPORT SYSTEM (IMSS)

Y. Nakamori
M. Ryobu
H. Fukawa
Y. Sawaragi

November 1985
WP-85-77

Working Papers are interim reports on work of the International Institute for Applied Systems Analysis and have received only limited review. Views or opinions expressed herein do not necessarily represent those of the Institute or of its National Member Organizations.

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
2361 Laxenburg, Austria

CONTENTS

PREFACE

As in natural sciences like physics the primary aim of a systems analytic study is to find a synthesis of formal and informal, mathematical and non-mathematical methods, procedures, approaches, etc., and to design a computer-based system as a qualitatively new tool for the analysis of concrete problems pertinent to a concrete real system under study.

This process of design should be able to incorporate different types of available knowledge and information about the real system. The non-quantifiable knowledge of people, who from their experience know many important properties of the real system, is often of a high value. Therefore, having efficient channels of communicating this type of knowledge into the process of design is very desirable.

This type of a communicating channel is one of the characteristic features of the modeling procedure described in this paper. It accepts two types of information about a real system: measurement data and also expert knowledge about the system's structure. The use of the highly interactive computer system based on this procedure is an iterative process in the course of which the subjective expert knowledge can be utilized to a great advantage.

This interactive system has been applied successfully by the IIASA *Regional Water Policies Project* (WAT) for the development of simplified models of complex groundwater-crop growth systems for their subsequent incorporation in the decision support system for the Southern Peel region in the Netherlands. This application will be described in a forthcoming paper and also in the final report of the WAT project.

Sergei Orlovski
Project Leader
Regional Water Policies Project

ABSTRACT

Key Words. Complex system, computer assistance, identification, man-machine interface, model simplification, structural modeling.

A computer-assisted mathematical modeling method that emphasizes the interaction between analysts and computers is presented. It combines algebraic and graph-theoretic approaches to extract a trade-off between human mental models and models based on the use of data collected from the system under study. The method is oriented to the modeling of the so-called "gray box" systems which often involve human behavioral aspects and also knowledge of the experts in relevant fields. By recursive dialogues with the computer, the modeler finds a system model which can be nonlinear with respect to descriptive variables. The structure of the computer program packages is also presented.

AN INTERACTIVE MODELING SUPPORT SYSTEM (IMSS)

Y. Nakamori[1], M. Ryobu[2], H. Fukawa[3] and Y. Sawaragi[2]

1. INTRODUCTION

Kalman (1983) has emphasized that "a model must explain real data; it must not be an artifact expressing the modeler's prejudices." He (1982) claims that "the principal modeling problems for the future are not statistical, but system-theoretical," and continues that "the immediate task is to begin developing prejudice-free modeling theory." His words are impressive and shocking for applied systems analysts. The state-of-the-art of the mathematical systems theory is, however, not so well developed that it can cope with complex large-scale systems which lie outside of the domain of validity of the physical laws. Such systems must involve human behavioral aspects and may not provide behavioral data sufficient both in quality and quantity. We reply to Kalman's addresses by quoting the assertion in Gaines (1984) that "the powerful methods of linear systems theory work not because they reflect reality but rather because we have built worlds of mechanical and electronical systems which are linear, and hence can be modeled, designed and controlled; outside technologically created reality linear systems theory has far less to offer in modeling the worlds of nature." On the other hand, it is also the fact that we often face situations where any classical, statistical procedures would not work adequately. To deal with such problems we must consider effective utilization of the existing theories and tools.

[1] On leave from the Department of Applied Mathematics, Faculty of Science, Konan University, Kobe, Japan.
[2] The Japan Institute of Systems Research.
[3] Department of Applied Mathematics and Physics, Kyoto University.

The process of mathematical modeling involves a certain number of stages and cycles. In a classical framework of system modeling three fundamental stages are considered: selection of the type of model, parameter estimation and the validation of the model. El-Sherief (1984) defines the system identification problem in his recent survey concerning multivariable system modeling as follows: "from a given set of input and output measurements (cause and effect), it is required to estimate a mathematical model within 'a specified class of models' which fits the measurements as closely as possible." He observes that "one important factor in identifying multivariable systems is the determination of the structural parameters; each type of model has its own structural parameters which must be known before an attempt is made to estimate the model parameters." This statement is related to the main point in Kalman (1980): "there is a one-to-one correspondence between the data and its canonical realization (model); with the same data each modeler must arrive at the same conclusion as any other modeler, except for a possibility inevitable but always irrelevant relabeling of the variables." But there certainly exist worlds that the (partial) realization theory would not be accepted yet, and under such uncertain circumstances we must make important decisions.

A great majority of tendency has been rapidly arising in modeling of badly posed systems in which emphasis lies on structure characterization instead of parameter estimation. In fact, Linstone et al. (1979) identify about 100 structural modeling techniques, and develop guidelines in the choice and proper use of 7 famous tools. They define a structural model as "any model which represents a complex system as a set of elements with relations – nearly always pairwise – linking some or all of them; and places the emphasis on the geometry or structure rather than on quantitative aspects of the relationships." Because generally decision-makers are not mathematicians or scientists a structural model is much more appropriate than others for learning experience. The structure of a system is fundamental to the understanding of what is happening. It gives new insights into the system to decision-makers and the modelers as well. Among many tools of structural modeling we extract the idea, for our purpose, from the Interpretive Structural Modeling (ISM) proposed by Warfield (1974); we have found in it the importance of a bird's eye view.

Our ultimate goal is to obtain some numerical relationships between system variables which should be, we believe, comprehensive or descriptive rather than intrinsic. Much attention has been also devoted to the extension of the classical image of modeling in uncertain environments. An unorthodox approach is known as the Group Method of Data Handling (GMDH) proposed by Ivakhnenko (1968). It is based on heuristic principles of self-organization and relies on bioengineering concepts. Despite the energetic research activities of Ibakhnenko and his colleagues after its introduction, the method seems to be far from world-acceptance. A critique is the following: by application of the self-organization method, the computer itself finds a unique model, but ignores any theories and consensuses developed in the relevant field. Look, this is a good example of the lesson for applied systems analysts. If we define the direction of new systems analysis as a discipline that provides decision support systems in any fields of human activities, then of vital importance is communication at every level, for instance dialogues between citizens and mathematicians, between economists and system analysts, between experts and decision-makers between people and computers, and between mental images and the data. We will borrow a part of Ibakhnenko's idea, but we do not necessarily rely on the whole process of heuristic self-organization.

The method presented in this paper consists of a combined modeling technique of algebraic and graph-theoretic approaches, and related man-machine interfaces. A new simulation model must be comprehensible, flexible and simple but appropriately complicated for the purpose of prediction or decision-making. But neither exactly defined stopping rule nor comparison criterion is imposed in our method. The pessimism of untouchability of the real structure does not allow to rely entirely on any traditional, statistical criteria because most of them have been invented to measure distance between the model and the real system. Some of them are, however, used in our method just as reference material, whenever required. Our mainpoint is how to balance, with computer assistance, the experts' mental models with those which the data tells.

In the next section we describe the outline of the method, and then in Section 3 we present the main part of this paper involving multistage dialogues in the modeling process and related man-machine interfaces. To make the paper self-contained, the details of graph-theoretic and algebraic phases of modeling will be presented in Sections 4 and 5, respectively, which are related to the principal computer program packages. Finally, in Section 6 we describe a personal computer software of the modeling support system.

2. STRUCTURE OF THE METHOD

PROBLEM. Suppose we have a real object to obtain a mathematical model. We introduce a set of names of variables:

$$S=\{x_i ; i=1,2,\ldots,n\} \quad ,$$

and suppose we have a measurement data table:

$$X=(x_{ij}) \ , \quad i=1,2,\ldots,n \ , j=1,2,\ldots,N,$$

where x_{ij} represents j-th measurement of i-th variable.

We introduce a cause-effect relation B, on the product set $S \times S$, defined by

$$(x_i, x_j) \in B \text{ if and only if } x_i \text{ influences } x_j \quad ,$$

or, equivalently, a matrix $A = (a_{ij})$ defined by

$$a_{ij} = \begin{cases} 1 & \text{if } (x_i, x_j) \in B \\ 0 & \text{otherwise} \end{cases} \quad .$$

This matrix describes characteristic set of the relation B and is called the *adjacency matrix*. The matrix A represents a type of our knowledge about the dependency relation between system variables, and is not determined clearly at the initial stage of the modeling process.

The problem is to obtain a mathematical model of the system in terms of a set of equations which governs the elements of S using the measurement data X and the information A.

OUTLINE. The modeling process consists of three different but interdependent stages of dialogues.

The first stage dialogue is required for preparation of the modeling, including input of measurement data and the initial version of cause-effect relation on the set of variables, transformation of variables, data screening, and refinement of the cause-effect relation.

