TOWARD ADVANCED COMPUTER-ASSISTED MODELING

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

April 1986
CP-86-17

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
2361 Laxenburg, Austria

Preface

A mathematically elaborated modeling method alone cannot develop useful models of large-scale systems that involve human activities. What is needed as input to the model-building process, besides measurement data, is the knowledge of experts in relevant fields. The problem is, then, what types of knowledge should or can be included in the modeling process and, more important, how do we manage them. The interactive method of data handling (IMDH) presented in this paper develops linear models of complex systems through recursive interaction with the computer, systematically introducing the expert's knowledge about the structure of the underlying system. It should be emphasized that the more one repeats dialogues with the computer, the more effectively knowledge can be used to develop and refine the model.

Contents

TOWARD ADVANCED COMPUTER-ASSISTED MODELING

Y. Sawaragi[1], *H. Fukawa*[2],
M. Ryobu[3], *and Y. Nakamori*[4]

1. INTRODUCTION

One of the difficulties in identifying multivariable, large-scale systems is the determination of structural parameters, i.e., the assumption of the forms of equations. In every mathematical modeling context (e.g., Sage, 1977; Beck, 1979; Mehra, 1980), the importance of the structure formulation is stressed before determination of the system parameters. But great difficulties are encountered in extending the existing methodology to ill-structured problems.

In uncertain environments that involve experimentation and physical laws, two types of approaches can be used to identify the optimum structure. One approach is to select a desirable structure from a set of candidate structures using certain criteria, such as Bayesian comparison (Kashyap, 1977) or pattern recognition (Vansteenkiste *et al.*, 1979). The second approach is to compound a complex structure from a combination of simple structures, starting from a linear structure (Young, 1977) or a nonlinear basic function of coupled variables (Ivakhnenko,

1. Department of Computer Science, Kyoto Sangyo University, and The Japan Institute of Systems Research, Kyoto, Japan
2. Department of Applied Mathematics and Physics, Kyoto University, Kyoto, Japan
3. The Japan Institute of Systems Research, Kyoto, Japan
4. IIASA, on leave from Department of Applied Mathematics, Konan University, Kobe, Japan

1968). A doubt, however, remains about the applicability of these approaches to system-determined (Kalman, 1980) problems wherein the qualitative aspects tend to dominate.

Data fitting of the regression type that is often used in econometric modeling requires trial-and-error methods in selecting a set of explanatory variables. The stepwise or all-subset techniques implemented in a computer reduce the burden on human effort to some extent. But the interpretation of the results is still a large task because of difficulties in checking the validity of the hypothesis testing and in giving meaning to regression coefficients. Rethinking of the resultant equations is not feasible when the number of equations is large and the cause effect relationships between variables are not known exactly in advance. Moreover, experience has taught us that statistical reliability does not ensure applicability. To avoid unnecessary complication and operational insignificance (Altman, 1980), structural considerations are crucial even in data fitting of the regression type.

As far as linear modeling is concerned, there is the idea that identification should depend on the data and only on the data (Kalman, 1980, 1983). But the majority of practical opinion emphasizes that it is difficult to build a model that does not reflect the outlook and bias of the modeler (e.g., Sage, 1977). The tendency for practitioners to have doubts about the mathematics and statistics is undeniable. A large range of complexity is methodologically undeveloped in the sense that neither analytical nor statistical methods are adequate for dealing with the systems that occur in this range (Klir, 1985). Thus, model building in uncertain environments calls for craft skills (Majone, 1984), where the word craft is used here to describe the mixture of science and art that is essential for successful application.

The most fascinating way to reflect the practical knowledge and experience of analysts and experts on model building is computer-assisted analysis, which can develop their ideas and exercise their judgment and intuition. Concepts for advanced computer-assisted modeling of different viewpoints are flourishing (e.g., Klir, 1979; Oren, 1979; Zeigler, 1984) and the actual design and implementation of interactive modeling systems has become quite active (Gelovani and Yurchenko, 1983; Fedorov *et al.*, 1984), stimulated by the rapid development of computers. The advance of computer graphics has facilitated the further development of computer-assisted modeling.

The interactive method of data handling (IMDH) presented in this paper was born in such an atmosphere, through a challenge to duplicate experts' mental models in the form of mathematical equations. But the practical problems possess their own characteristic features and await different developments. We are far from the utopia where any kind of model can be immediately developed with computer assistance. We begin in this paper with the linear modeling of a system in which the qualitative aspect is dominant, but for which extensive knowledge and cumulated experience are available. There are three categories of models, depending on the use: descriptive models, predictive or forecasting models, and planning models. Efforts to develop the methodology of modeling in order to increase decision-making capability have been made by several authors (e.g., Elzas, 1983; Zeigler, 1984). Although our ultimate goal is in this direction, the present version of IMDH is aimed at developing predictive or forecasting models.

In building a predictive or forecasting model we must separate cause from effect. The graph theoretic approach has been of great benefit in introducing assymmetric causal dependence, in which the information as to which variables appear in which equations is replaced by a directed graph with variables as nodes (e.g., Lady, 1981). Although graph theoretic techniques seem to have played a full part only in the structuring of societal systems (Harary *et al.*, 1965; Roberts, 1976; Warfield, 1976, 1982; Linstone *et al.*, 1979; Lendaris, 1980), wide applications are also reported in several fields, e.g., model simplification (Lady, 1981; Warfield, 1981), linear systems theory (Tao and Hsia, 1982; Reinschke, 1984), and economic modeling (Royer, 1980).

IMDH is a new type of linear modeling procedure with computer assistance. It requires that all the responsibility for judgments as to the structure of the model, the goodness of fit, the order of the system, and the predictive power should be attributed to the analysts and the experts, instead of using statistical or information theoretical criteria. IMDH has two extremely different features from the traditional linear modeling methods.

First, it uses a self-organization method, instead of the stepwise or all-subset procedures, in selecting explanatory variables, which makes the modeling time considerably shorter and tolerates the scarcity of data points. The self-organization method used here is a modified version of the group method of data handling (Ivakhnenko, 1968, 1970, 1971; Ivakhnenko *et al.*, 1979), that is based on heuristic principles of self-organization and relies on bioengineering concepts.

Second, instead of hypothesis testing or information theoretic criteria (Akaike, 1976; Rissanen, 1976) we use the graph theoretical techniques that are, to some extent, similar to some aspects of interpretive structural modeling (Warfield, 1974). The digraph gives insight into the cause effect relationships present in the linear model. It facilitates the interaction between analysts and the computer and then makes rethinking of the model equations quite easy.

IMDH effectively reflects the experts' knowledge on the model and assists analysts and experts to modify the model efficiently. Through the modeling process IMDH enlightens analysts about the underlying complex system, because the process of model building itself is a learning experience. IMDH accepts reactions of the analysts flexibly, and finally finds an elaborate model useful for the purpose in hand.

2. MODELING INFORMATION

The first craft required is the selection of descriptive variables. Let us write

$$X = \{ x_1, x_2, \ldots, x_m \}$$

as the set of variables chosen by analysts or experts. The set X can include nonlinear reexpressions or time-delayed variables of initial variables. Following the traditional usage, we use the term linear model to describe a set of equations whose structural parameters are embedded linearly. Reexpression and time-shifting enable us to analyze nonlinear relationships and multiple autoregressive processes, respectively.

A rigid assumption is imposed here that the corresponding data is complete in the sense that they are screened in advance to avoid multicollinearity or the influence of outliers. This does not imply that all the data should be measured absolutely correctly. Soft observation is allowed to compensate for lacking or extraordinary data. Hereafter, we use the term observation instead of measurement, meaning that observation includes data estimated or modified by the experts. Let us write the observation sequence for the variable x_i as

$$D_i = col(d_{i1}, d_{i2}, \ldots, d_{in}), \quad i = 1, 2, \ldots, m$$

and the whole observation table as

$$D = (D_1, D_2, \ldots, D_m).$$

Other modeling information involved is qualitative, i.e., the mental images of analysts or experts, among which the pairwise cause effect relationships are fed to the computer in a matrix form. Let us write $C = (c_{ij})$, $i, j = 1, 2, \ldots, m$, as an incidence matrix that characterizes the pairwise cause effect relationships. In principle, the elements of C are defined by

$$c_{ij} = \begin{cases} 0 \text{ if } x_i \textit{ never affects } x_j, \textit{ or } i = j \\ 2 \text{ if } x_i \textit{ certainly affects } x_j \\ 1 \textit{ otherwise} \end{cases}$$

A basic assumption of our argument is that much of the structure of the underlying system is ambiguous. Because both the complexity and ambiguity of an object depend on the interests and capabilities of the individual, filling in the incidence matrix is also a craft. But in-depth considerations are not required initially, rather, the way to introduce such relationships should be negative. Here, negative means that the modeler should enter into the computer a part of his knowledge only, putting the 0s and 2s in the right places. The remaining ambiguities are resolved after some iterative modeling sessions.