The second stage dialogue is devoted to find a trade-off between the measurement data and the modeler's knowledge about dependencies between the variables. Based on the measurement data and the initial version of the cause-effect relation, using the option of regression method, the computer finds a linear model and the corresponding digraph model. The modeler modifies the new relation referring to these computer models and his knowledge. The process continues repeatedly until no change occurs or the modeler is satisfied with the modified relation.

The third stage dialogue is related to model simplification and elaboration. Model simplification is based on the use of equivalence relation, and model elaboration is an application of regression analysis including the hypothesis testing on estimated coefficients, and examinations of the explanatory and predictive powers of the model.

Figure 1 shows the structure of the modeling process.

3. INTERACTIVE MODELING METHOD

3.1. The First Stage Dialogue

The first stage of the modeling process consists of the following steps that are necessary for preparation of the modeling.

1. *Data Input*

We call the triplicate (S,X,A) the modeling knowledge which is fed into the computer at the first step. The manner of filling the adjacency matrix A should be negative. Here negative means that the modeler should enter the computer a part of his knowledge, putting 0's at the right places. The rest of entries in A will be filled with 1's by the computer. The underlying idea is that we should inquire into strength of relationship between every pair of variables except those which are definitely irrelevant.

In filling the adjacency matrix $A = (a_{ij})$, we allow to use an extension of binary relation:

$$a_{ij} = \begin{cases} 2 & \text{if } x_i \text{ certainly influences } x_j \\ 0 & \text{if } x_i \text{ never influences } x_j \\ 1 & \text{otherwise} \end{cases} .$$

There is no difference between 1 and 2 in digraph modeling. They are treated differently in choosing explanatory variables in linear modeling, i.e., the variables indicated by 2 are regarded as the core variables and those indicted by 1 the optional variables. We redefine the cause-effect

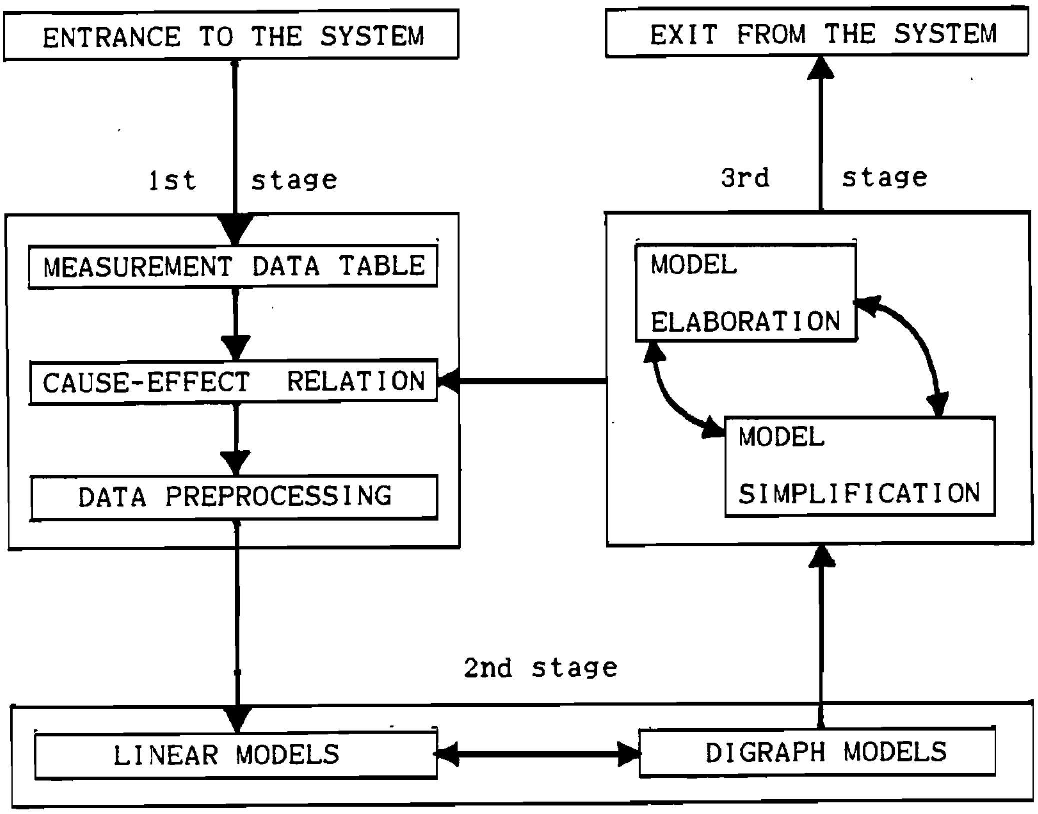

Figure 1: The structure of the modeling process

relation B as follows:

$$(x_i, x_j) \in B \text{ if and only if } a_{ij} \neq 0 \quad .$$

We have another option of filling the matrix A. The relation considered is the cause and effect that is not necessarily transitive. But it may be quite feasible to employ the assumption of transitivity to develop a linear model. The modeler can choose the option of a transitive embedding method which is a modified version of that in Warfield (1976). The resulting matrix A is a transitive matrix with elements 0 or 1. The advantage of this method is that it can reduce the number of pairwise comparisons remarkably. One caution in using this method is that the modeler should consider indirect cause-effect relationships as well as direct ones.

2. *Transformation of Variables*

The modeler can feed into the computer another type of knowledge: time-lag effects or functional relationships. The resulting model can be fairly complicated by transformation of variables. Transformation is also needed to make distributions of variables symmetric because, according to Hartwig and Dearing (1979), non-symmetric distributions and non-linear relationships often exist together. If every distribution of variables is roughly symmetric, then we will have a high chance to obtain a linear model. For this purpose the computer helps the modeler offering possible transformations. The options of these include:

$$y = exp(x)\ ,\quad y = log(x)\ ,\quad y = \log\,(x/(1-x))\quad ,$$

$$y = x^n\quad (n\text{: given integer})\ ,\quad y(t) = x(t-1)\quad (t\text{: time})\quad ,$$

$$y = a + bx_1 + cx_2 + dx_1x_2\quad (a, b, c \text{ and } d\text{: given constants})\quad ,$$

and their combinations. Needless to say, some transformations have constraints with respect to the range of numbers. The computer provides histograms of the original and transformed variables to assist the modeler's judgement. The modeler can choose a transformation by which the resulting new variable has a satisfactorily symmetric histogram after several examination.

When a transformation is done, the computer modifies the modeling knowledge (S, X, A) in the following way. If a single variable is transformed by a formula except the time-lag operation, then the corresponding row of the data matrix X is simply rewritten with the transformed numbers. Otherwise, a new variable is added to the set S, a new row is added to the data matrix X for the transformed data, and the adjacency matrix is extended in such a way that the new row (resp. column) is given by Boolean addition of the corresponding rows (resp. columns) of original variables.

3. *Cause-Effect Relation*

If the modeler wants to look at the structure of his mental model, then the computer will show the digraph of hierarchy based on the adjacency matrix A, taking its transitive closure and extracting the skeleton. The process of obtaining a digraph will be explained in detail in Section 4. Moreover, if the modeler wants to check the relationship between a pair of variables, then the computer will show two dimensional scatter plots. The modeler can change the relation characterized by A referring to these information.

If the objective of modeling is not just description but control or prediction of the real system, then the control variables or the variables whose data can be obtained accurately should be placed in appropriate positions in the hierarchy. The introduction of the modeler's assumptions or prejudices at this stage should be as little as possible.

4. *Linear Relation*

The computer checks and displays pairs of variables which have high correlation coefficients. To avoid the problem of multicollinearity and also to simplify the model, it is recommended that one of the pair is set aside when they are supposed to be linearly dependent. If the pair x_i and x_j is such a pair and the modeler wants to exclude x_j, then the j-th row and column of A will be rewritten as

$$a_{ji} = 3 \quad \text{and} \quad a_{jk} = 0\,, \; k \neq i\,, \; k \neq j$$
$$a_{ij} = 2 \quad \text{and} \quad a_{kj} = 0\,, \; k \neq i\,, \; k \neq j \quad ,$$

where 3 is treated the same as 1 or 2 in the digraph modeling, but it is treated as 0 in the linear modeling. Thus, x_j will be explained only by x_i and x_j will not be an explanatory variable for x_i.

The underlying idea is that if we put $a_{ji} = 2$, then there will be a high possibility that x_i is also explained only by x_j and this is not interesting. If there are more than two variables that are highly correlated, then the modeler can remove some of them in the same manner.

If there are some known relationships between variables in terms of linear equations, then the modeler can enter the facts into the computer. This information will save time in the second stage dialogue.

5. *Data Screening*

If at some step the modeler wants to check distributions or outliers of the data for some variables, the computer assists the modeler by showing the list of candidates of outliers, histograms or scattergrams. The modeler can designate the case numbers which he does not want to use in modeling.

After the first stage dialogue, the set of variables S and the data matrix X are fixed and will be used in the next stage as they are. The adjacency matrix A which contains the modeling knowledge obtained up to this stage is alone open for further modification. Figure 2 shows a visual description of the first stage dialogue.

3.2. The Second Stage Dialogue

The purpose of this stage is to elaborate the cause-effect relations which are summarized in the adjacency matrix. A series of reciprocal considerations and calculations by the modeler and the computer will continue until at least one of them recognizes that the further repetition would not improve the model. The information exchange process is the following:

1. *Selection of Regression Method*

The modeler should choose one of the options of regression methods with self-selection of explanatory variables, which will be used in the next step. The options of these include:

- the forward selection procedure,
- the backward elimination procedure,
- the all possible selection procedure, and
- the group method of data handling.

The last one is most recommended in our method because we have in mind the real world that can hardly provide the data with which the traditional, statistical inferences work well. If this method is selected, the computer asks the modeler about the data division into the training and testing sets. We use a modified or simplified version of this method, i.e., ***the partial descriptions*** will be written in a linear form (linear in variables). This point will be discussed in Section 5.1.

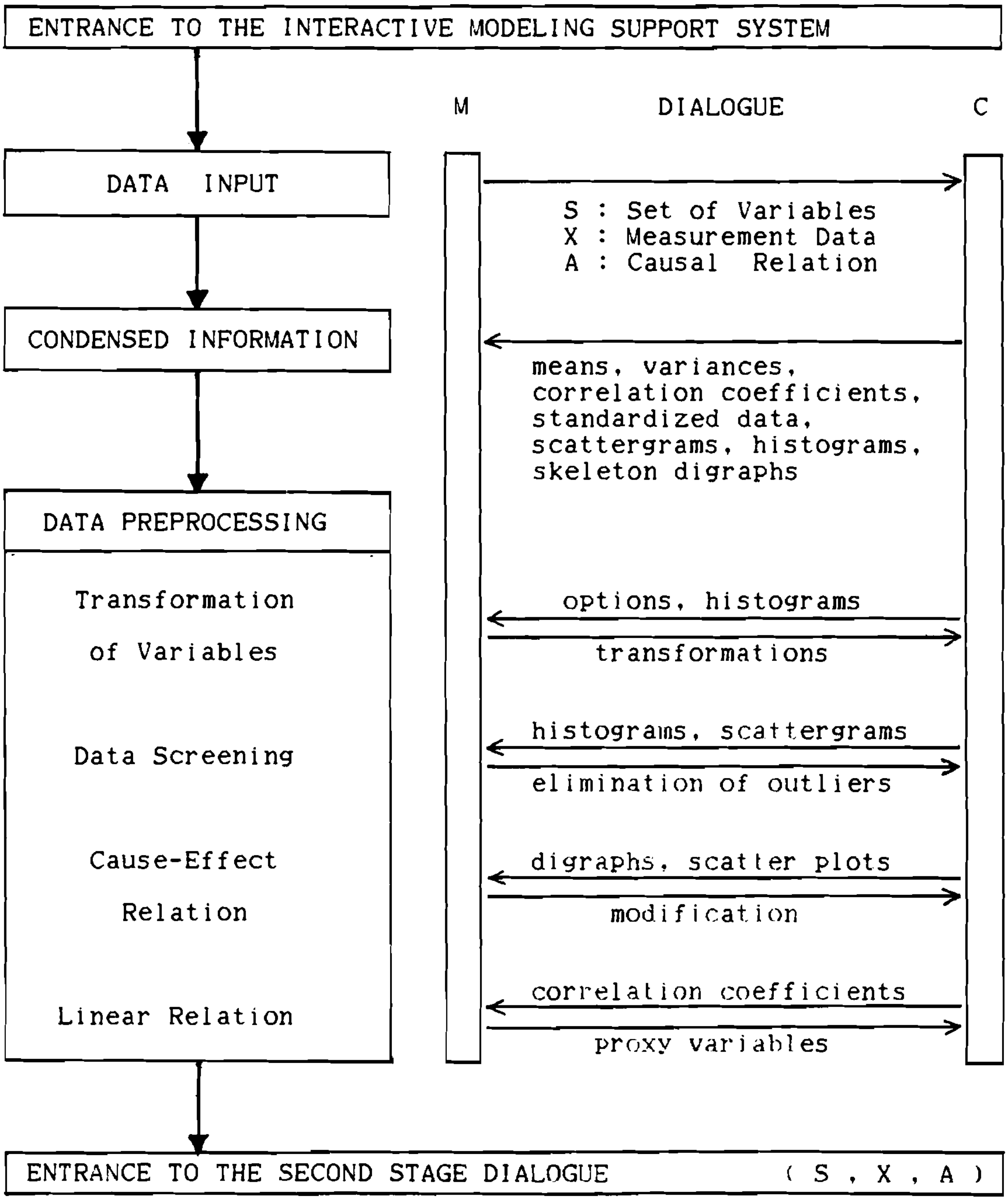

Figure 2: The first stage dialogue (M: modeler, C: computer)

2. *Estimation of Linear Relationships*

First the computer prepares an $n \times (n+1)$ vacant array $\bar{C}$:

$$\bar{C} = (c_{ij}) , \quad i = 1,2 , \ldots , n , \quad j = 0,1 , \ldots , n \quad .$$

For the convenience we write c_0 for the first column vector of $\bar{C}$ and C for the remaining $n \times n$ matrix, i.e.,

$$\bar{C} = (c_0,C) \quad .$$

The array $\bar{C}$ is prepared for the *coefficient* matrix of a set of linear equations which the computer will search from now on:

$$x = c_0 + Cx \quad \text{with} \quad c_{ii} = 0 , \ \text{all} \ i \quad ,$$

where x denotes the n-column vector whose components correspond to the names of variables $x_1,x_2 , \ldots , x_n$.

By the selected automatic modeling method, the computer will estimate the row vectors of $\bar{C}$ one by one referring to the matrix A (which can be converted to a transitive matrix before going into modeling) in the following way:

- Suppose now a turn is i-th row vector of $\bar{C}$.

- The computer refers the i-th column of the adjacency matrix A, and defines two subsets of S:

$$S_i^2 = \{x_j \ ; \ a_{ji} = 2 \ , \ j \neq i \}$$
$$S_i^1 = \{x_j \ ; \ a_{ji} = 1 \ , \ j \neq i \} \quad .$$

The variables x_j in S_i^2 are always chosen as ***explanatory variables*** for x_i, and those in S_i^1 are candidates of explanatory variables for x_i ***(the explained variable)***. Let us call S_i^2 the set of ***core variables*** and S_i^1 the set of ***optional variables***.

- By the specified method the computer will find the ***best fit*** equation (the meaning of the best fit will be explained in Section 5):

$$x_i = \hat{c}_{i0} + \sum_{x_j \in S_i^1 \cup S_i^2} \hat{c}_{ij} x_j$$

where $\hat{c}_{i0}$, $\hat{c}_{ij}$ are estimated coefficients.

- The computer substitutes:

$$c_{i0} = \hat{c}_{i0} \quad ,$$
$$c_{ij} = \begin{cases} \hat{c}_{ij} & \text{if} \ x_j \in S_i^1 \cup S_i^2 \\ 0 & \text{otherwise} \end{cases} \quad .$$

Thus the computer will have found a linear model:

$$M_C = (S,\bar{C}) \quad .$$

3. *Extraction of Skeleton*

A linear manifold indicates a relationship between variables, but it does not tell about the cause-effect relationships. But, here we impose on the computer a heuristic assumption: *In the linear model*

$$M_C = (S, \bar{C}) \quad \text{or} \quad x = c_0 + Cx$$

the explanatory variables are regarded as the causal variables to the explained variables.

Following this assumption the computer modifies the adjacency matrix $A = (a_{ij})$ referring to $C = (c_{ij})$ as follows: Let us denote the new i, j entry of A by a_{ij}^n and the old i, j entry by a_{ij}^o.

$$a_{ij}^n = \begin{cases} 0 & \text{if } c_{ji} = 0 \text{ and } x_i \in S_j^1 \\ a_{ij}^o & \text{otherwise} \end{cases} \quad .$$

Thus some of 1's in A will turn into 0's. The corresponding relation B is then defined by

$$(x_i, x_j) \in B \text{ if and only if } a_{ij} \neq 1 \quad .$$

Let us introduce a digraph D defined by

$$D = (S, B)$$

where the elements of S are identified as *vertices* and those of B *arcs*. The vertices are represented by points and there is a directed line heading from x_i to x_j if and only if (x_i, x_j) is in B. Let $\bar{B}$ denote the transitive closure of B, i.e.,

(a) $\bar{B}$ contains B,

(b) $\bar{B}$ is transitive, and

(c) $\bar{B}$ is the minimal relation satisfying (a) and (b).