Starting with this *a priori* information, we find a set of linear equations:

$$x_i = a_{i0} + \sum_{x_j \in X_i} a_{ij} x_j, \qquad i = 1, 2, \ldots, m$$

where $X_i = X - \{x_i\}$, $i = 1, 2, \ldots, m$, with the hope that it could describe the underlying complex system and be capable of predicting the behavior of the system. We say that x_j is an explanatory variable for x_i if $a_{ij} \neq 0$, and that x_i is an explained variable if $a_{ij} \neq 0$ for at least one $j (\neq 0)$.

The modeling sessions are divided into two main stages. The first stage is devoted to finding a trade-off structure between the experts' mental models and the computer models. The self-organization method is used to obtain linear equations and graph theoretic techniques are used for interaction. The required human input is knowledge of the structural image of the system. This stage includes part of the model verification, because the modeler should judge whether the model behaves, in general as he intends.

The second stage is concerned with judgments about the validity of the model in terms of its explanatory and predictive powers. Prepared materials are residual plots and predictions. To check the predictive power, some of the original data are left unused during the model building. But data concerning the results of policies not implemented are generally not available, so scenario analyses are prepared. Here, both cumulative experience and deep insight into the system are required.

Even properly tested models can turn out to be inapplicable if sudden jumps occur in some variables. The validity of a model of the black-box type is usually assured only when the explanatory variables change within the data range used in the modeling, having nearly equal correlations with each other. Since any mathematical model is fatally tentative, the modeling sessions in IMDH are endless in principle. All of the modeling knowledges:

$$\{ X, D, C, \textit{computer models}, \textit{mental images} \}$$

will be refined in modeling sessions tomorrow and so be different from those of today.

3. MODELING PROCEDURES

The first task of the computer is to select the explanatory variables and estimate the coefficients in each equation using the information $\{X, D, C\}$. Let us define two subsets of X_i as follows:

$$X_i^c = \{ x_j : c_{ji} = 2 \}$$

$$X_i^o = \{ x_j : c_{ji} = 1 \}$$

The elements of X_i^c are always chosen as explanatory variables and those for X_i^o are candidates of explanatory variables in x_i. Let us call X_i^c the set of core variables and X_i^o that of optional variables, as is usual in statistical terminology. The modeler can divide the observation set D into two sets D_b and D_c ; the former is used for model building and the latter for checking the predictive power. The division can be done arbitrarily as long as the number of data points in D_b is enough to determine the parameters in the model.

First, the coefficients in equations of the form

$$x_i = a_{i0} + \sum_{x_j \in X_i^c} a_{ij} x_j$$

are estimated by the method of least squares for the variables x_i for which the core sets X_i^c are nonempty. Then, the residuals are calculated for these variables; let us write the residual variables as x_i again, noticing that the definite influences have already been accounted for. Finally, the self-organization method is used to select additional explanatory variables for the variables x_i for which the optional sets X_i^o are nonempty. The final form of the equations is written as

$$x_i = a_{i0} + \sum_{x_j \in X_i^c} a_{ij} x_j + \sum_{x_j \in X_i^o} a_{ij} x_j$$

for the variables x_i, with the unions $X_i^c \cup X_i^o$ being nonempty.

The self-organization method implemented in the computer is a modified version of the group method of data handling proposed by Ivakhnenko (1968) and can be regarded as a specific algorithm of computer artificial intelligence. The main idea was inspired by the process of crossing and selecting plants to obtain the best possible hybrid after raising several generations of the plants. We have adopted this idea in linear modeling and now explain the self-organization method used here.

Suppose that $\{x_1, x_2, \ldots, x_p\}$ is a set of candidates of explanatory variables for the variable y. The problem is to select an optimal subset of explanatory variables by which y could be explained satisfactorily in terms of a linear equation. The process consists of several layers and in each layer new variables are introduced as hybrids of a pair of variables from the previous layer. Denote by x_i^k and D_0^k the candidate of explanatory variable and the data set for model building in the kth layer, respectively. The observation set D_0^k is divided further into the training set D_{01}^k and the testing set D_{02}^k ; the former is used for model development and the latter for selection of the partial descriptions, i.e., better hybrids. The algorithm can be summarized as follows.