Suppose the variable set S can be divided into m *equivalence classes* $E_1, E_2, \ldots, E_m$. Here an E_p is defined by

$$x_i \, , \, x_j \in E_p \text{ if and only if } (x_i, x_j) \, , \, (x_j, x_i) \in \bar{B} \quad .$$

By the graph-theoretic terminology an equivalence class is called a *strong component* or *a cycle set* of the digraph D. For details see Section 4.1. Then we can define new sets:

$$\bar{S} = \{E_p \, ; \, p = 1, 2, \ldots, m\} \quad ,$$

$$B^* = \{(E_p, E_q) \, ; \text{ some } (x_i, x_j) \in \bar{B} \, , \, x_i \in E_p \, , \, x_j \in E_q\}$$

and the corresponding digraph is called the *condensation digraph*:

$$D^* = (\bar{S}, B^*) \quad .$$

Finally we introduce the *skeleton digraph* $\bar{D}$ which is a minimum-arc *subdigraph* of D, for which removal of any arc would destroy reachability present in the relation. Actually the above process is carried out by some matrix operations in the computer. The details will be described in Section 4.2.

After all, the computer will have found the digraph model:

$$M_D = (\bar{S},\bar{D}) \quad .$$

This is a visual version of the linear model M_C. The digraph model is uniquely led from the linear model by the heuristic assumption, but the reverse is not true.

4. *Information Exchange*

Now the computer has a linear model M_C and the corresponding digraph model M_D. This step is devoted to the learning experience for both the modeler and the computer. Showing its digraph model M_D, the computer asks the modeler modification of the relation present in the linear model. The allowable amendments to the digraph model and their reflection on the adjacency and reachability matrices are summarized as follows; if an amendment affects the skeleton matrix, the digraph model is immediately modified.

(1) Format Amendments to Hierarchy

To facilitate interpretation of the relation, the modeler can amend the format of hierarchy that affects only the skeleton matrix. Such amendments include replacements of vertices, the contraction of vertices in different levels and the pooling of vertices in the same level. The vertices contracted or pooled are drawn in different colors to distinguish them from the strong components.

(2) Substantial Amendments to Cycles

The modeler can look at the adjacency structure of each cycle (strong component) and modify it by adding or removing arcs. Addition of an arc to a digraph map of a cycle has no effect on the reachability but corresponds to replacing a 0 in the adjacency matrix with a 1. On the other hand, removal of an arc causes the reverse operation on the adjacency matrix. When an arc is removed, the computer finds the transitive closure of the revised adjacency matrix and rewrite the reachability matrix. But an arc removal from a cycle sometimes preserves the universal reachability. If the cycle clipping is desired, a cycle can be divided into two strong components which can be either in the same level or in different levels. When a cycle is clipped by this manner, the corresponding interconnecting entries between divided strong components in the reachability matrix, and also those of the adjacency matrix filled with 1's, will be replaced by 0's. The modeler should pay careful attention to the cycles forming the ***vertex basis***. Here the vertex basis of a digraph is the set of vertices which consists of all vertices with no incoming arcs. The variables in the vertex basis should be measurable with relatively small measurement errors and should be appropriate as the control variables.

(3) Substantial Amendments to Hierarchy

Addition of a new arc to the hierarchy causes the same change in the adjacency and reachability matrices in such a way that all 0's between two strong components are replaced by 1's. But the latter matrix may not be reachable; hence the computer finds the transitive closure of the revised matrix. Removal of an arc from the hierarchy often affects the reachability. If an arc is removed, the adjacency matrix is first modified by replacing all 1's between two strong components with 0's. Then the computer finds the transitive closure of the revised adjacency matrix and see it thereafter as the transitive matrix. If an arc removal causes the violence of the total reachability necessary in the system, the modeler should compensate it by adding appropriate arcs.

Even the expert can hardly tell whether the obtained linear equations are appropriate or not because of the difficulties of checking validity of the hypothesis testing and giving meaning to regression coefficients. Therefore the linear equations are not shown here. But the direct modification of the adjacency structure between cycles is sometimes required. We prepare another program for this purpose.

(4) Amendments to Adjacency Structure

The computer exhibits the columns of the adajcency matrix A one by one which may present the linear relationship of variables. The modeler can change 0's to 1's in each column, and vice versa. Moreover he can write 2's at some entries if the indicated variables should always be necessary as the explanatory variables, i.e., the core variables. If any change is done, the reachability matrix is recalculated and the revised digraph is shown.

If the modeler does not change any relationships, then the modeling process will proceed to the third stage dialogue. Otherwise, the second stage dialogue will be repeated again. In this case the modeler can inform the computer the linear relationships with which he is already satisfied for saving time. He can substitute the reachability matrix for the adjacency matrix to find further possibilities in the linear modeling. Figure 3 sketches this stage of dialogue.

3.3. The Third Stage Dialogue

This stage consists of two modes:

- model elaboration, and
- model simplification.

The modeler can move from one mode to another at any time he wants.

1. *Model Elaboration*

If the modeler considers that he has enough data and that their statistics are meaningful, then he can elaborate the computer model by the classical regression analysis. Even if he has used the group method of data handling at the second stage, it is recommended in Ivakhnenko et al. (1979) that the coefficients of all the models upon comparison and selection can be reestimated using the minimum mean square error method applied to the whole data table. In this mode the modeler must designate an explained variable, then the computer will reestimate the coefficients of the linear equation and provide the following statistics:

- standard errors of estimated coefficients,
- t-ratios of estimated coefficients,
- standard deviation of residuals,
- F-ratio against a null hypothesis, and
- controlled determination coefficient.

All the above statistical terminologies will be explained in Section 5.2. Moreover, the computer supplies the routines:

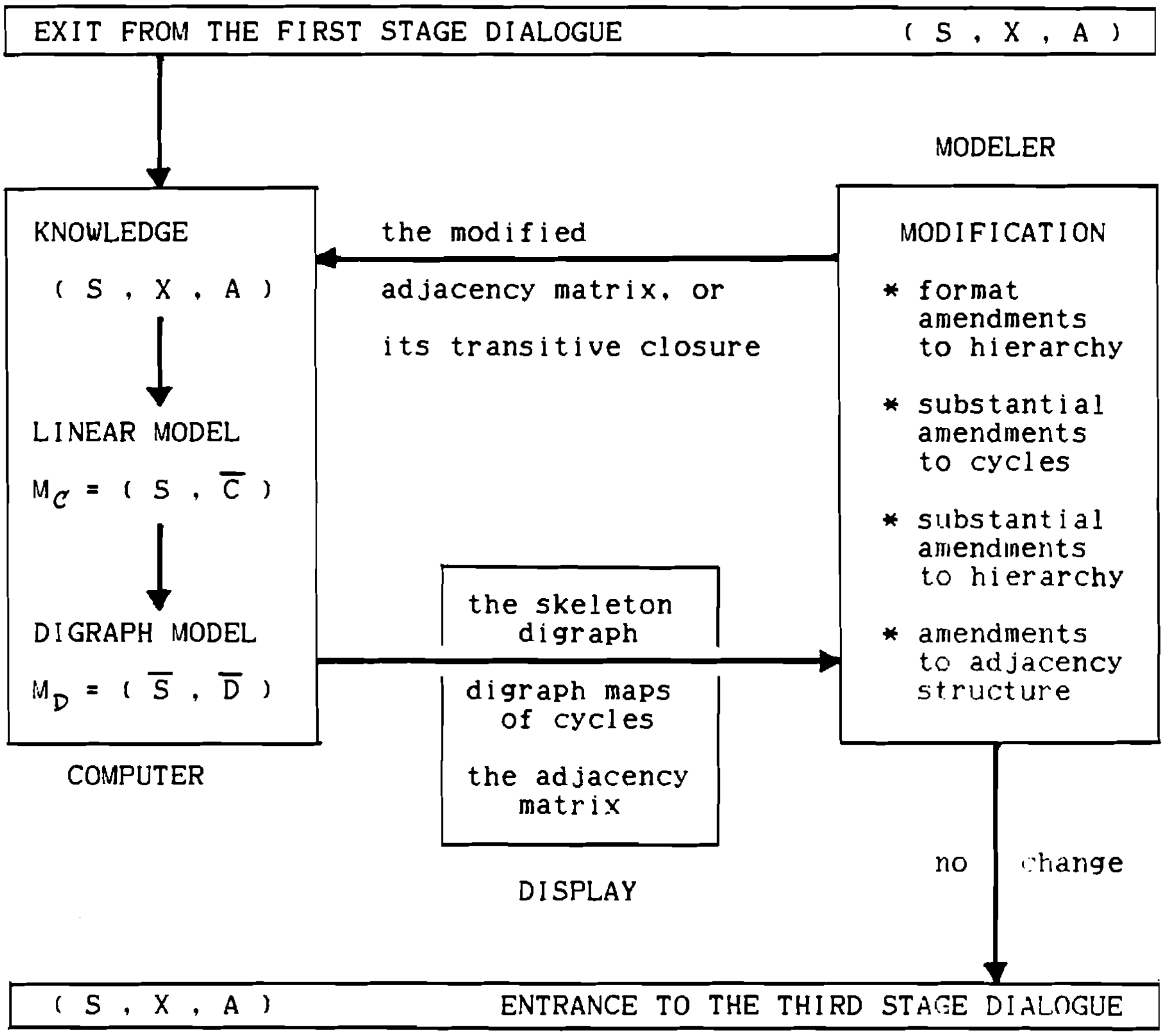

Figure 3: The second stage dialogue

- residual plots,
- multicolliearity checking, and
- prediction, if a new data set is available.