Algorithm of the Self-Organization Method.

Step 1. Set $k = 1$.

If $p > 1$ then go to step 2.

Otherwise estimate the coefficients of the equation:

$$y = a_0 + a_1 x_1$$

by the method of least squares with the data D_{61}^{1} .

Go to step 6.

Step 2. Estimate the coefficients of linear equations in the form:

$$y = a_{i0} + a_{i1} x_s^k + a_{i2} x_t^k$$

using the training data set D_{61}^{k} and applying the method of least squares, where i changes from 1 to ${}_nC_2$, while s moves from 1 to $p - 1$ and t from $s + 1$ to p. Note that i and the pair (s, t) have one-to-one correspondence.

Step 3. Denoting by f_i^k the estimated linear functions, let

$$y_i = f_i^k(x_s^k, x_t^k) \qquad i = 1, 2, \ldots, {}_nC_2$$

Calculate the mean square errors between y and the y_is, applying the testing data set D_{62}^{k} .

Step 4. Let

$$q = \begin{cases} p/2 & \text{if } p \text{ is even} \\ (p+1)/2 & \text{if } p \text{ is odd} \end{cases}$$

Select q functions among all of the f_i^ks so that the selected ones provide smaller mean square errors than the others.

If $q = 1$ go to step 6.

Step 5. Let $p = q$. Denote again the selected functions by

$$f_1^k, f_2^k, \ldots, f_p^k$$

Define the hybrid variables for the next layer:

$$x_i^{k+1} = f_i^k(x_s^k, x_t^k) \quad i = 1, 2, \ldots, p$$

and use these equations to generate new data sets D_{b1}^{k+1} and D_{b2}^{k+1} .

Let $k = k+1$. Return to step 2.

Step 6. Find a function among those obtained in all the layers that has the minimum mean square error; this is the final approximation. Express this final approximation using the original variables by successive substitution.

Obviously, if the number of candidates p is less than three, they are chosen unconditionally. In other words, if the number of elements in the optional set X_i^o is less than three, these elements are treated as if they belong to the core set X_i^c .

From the practical viewpoint, the smaller the number of explanatory variables is, the better. In regression or time series analysis, the problem of determination of the order of the equation is stimulating and intensive research. From our experience, the self-organization method described here chooses a moderate number of explanatory variables that are, for some reason, difficult to explain in terms of mathematical terminologies.

4. STRUCTURAL ANALYSIS

Even the experts can hardly tell whether the obtained linear model is appropriate or not because the coefficients of a linear model do not necessarily have practical meaning. Therefore, we extract the structure of the linear model in the form of digraphs and show these to the experts to assist their judgments.

Let X be the set of variables again and R be a relation on $X \times X$ defined such that (x_i, x_j) is in R if and only if x_i is an explanatory variable for x_j in the linear model. We introduce a digraph

$$G_R = (X, R)$$

where the elements of X are identified as vertices and those of R as directed lines. The vertices are represented by points and there is a directed line, called an arc, heading from x_i to x_j if and only if (x_i, x_j) is in R.

If there is a path from x_i to x_j, we say x_j is reachable from x_i and write

$$x_i \mapsto x_j$$

where the path is a sequence:

$$x_i, (x_i, x_{k_1}), x_{k_1}, \ldots, (x_{k_t}, x_j), x_j$$

If $x_i \mapsto x_j$ and $x_j \mapsto x_i$, we write

$$x_i \cong x_j$$

The digraph G_R is transitive, i.e.,

$$\text{if } x_i \mapsto x_j \text{ and } x_j \mapsto x_k \textit{ then } x_i \mapsto x_k$$

Hence the equivalence law holds with respect to $\cong$, i.e.,

(i) $x_i \cong x_j$

(ii) $x_i \cong x_j \rightarrow x_j \cong x_i$

(iii) $x_i \cong x_j,\ x_j \cong x_k \rightarrow x_i \cong x_k$

Let X' be the quotient set of X with respect to $\cong$, i.e.,

$$X' = X/\cong := \{x'_1, x'_2, \ldots, x'_{m'}\}, \quad m' \le m.$$

We can now define the condensation digraph G_C of G_R, identifying X' as the vertex set. We draw an arc from x'_p to x'_q if and only if $p \neq q$ and, for some vertices, $x_i \in x'_p$ and $x_j \in x'_q$, there is an arc from x_i to x_j in G_R. Finally, we obtain a skeleton digraph G_S, which is a minimum-arc subdigraph of G_C from which removal of any arc would destroy the reachability present in G_C. We show these digraphs to the experts in a session of IMDH and seek modification of the structure of the model.