The modeler can elaborate the computer model by adding or removing some explanatory variables referring to these statistics. If the modeler wants the data preprocessing, he can call the subroutines in the first stage:

- transformation of variables, and
- data screening.

2. *Model Simplification*

Because the variables in an equivalence class may be connected by a linear relationship, it is desirable to choose proxy variables for model simplification and elaboration as well. The modeler can extract some proxy variables in each equivalence class to simplify the computer model in the following way.

(1) If two or more explanatory variables in a linear equation come from the same equivalence class, then the modeler can examine model simplification by choosing one or a few proxy variables and removing the rest. The computer will reestimate the coefficients of the equation and calculate some statistics mentioned in the model elaboration mode. The modeler can ask the computer to choose other variables as the proxy variables repeatedly, and if he is satisfied with one of the results, he will obtain a simplified model.

(2) If the explanatory variables in a linear equation come from many equivalence classes, then the modeler can examine further simplification so that the explanatory variables will come from a small number of equivalence classes, as long as the simplification does not destroy the reachability present in the model developed at the second stage.

Figure 4 shows the flow chart of the third stage dialogue. *The modeler can return to the first stage dialogue if he wants to reconstruct the model by using alternative tools equipped in the computer.*

4. GEOMETRIC PHASE OF MODELING

Having in mind that our final goal is to extract numerical properties of a complex system, we place the emphasis on the quantitative aspects of the relationships. One important thing involved in developing geometric models is the learning experience about the potential variables and their interactions. Lack of understanding of the structure of the underlying systems often leads us to the wrong conclusion.

Let us recall the notations: S denotes the set of descriptive system variables:

$$S = \{x_1, x_2, \ldots\ldots, x_n\}$$

and B a cause-effect relation of these variables:

$$B = \{(x_i, x_j) ; x_i, x_j \in S \text{ and } x_i \text{ affects } x_j\} \quad .$$

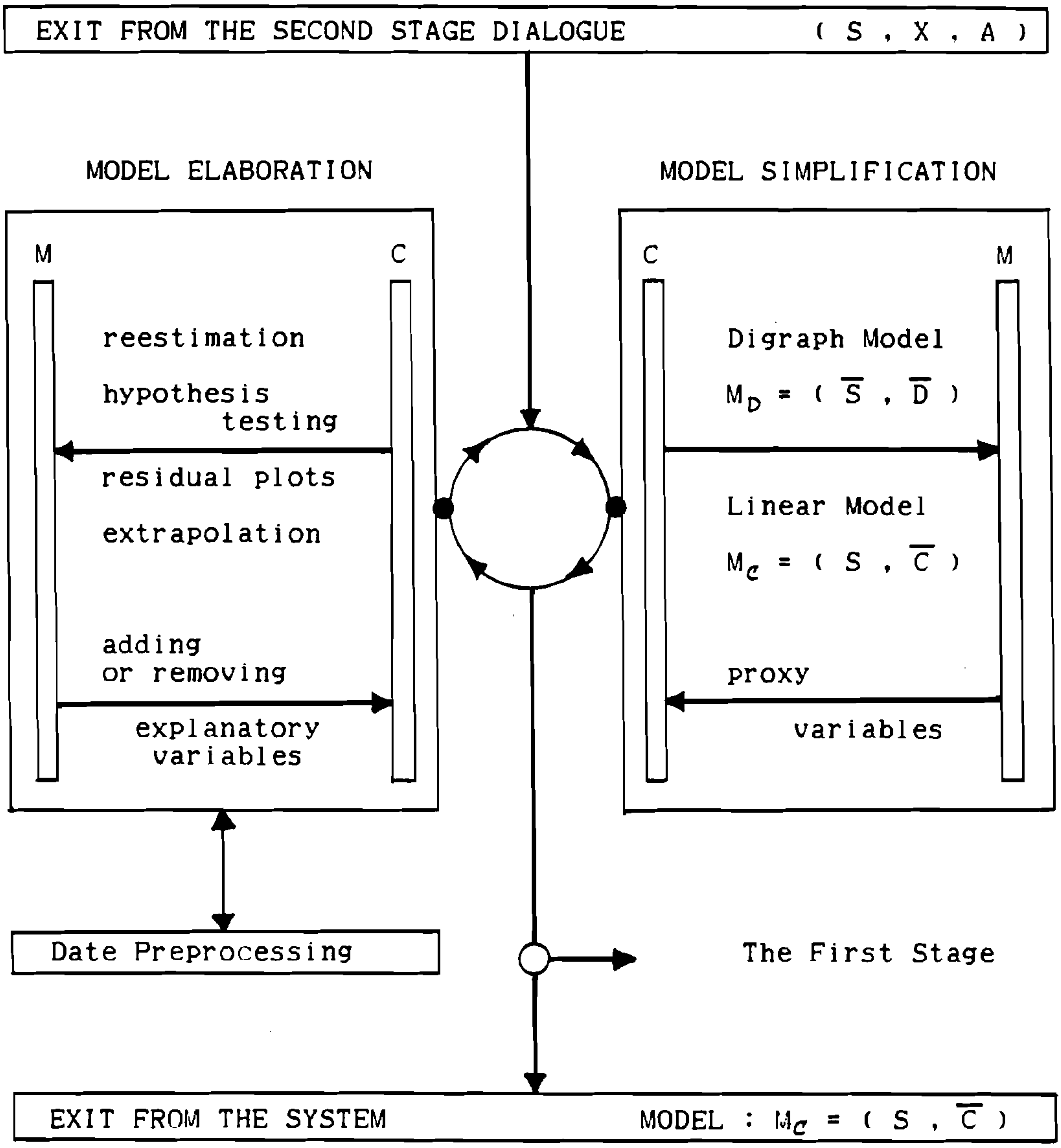

Figure 4: The third stage dialogue (M: modeler, C: computer)

4.1. Graph Theoretic Concepts

The foundation of structural modeling is provided by the graph theory which has been impressively developed by Harary et al. (1965) and Roberts (1976): the latter is the original text of the following description.

Define a directed graph or digraph D as a pair (S,B), where S is the set of elements and $B \in S \times S$ is a subset of ordered pairs of elements. We use the notation $S(D)$ and $B(D)$ for the vertex set and the arc set of D, respectively. The vertices are represented by points and there is a directed line leading from x_i to x_j if and only if (x_i,x_j) is in B. If there is an arc from vertex x_i to vertex x_j, we shall say that x_i is *adjacent* to x_j. We say x_j is *reachable* from x_i if there is a *path* from x_i to x_j. A *path* is a sequence

$$y_1 \,,\, (y_1,y_2) \,,\, y_2 \,,\, (y_2,y_3) \,,\, \ldots \,,\, (y_t,y_{t+1}) \,,\, y_{t+1}$$

where $t \geq 0$, $\{y_1,y_2, \ldots, y_{t+1}\}$ is a subset of S and each (y_i,y_{i+1}) is in B. A path is called *closed* if $y_{t+1} = y_1$. If the path is closed and the vertices $y_1,y_2, \ldots, y_t$ are distinct, then the path is called a *cycle*. An arc from a vertex to itself is called a *loop*.

A digraph D is *strongly connected* or *strong* if for every pair of vertices x_i and x_j, x_i is reachable from x_j and x_j is reachable from x_i. A *subdigraph* of D is a digraph whose vertex set is a subset of $S(D)$ and whose arc set is a subset of $B(D)$. A *strong component* or a *cycle set* of D is a maximal strongly connected subdigraph, where maximal means that if we add more vertices, the resulting generated subdigraph is not strongly connected. The vertices in a strong component form an equivalence class, i.e., they are connected with each other by a reflexive, transitive and symmetric relation. Note that a single vertex may constitute a strong component, and each vertex is in one and only one strong component.

We can now define a new digraph D^*, the *condensation digraph* of D as follows. Let $E_1,E_2, \ldots, E_m$ be the strong components or proxy vertices. Then

$$S(D^*) = \{E_1,E_2, \ldots, E_m\} \quad ,$$

and we draw an arc from E_p to E_q if and only if $p \neq q$ and for some vertices $x_i \in E_p$ and $x_j \in E_q$, there is an arc from x_i to x_j in D.

A collection V of vertices of a digraph D is called a *vertex basis* of D if every vertex not in V is reachable from some vertex in V and V is minimal. Here, minimal means that no proper subset of V can reach all vertices. The concept of the vertex basis is important from the control-theoretic view point and some theorems have been established:

(1) *The condensation digraph D^* of a digraph D is acyclic, i.e., it has no cycles.*

(2) *An acyclic digraph has a unique vertex basis, consisting of all vertices with no incoming arcs.*

(3) *Let V^* be the unique vertex basis of D^*. Then the vertex basis of D are those sets V consisting of one vertex from each strong component of D which is in V^*.*

(4) *Every two vertex bases of a digraph have the same number of vertices.*

A *skeleton digraph* $\bar{D}$ is a minimum-arc subdigraph of D^*, in which each strong component or a cycle in D has been replaced with a proxy vertex, and from which removal of any arc would destroy reachability present in the relation. A skeleton digraph gives insight into the hierarchical structure of the underlying system.

4.2. Structural Modeling

System structures in terms of the graph-theoretic terminologies can be conveniently summarized using suitable binary matrices. The process of structural modeling is a series of steps of matrix operations; a brief description is presented below.

Suppose $D = (S,B)$ is a digraph. The adjacency matrix A associated with D is the matrix (a_{ij}) defined by

$$a_{ij} = \begin{cases} 1 & \text{if } (x_i, x_j) \in B \\ 0 & \text{otherwise} \end{cases} .$$

One of the important properties of the adjacency matrix is:

If D is a digraph with adjacency matrix $A = (a_{ij})$, then i,j entry of A^t gives the number of paths of length t in D which lead from x_i to x_j.

The *reachability matrix*, or the *transitive closure* $R = (r_{ij})$ of A is the matrix defined by

$$r_{ij} = \begin{cases} 1 & \text{if } x_j \text{ is reachable from } x_i \\ 0 & \text{otherwise} \end{cases} .$$

Note that each vertex is reachable from itself, since x_i alone is a path, so $r_{ii} = 1$, all i. The reachability matrix can be expressed in terms of the adjacency matrix:

$$R = I + A + A + \cdots + A^{n-1} = (I+A)^{n-1}$$

where all the operations are *Boolean*. It is obvious from the definition that the reachability matrix describes *reflexive, transitive* relation, i.e., a *partial ordering relation*.

Many authors have developed partitioning and tearing methods on the reachability matrix in order to construct an interpretive structural model. Efficient procedures are found in Warfield (1976) and Sage (1977). After several partitions and rearrangements of the reachability matrix, one can obtain a *standard* or *canonical* form which is a lower block triangular matrix. This matrix can be converted into a condensation matrix in which the rows and columns of all the same levels, i.e., the cycle sets or the strong components in D are deleted except one, that one being identified as the proxy element.

As far as the extraction of strong components is concerned, the following theorem is useful.

Suppose D is a digraph with the reachability matrix $R = (r_{ij})$. Then:

(1) *The strong component containing a vertex x_i is given by the entries of* 1 *in the i-th row (or column) of $R \times R^T$, where R^T is the transpose of R and the product is the elementwise product, i.e., $R \times R^T = (r_{ij} \times r_{ji})$.*

(2) *The number of vertices in the strong component containing x_i is the i-th diagonal entry of R^2.*

The *skeleton matrix* (s_{ij}) is a condensation matrix in which all diagonal entries are 0, and the entries of 1 are changed into 0 until any additional entry would destroy reachability present in the condensation matrix. An efficient algorithm to find the skeleton matrix is presented in Warfield (1976). The relation modeled is ***asymmetric***, i.e., an entry $s_{ij} = 1$ implies $s_{ji} = 0$, and no cycle is found in the structure. The structural model of such a ***transitive, asymmetric*** relation is called a ***hierarchy***.

These model exchange isomorphisms describe the process by which primitive (mental) models are ultimately transformed into clearly articulated interpretive structural models. One of the greatest advantages of this process is that it gives the modeler insight into the structure itself. As insight is gained, the modeler may want to correct earlier aspects of the model.

5. ALGEBRAIC PHASE OF MODELING

Our method requires the program packages for the procedures of self-selection of explanatory variables at the second stage. The classical regression analysis is also used at the third stage for model simplification and elaboration.

5.1. Self-Organization Method

If the modeler has enough data, the following self-selection procedures are recommended:

- the forward selection procedure,
- the backward elimination procedure, or
- the all possible selection procedure.

The selection criterion (goodness of fit) used in these procedures is usually the controlled determination coefficient . A drawback of these procedures is that they need a fairly long time for calculation when the number of candidates of explanatory variables is large. If the modeler does not have enough data, or he wants a quick search for a linear model at the second stage, then he can choose a linear version of

- the group method of data handling.

We give below a brief summary of this method.

As mentioned in the introduction we are against some aspects of this method: "For the discovery of laws it is not necessary for the human operator to specify the set of explanatory variables, the input and output variables, the control variables, and the disturbances, etc. All of this is done by the computer (Ivakhnenko et al. 1979)." It is a matter of common knowledge that even apparently irrelevant variables could be approximately embedded in a linear (or nonlinear) equation. The reason why we use a part of this method is that we are supposing the objective systems as those which could hardly provide adequate data with which mathematics or statistics would work well to develop fantastic models acceptable to every person.

Ivakhnenko's idea is the following:

- If the data are not too variable, the computer itself can find the best unique model for prediction or the best one exhibiting cause-effect relationships.

- By application of the self-organization method, the computer should be able to objectively discover the natural law that exists in the object under study.

A prototype of the group method of data handling can be described as follows:

- The model to be found is the ***complete description***, where the explained variable is a nonlinear function of all the explanatory variables and their time-delayed variables. This complete description is found by several layers of approximation.

- At the first layer of selection the complete description is substituted by some ***partial descriptions*** which are nonlinear functions of every possible combinations of pairs of the explanatory variables and their time-delayed variables. The values of the partial description coefficients (goodness of fit) can be found by the mean squares error method. Then some of the partial descriptions are chosen such that the errors of selected ones are less than a specified threshold value.

- At the second layer of selection, the selected partial descriptions at the first layer play the roles of explanatory variables. The estimation of coefficients and the choice of some partial descriptions (the number should be less than that of the first layer) are repeated again.

- The number of selection layers increases as long as the lower value of the criteria is decreasing. Thus the process is continuously repeated with the imposition of ever more rigid thresholds so that finally a unique model is selected. When the model complexity gradually increases, the selection criterion passes through a minimum, and thus obtains the model of optimal complexity.

- The above process is the mathematical counterpart of the process used by a gardener in selectively raising various species for the purpose of obtaining a hybrid type that has desired properties.

A variety of heuristic criteria and algorithms are proposed by Ivakhnenko and his followers. The modeler must specify a criterion, an algorithm, some types of partial descriptions, etc. They are summarized as follows (we omit the explanation of terminologies).

- The operator (they call the modeler just as an operator) must convey to the computer a criterion of model selection according to his purpose, for example,

- the regularity criterion,
- the minimum-of-bias criterion,
- the combined criteria, and
- the balance-of-variance criterion.

- The operator must reduce the amount of data used in

  - model development (training set),

where coefficients are estimated by the mean squares error method, and use the rest in

- model verification (testing set),

i.e., selection of the partial descriptions.

- The operator must specify the list of feasible reference functions, such as

  - polynomiales,
  - rational fractions,
  - harmonic series, etc.

- The operator must specify the simulation environment, that is, a list of possible explanatory variables and their time-delayed variables.

- The operator must determine an algorithm for model sifting, for example,

  - the multilayer threshold algorithm,
  - the combinatorial algorithm, or
  - the adaptive learning network algorithm.

According to Ibakhnenko et al. (1979), there already exist about 100 algorithms. This fact itself tells how heuristic this method is.

In our method we use (heuristically)

- the regularity criterion,
- the multilayer threshold algorithm

which we have already described as a prototype of the group method of data handling. We restrict the partial descriptions to linear equations (linear in variables). It should be noted that in this paper a linear model means that the unknown parameters in each equation are embedded linearly. Because the modeler can transform variables as mentioned in Section 3.1, he can construct nonlinear models (nonlinear in the original variables). The reason of our constraint on the partial description that they should be linear in variables is that if we permit nonlinear equations for the partial descriptions, by application of the self-organization method the computer will often find a nonlinear equation with very high degree as the best model which cannot be interpreted at all.