This process of digraph modeling is carried out in the computer by a series of matrix operation steps. Many descriptions in the literature for obtaining skeleton digraphs are very complicated. We show here simple and efficient algorithms, including transitive closure, part division, hierarchical ordering, matrix condensation, and skeletonizing. Let us use the same notation R for the corresponding matrix to the relation R, defining that $R = (r_{ij})$, $i, j = 1, 2, \ldots, m$, and

$$r_{ij} = \begin{cases} 1 & \text{if } (x_i, x_j) \textit{ is in the relation } R, \textit{ or } i = j \\ 0 & \textit{otherwise.} \end{cases}$$

An interesting fact used in the matrix condensation is that if R is a reachability matrix, then the following are equivalent:

(i) $x_i \cong x_j$

(ii) the ith row (respective column) and the jth row (respective column) are identical.

The list of prepared arrays and their initial values are:

$R = (r_{ij}),\ i, j = 1, 2, \ldots, m$: the given incidence matrix

$S = (s_{ij})$: the skeleton matrix with undefined size $n \times n$

$Q = (q_{ij}),\ q_{ij} = r_{ij},\ i, j = 1, \ldots, m$: a dummy matrix

$v = (v_i),\ v_i = i,\ i = 1, 2, \ldots, m$: the index set

$a = (a_i),\ a_i = 0,\ i = 1, 2, \ldots, m$: the part indicator

$b = (b_i),\ b_i = 0,\ i = 1, 2, \ldots, m$: the level indicator

$c = (c_i),\ c_i = 0,\ i = 1, 2, \ldots, m$: the group indicator

$q = (q_i),\ q_i = 0,\ i = 1, 2, \ldots, m$: a dummy vector

The final values of arrays are: R becomes the transitive closure of the original one and its rows and columns are arranged in the hierarchical order. Rearranged variables are stored in the index set v, and arrays a, b, and c store the parts, levels, and groups to which the corresponding variables belong, respectively. The algorithms to develop a digraph model are summarized as follows.

Algorithm for Transitive Closure.

Step 1. Set $i = 0,\ s = 0$.

Step 2. Let $i = i + 1$. Set $j = 0$.

Step 3. Let $j = j + 1$. Set $t = 0,\ k = 0$.

Step 4. Let $k = k + 1$.

If $r_{ik} \times q_{kj} = 1$, then let $t = 1,\ k = m$.

If $k < m$, then repeat step 4,

otherwise if $t = 1$ and $r_{ij} = 0$, then let $r_{ij} = 1,\ s = 1$.

If $j < m$, then return to step 3,

otherwise if $i < m$, then return to step 2,
otherwise if $s = 1$, then return to step 1,
otherwise stop.

Algorithm for Part Division.

Step 1. Let $q_{ij} = \max\{ r_{ij}, r_{ji} \}$, $i, j = 1, 2, \ldots, m$.

Step 2. Take the transitive closure of $Q = (q_{ij})$.

Step 3. Set $part = 1$.

Step 4. Let $i = i + 1$.
If $i > m$, then go to step 6,
otherwise if $a_i \neq 0$, then repeat step 4,
otherwise let $a_i = part$, and set $j = i$.

Step 5. Let $j = j + 1$.
If $j < m$ and $a_j \neq 0$, then repeat step 5,
otherwise if $q_{ij} = 1$, then $a_j = part$.
If $j < m$, then repeat step 5,
otherwise if $i < m$, then $part = part + 1$ and return to step 4.

Step 6. Let $part = \max\{ a_i \}$.
If $part = 1$, then stop, otherwise set $s = m$.

Step 7. Let $s = s - 1$. Set $t = 0$.

Step 8. Let $t = t + 1$.
If $a_t > a_{t+1}$, then
swap a_t and a_{t+1}, swap v_t and v_{t+1},
swap r_{tj} and $r_{t+1,j}$, $j = 1, 2, \ldots, m$,
swap r_{jt} and $r_{j,t+1}$, $j = 1, 2, \ldots, m$.
If $t < s$, then repeat step 8,
otherwise if $s > 1$, then return to step 7,
otherwise stop.

Algorithm for Level Division.

Step 1. Set $level = 0$, $part = 0$, $t = 0$.

Step 2. Let $part = part + 1$. Set $s = t + 1$, $c = 0$, $d = 0$.

Step 3. Let $t = t + 1$.

If $a_t = part$, then let $c = c + 1$,

and if $t < m$, then repeat step 3.

If $a_t \neq part$, then let $t = t - 1$.

Set $h = t$.

Step 4. Let $level = level + 1$. Set $i = s - 1$.

Step 5. Let $i = i + 1$.

If $i > t$, then go to step 9,

otherwise if $b_i \neq 0$, then repeat step 5,

otherwise set $r = 0$, $a = 0$, $j = s - 1$.

Step 6. Let $j = j + 1$.

If $j > t$, then go to step 7,

otherwise if $b_j \neq 0$, then repeat step 6,

otherwise let $r = r + r_{ij}$ and $a = a + r_{ij} \times r_{ji}$.

If $j < t$, then repeat step 6.

Step 7. If $r = a$, then let $d = d + 1$, $q_d = i$.

If $i < t$, then return to step 5, otherwise set $l = 0$.

Step 8. Let $l = l + 1$.

If $b_{q_l} = 0$, then let $b_{q_l} = level$.

If $l < d$, then repeat step 8,

otherwise if $d < c$, then return to step 4.

Step 9. Let $h = h - 1$. Set $k = s - 1$.

Step 10. Let $k = k + 1$.

If $b_k > b_{k+1}$, then swap b_k and b_{k+1},

swap a_k and a_{k+1}, swap v_k and v_{k+1},

swap r_{kj} and $r_{k+1,j}$, $j = 1, 2, \ldots, m$,

swap r_{jk} and $r_{j,k+1}$, $j = 1, 2, \ldots, m$.

If $k < h$, then repeat step 10,

otherwise if $h > s$, then return to step 9,

otherwise if $t < m$, then return to step 2,

otherwise stop.

Algorithm for Group Division.

Step 1. Let $group = 1$, $level = 0$, $t = 0$.

Step 2. Let $level = level + 1$. Set $s = t + 1$.

Step 3. Let $t = t + 1$.
If $t < m$ and $b_t = level$, then repeat step 3.
If $b_t \neq level$, then let $t = t - 1$.
Set $h = t$.

Step 4. Set $i = s - 1$.

Step 5. Let $i = i + 1$.
If $i > t$, then go to step 8,
otherwise if $c_i \neq 0$, then repeat step 5,
otherwise set $c_i = group$, $j = i$.

Step 6. Let $j = j + 1$.
If $j > t$ then return to step 5,
otherwise if $c_j \neq 0$, then repeat step 6,
otherwise set $q = 0$, $c = 0$.

Step 7. Let $q = q + 1$.
If $r_{iq} = r_{jq}$, then $c = c + 1$.
If $q < m$, then repeat step 7,
otherwise if $c = m$, then $c_j = group$.
If $j < t$, then return to step 6,
otherwise if $i < t$, then $group = group + 1$ and return to step 5.

Step 8. Let $h = h - 1$. Set $k = s - 1$.

Step 9. Let $k = k + 1$.
If $c_k > c_{k+1}$, then
swap c_k and c_{k+1}, swap b_k and b_{k+1},
swap a_k and a_{k+1}, swap v_k and v_{k+1},
swap r_{kj} and $r_{k+1,j}$, $j = 1, 2, \ldots, m$,
swap r_{jk} and $r_{j,k+1}$, $j = 1, 2, \ldots, m$.
If $k < h$, then repeat step 9,
otherwise if $h > s$, then return to step 8,
otherwise if $t < m$, then return to step 2,
otherwise stop.

Algorithm for Condensation and Skeletonizing.

Step 1. Set $q_1 = 1$, $i = 1$.

Step 2. Let $i = i + 1$.
If $c_i = c_{i-1}$, then let $q_i = 0$, otherwise let $q_i = 1$.
If $i < m$, then repeat step 2.