5.2. Classical Procedures

Suppose now x_i is chosen for an explained variable, then from the adjacency matrix $A = (a_{ij})$ we have

$$S_i = \{j ; a_{ji} = 1, \text{ or } a_{ji} = 2 \text{ , } j \neq i \}$$

which corresponds to the union of core and optimal sets of explanatory variables for x_i. Let us introduce an N-column vector:

$$y = (y_1, y_2, \ldots, y_N)^T \text{ , where } y_j = x_{ij} \text{ , } j = 1,2, \ldots, N \quad ,$$

and the relabeled data matrix corresponding to S_i:

$$Z = (x_{jk})^T \text{ , } j = 0,1, \ldots, p \text{ , } k = 1,2, \ldots, N \quad ,$$

with $x_{0k} = 1$, all k, where $p = |S_i|$, the number of elements in S_i. In the classical regression analysis the disturbances in data are usually taken into account only for the explained variable. We introduce the noise of the explained variable as an N-column vector:

$$u = (u_1, u_2, \ldots, u_N)^T \quad ,$$

with assumptions:

$$E(u) = 0 \text{ , Var } (u) = E(uu^T) = \sigma^2 I \text{ , } u_i \sim N(0, \sigma^2) \text{ , } \sigma^2\text{: unknown} \quad ,$$

where $E(\cdot)$ denotes the expectation and $N(\cdot,\cdot)$ the normal distribution. We write the coefficients to be estimated as a $(p+1)$-column vector:

$$\beta = (\beta_0, \beta_1, \ldots, \beta_p)^T \quad .$$

By applying the least squares method we search the best approximation of the unknown vector in the set of assumed linear equations:

$$y = Z\beta + u \quad .$$

The least square estimator b of β is given by

$$b = (Z^T Z)^{-1} Z^T y \quad ,$$

if $Z^T Z$ is nonsingular. The estimator of y and the residual are given by

$$\hat{y} = Zb \text{ , } e = y - \hat{y} = y - Zb \quad ,$$

respectively. The unbiased estimator of the variance σ^2 is given by

$$s^2 = \frac{e^T e}{N-p-1} = \frac{y^T y - b^T Z^T y}{N-p-1} \quad ,$$

where the number $N - p - 1$ is called the ***degrees of freedom.***

The estimator b is independent of s^2 and

$$b \sim N(\beta \text{ , } \sigma^2 (Z^T Z)^{-1}) \quad .$$

This means that b is unbiased, and it is well known that the least square estimator has the minimum variance in all unbiased estimators which are linear with respect to measurements. Given a new measurement vector:

$$z = (1 , z_1 , z_2 , \ldots , z_p) \quad ,$$

the prediction of the explained variable is given by

$$\hat{y}_z = zb$$

with variance:

$$\text{var } (\hat{y}_z) = (z^T (Z^T Z)^{-1} z + 1)\sigma^2 \quad .$$

The standard error s.e. $(\hat{y}_z)$ of $\hat{y}_z$ is the square root of var $(\hat{y}_z)$ where σ^2 is substituted by s^2. The *confidence limit* with significance level α for $\hat{y}_z$ is given by

$$\hat{y}_z \pm t(N-p-1, \alpha/2) s.e.(\hat{y}_z) \quad ,$$

where $t(p,q)$ is the $(1-q)$ percentile point of the t-distribution with degrees of freedom p.

The t-ratio or t-statistic is defined by

$$t = \frac{b_i - \beta_i^0}{s\sqrt{c_{ii}}} \text{ , where } c_{ij}: i,j \text{ entry } of \ (Z^T Z)^{-1} \quad .$$

If the null hypothesis $H_0(\beta_i = \beta_i^0)$ is true, then this statistic follows the student's t-distribution with degrees of freedom $N-p-1$. The statistic $s\sqrt{c_{ii}}$ is an unbiased estimator of $\sqrt{\text{var}(b_i)}$ and called the ***standard error*** of b_i. In case that $\beta_i^0 = 0$, the confidence limit with significance level α is given by

$$b_i \pm t(N-p-1, \alpha/2) s \sqrt{c_{ii}} \quad .$$

On the other hand the *F-ratio* is used for another type of hypothesis testing:

- all or some of the regression coefficients are zeros,
- two or more regression coefficients are identical.

The original model is called the full model (FM) and a model in which some coefficients are specified is called a reduced model (RM). Let $\hat{y}_i$, $\hat{y}_i'$ be the estimates by FM, RM, respectively. *Sums of squares due to error* are defined by

$$\text{SSE(FM)} = \sum_i (y_i - \hat{y}_i)^2 \quad ,$$

$$\text{SSE(RM)} = \sum_i (y_i - \hat{y}_i')^2 \quad ,$$

respectively. Assume that the RM contains k parameters to be estimated. Then F-ratio is defined by

$$F = \frac{[\text{SSE(RM)} - \text{SSE(FM)}]/(p+1-k)}{\text{SSE(FM)}/(N-p-1)} \quad ,$$

which follows the F-distribution with degrees of freedom $(p+1-k \ , N-p-1)$. If the value of F-ratio is less than a given percentile point of F-distribution with degrees of freedom $(p+1-k \ , N-p-1)$, then the null hypothesis will be rejected.

It should be noted that the hypothesis testing is meaningful only when the assumptions on the error term is valid. To check this, the modeler should look at the residual plots.

The coefficient of multiple correlation R is the sample correlation coefficient between y and $\hat{y}$, and often used for the goodness of fit of the regression equation. The square of R is called the ***determination coefficient*** given by (after a little manipulation)

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \text{ , where } \bar{y} = \frac{1}{N} \sum_i y_i \quad .$$

This statistic is the ratio of the generated variation to the total variation. In other words, R^2 identifies a good regression, in the sense of the estimated function contributing most to the total variation of the explained variable. If we add a further explanatory variable to an equation, R^2 increases in return for the decrease of degrees of freedom. The variation in residuals will be smaller, and the confidence interval will be wider. The most often used criterion for model selection is then the *controlled determination coefficient* defined by

$$\bar{R}^2 = 1 - \frac{[\sum (y_t - \hat{y}_t)^2]/(N-p-1)}{[\sum (y_t - \bar{y})^2]/(N-1)} = 1 - \frac{N-1}{N-p-1}(1-R^2) \quad .$$

This statistic is useful in prediction application, where we want a set of explanatory variables which minimizes residual variance.

6. COMPUTER PROGRAM PACKAGE

We have developed a program package for the modeling support system, which runs on a personal microcomputer. The system structure is described in Figure 5. The list of subprograms in the program package, with functions and hierarchical levels in the program structure, is shown in Table 1, and data files for modeling information to be prepared or generated are summarized in Table 2.

7. CONCLUDING REMARKS

The two well known methods to infer causal relationships from non-empirical data are Blalock's causal inference (Blalock, 1972) and the path analysis (for instance, Kenny, 1979). The former is used for verification of hypothetic causal models, and the latter for analysis of strengths of causal relationships in assumed models. Both fall under the category of correlation-regression analysis, and could not infer causal relationships completely. To express non-symmetric causal relationships, a set of linear equations which we described in Section 3.2 is often adopted. If the relationships are asymmetric and acyclic (the so-called recursive system), the treatment of such a system is relatively easy. Otherwise, some of the regression coefficients should be specified before solving the problem (see, for instance, Johnson, 1972). If the whole variables in the system can be successfully divided into output (endogenous) and input (exogenous) variables, one can use a model written in a set of linear simultaneous equations. The so-called simultaneous equation estimation has been employed in econometrics for quite some time now. This method also requires a priori model specification, and an error in model formulation can easily influence the validity of the total model.

Most of the theoretical approaches in modeling analysis in econometrics, ecology and sociology seem to enter too many mathematical constraints in return for removing human knowledge. The proposed method in this paper is not mathematics-oriented but application-oriented. The interactive modeling support system is a tool for enlightening both the computer and the modeler about the underlying complex system. The main point is how effectively extract reality from human mental models with computer assistance. Even Kalman (1983) states that "in the modeling context prejudice may sometimes be good and in fact most valuable, such as a brilliant

Table 1: The program package of the interactive modeling support system

code	stage	level	function
100	0	1	file control (open, create or erase)
200	0	1	the master menu for stage menus
300	1	2	the menu for the first stage dialogue
310	1	3	initialization of the modeling
320	1	3	the menu for measurement data input
321	1	4	filing the original measurement data
322	1	4	appending data of new variables
323	1	4	appending up-dated measurement data
324	1	4	correction of mistyped data
330	1	3	the menu for the relation input
331	1	4	filing the relation one-by-one
332	1	4	filing the relation by transitive embedding
333	1	4	modification of the cause-effect relation
334	1	4	calculation of the transitive closure
340	1	3	transformations of variables
350	1	3	digraph models of the initial relation
360	1	3	checking the relation by correlations
370	1	3	outlier checking or elimination
380	1	3	calculation of basic statistics
400	2	2	the menu for the second stage dialogue
410	2	3	the menu for the regression methods
411	2	4	the forward selection procedure
412	2	4	the backward elimination procedure
413	2	4	the all possible selection procedure
414	2	4	the group method of data handling
420	2	3	refinement of regression coefficients
430	2	3	digraph modeling
440	2	3	digraph models of the revised relation
450	2	3	amendments of the digraph model
500	3	2	the menu for the third stage dialogue
510	3	3	model simplification
520	3	3	the menu for model elaboration
521	3	4	hypothesis testing
522	3	4	residual plots
523	3	4	multicollinearity checking
524	3	4	estimation by the model
524	3	4	information of regression results
530	3	3	prediction based on new data
540	3	3	digraph modeling
600	1	3	the menu for modeling information
610	1	4	the initial version of the relation
620	1	4	the original measurement data
630	1	4	the standardized data
640	1	4	the averages and variances
650	1	4	the correlation coefficients
660	1	4	the menu for scatter diagrams
661	1	5	histograms and scattergrams
662	1	5	scatter plots between two variables
663	1	5	scatter plots between three variables
700	4	2	the current linear model

guess about the nature of the data." We admit that there is no unique way to complex-system modeling. But we believe that the proposed method certainly directs to the right way in this field. The development of the modeling support system is still in its first stage and some important issues are

Table 2: Data files for modeling information

code	contents
00	the list of systems in the disk
01	commonly used parameters
02	the list of names of variables
03	the list of outliers
04	the original data table
05	sample means of variables
06	sample variances of variables
07	the standardized data table
08	correlation coefficients
09	the initial adjacency matrix
10	the initial reachability matrix
11	the initial skeleton matrix
12	the revised adjacency matrix
13	the revised reachability matrix
14	the revised skeleton matrix
15	regression results (statistics)
16	the linear model (coefficients)
17	the data table for prediction

left for future study. They are, for example, the problems of non-pairwise relationships, non-binary relationships, intransitive relations, cumulative connections, dynamics and structural changes.

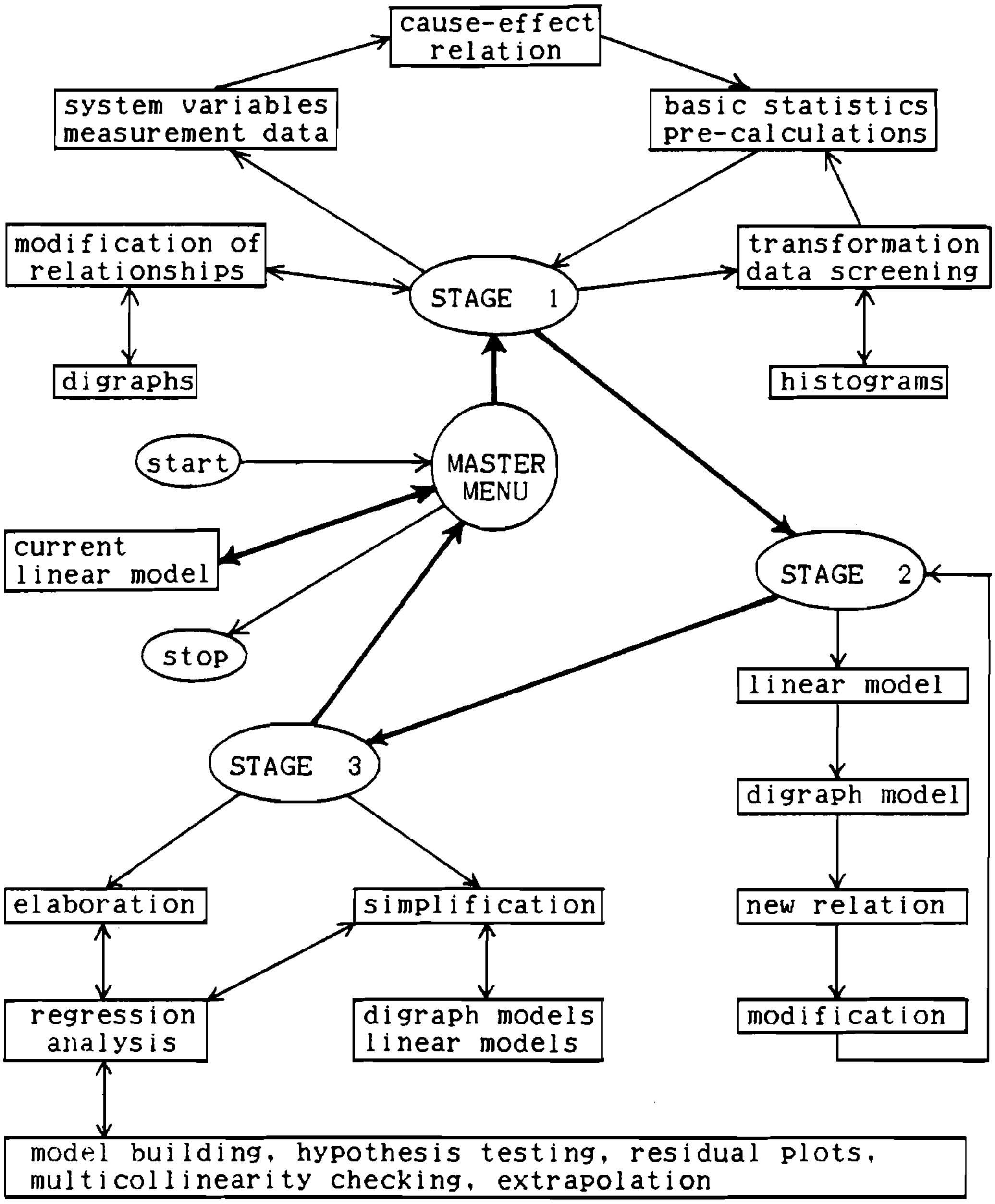
cause-effect relation
system variables measurement data
basic statistics pre-calculations
modification of relationships
STAGE 1
transformation data screening
digraphs
histograms
start
MASTER MENU
current linear model
STAGE 2
stop
linear model
STAGE 3
digraph model
elaboration
simplification
new relation
regression analysis
digraph models linear models
modification
model building, hypothesis testing, residual plots, multicollinearity checking, extrapolation

Figure 5: The interactive modeling support system

REFERENCES

Blalock, H.M., Jr. (1972). *Social Statistics*. McGraw-Hill, New York.

El-Sherief, H. (1984). Recent Advances in Multivariable System Modeling and Identification Algorithms and Their Applications. *Systems Research*, Vol. 1, No. 1, pp. 63-70.

Gaines, B.R. (1984). Methodology in the Large: Modeling All There Is. *Systems Research*, Vol. 1, No. 2, pp. 91-103.

Harary, F., Norman, R.Z. and Cartwright, D. (1965). *Structural Models: An Introduction to the Theory of Directed Graphs*. Wiley, New York.

Hartwig, F. and Dearing, B.E. (1979). *Exploratory Data Analysis*. Sage, London.

Ivakhnenko, A.G. (1968). The Group Method of Data Handling, a rival of the method of stochastic approximation. *Soviet Automatic Control*, Vol. 13, No. 3, pp. 43-55.

Ivakhnenko, A.G., Krotov, G.I. and Visotsky, V.N. (1979). Identification of the Mathematical Model of a Complex System by the Self-Organization Method. *Theoretical Systems Ecology*, E. Halfon (Ed.), Academic Press, New York, pp. 325-352.

Johnson, J. (1972). *Econometric Methods*. McGraw-Hill, New York.

Kalman, R.E. (1980). A System-Theoretic Critique of Dynamic Economic Models. *Int. J. of Policy Analysis and Information Systems*, Vol. 4, No. 1, pp. 3-22.

Kalman, R.E. (1982). Identification form Real Data. *Current Developments in the Interface: Economics, Econometrics, Mathematics*, M. Hazewinkel and A.H.G. Rinnooy Kan (Eds.), D. Reidel, pp. 161-196.

Kalman, R.E. (1983). Identifiability and Modeling in Econometrics. *Developments in Statistics*, Vol. 4. Academic Press, New York, pp. 97-136.

Kenny, D.A. (1979). *Correlation and Causality*. Wiley, New York.

Linstone, H.A., Lendaris, G.G., Rogers, S.D., Wakeland, W. and Williams, M. (1979). The Use of Structural Modeling for Technology Assessment. *Technological Forecasting and Social Change*, Vol. 14, pp. 291-327.

Roberts, F.S. (1976). *Discrete Mathematical Models with Application to Social, Biological, and Environment Problems*. Prentice-Hall, New Jersey.

Sage, A.P. (1977). *Methodology for Large-Scale Systems*. McGraw-Hill, New York.

Warfield, J.N. (1974). Toward Interpretation of Complex Structural Models. *IEEE Trans. Syst. Man Cybern.*, SMC-4, No. 5, pp. 405-417.

Warfield, J.N. (1976). *Societal Systems: Planning, Policy and Complexity*. Wiley, New York.