Step 3. Let $n = c_m$. Set $i = 0$, $k = 0$.

Step 4. Let $i = i + 1$.
If $i > m$, then go to step 6,
otherwise if $q_i = 0$, then repeat step 4,
otherwise let $k = k + 1$ and set $h = 0$, $j = 0$.

Step 5. Let $j = j + 1$.
If $j > m$, then return to step 4,
otherwise if $q_j = 0$, then repeat step 5,
otherwise let $h = h + 1$.
If $k \neq h$, then let $s_{kh} = r_{ij}$.
If $j < m$, then repeat step 5,
otherwise if $i < m$, then return to step 4.

Step 6. Set $i = 0$.

Step 7. Let $i = i + 1$. Set $j = i$.

Step 8. Let $j = j + 1$. Set $k = j$.

Step 9. Let $k = k + 1$.
If $s_{ji} \times s_{kj} = 1$, then let $s_{ki} = 0$.
If $k < n$, then repeat step 9,
otherwise if $j < n - 1$, then return to step 8,
otherwise if $i < n - 2$, then return to step 7,
otherwise stop.

The skeleton digraph can be drawn as follows. First we write elements of the group indicator c one by one in a circle from top to bottom, except for the same elements as appeared before. Then we draw an arc between the circles if the corresponding entry of the skeleton matrix is 1. Finally, we amend the format of the hierarchy to facilitate interpretation of the skeleton.

5. INTERACTIVE MODELING

Here we summarize the whole process of IMDH. As mentioned already, the modeling sessions consist of two main stages. The first stage is devoted to finding a trade-off structure between the computer models and the experts' mental models. The dialogue continues until the cause effect relation in the computer model becomes satisfactory. The second stage is related to judgments of the explanatory and predictive powers of the computer model obtained in the first stage. If the model is not satisfactory, then the modeling process is repeated from the beginning. The whole process is schematized in Figure 1 and the dialogues are summarized as follows.

The First Stage Dialogue.

Step 1. (*Expert*) edits the set of descriptive system variables and prepares the observation table.

Step 2. (*Expert*) introduces the cause effect relationships between variables.

Step 3. (*Computer*) finds a linear model, i.e., a set of linear equations using the self-organization method.

Step 4. (*Computer*) displays the cause effect relationships embedded in the linear model in terms of hierarchical digraphs.

Step 5. (*Expert*) amends the digraph by adding or removing arcs in it, if necessary. If the amendments cause changes in the cause effect relationships in the linear model, then the modeling session returns to step 2, otherwise it proceeds to the second stage dialogue.

The Second Stage Dialogue.

Step 6. (*Computer*) provides residual plots and predictions, and also assists the scenario analysis.

Step 7. (*Expert*) looks for the equations that have weak explanatory and predictive powers. If there are such equations, the modeling session returns to the beginning.

There are several points that are fascinating in computer-assisted modeling and require sophisticated computer software for effective interaction. They include:

(1) Data screening and transformation of variables.

(2) Introduction of the initial version of cause effect relationships.

(3) Format of and substantial amendments to digraphs.

(4) Reflection of amendments in the digraphs on the incidence matrix.

(5) Graphic displays of the residuals and predictions.

(6) Interactive scenario analysis.

We are now developing the computer software for the method proposed in this paper. The detailed treatments of these points are described in a separate publication (Nakamori, *et al.*, 1985).

As an important application of IMDH, we have been engaged in a regional economic-forecasting model for Kyoto, Japan. Here we present a brief summary of a result obtained using IMDH. The selected variables are shown in Table 1. Besides these original variables, one- and two-year time-delayed variables are taken into consideration. After four-time repetitions of the process of IMDH, we and the experts reached a final agreement on the incidence matrix, as shown in Table 2, where the time-delayed variables are assumed to have the same dependencies as the original ones. From this matrix the forecasting model was obtained, as shown in Table 3, and the corresponding digraph is shown in Figure 2.

The data used in the model is from 1960 and 1976 and the predictions of the obtained model are summarized in Table 4. This result is fairly satisfactory from the viewpoint of the consumed time for modeling, which was about 27 hours, including calculations and discussions. Generally, it is very difficult to modify a large-scale model once obtained because of the cost and time. IMDH overcomes this difficulty.

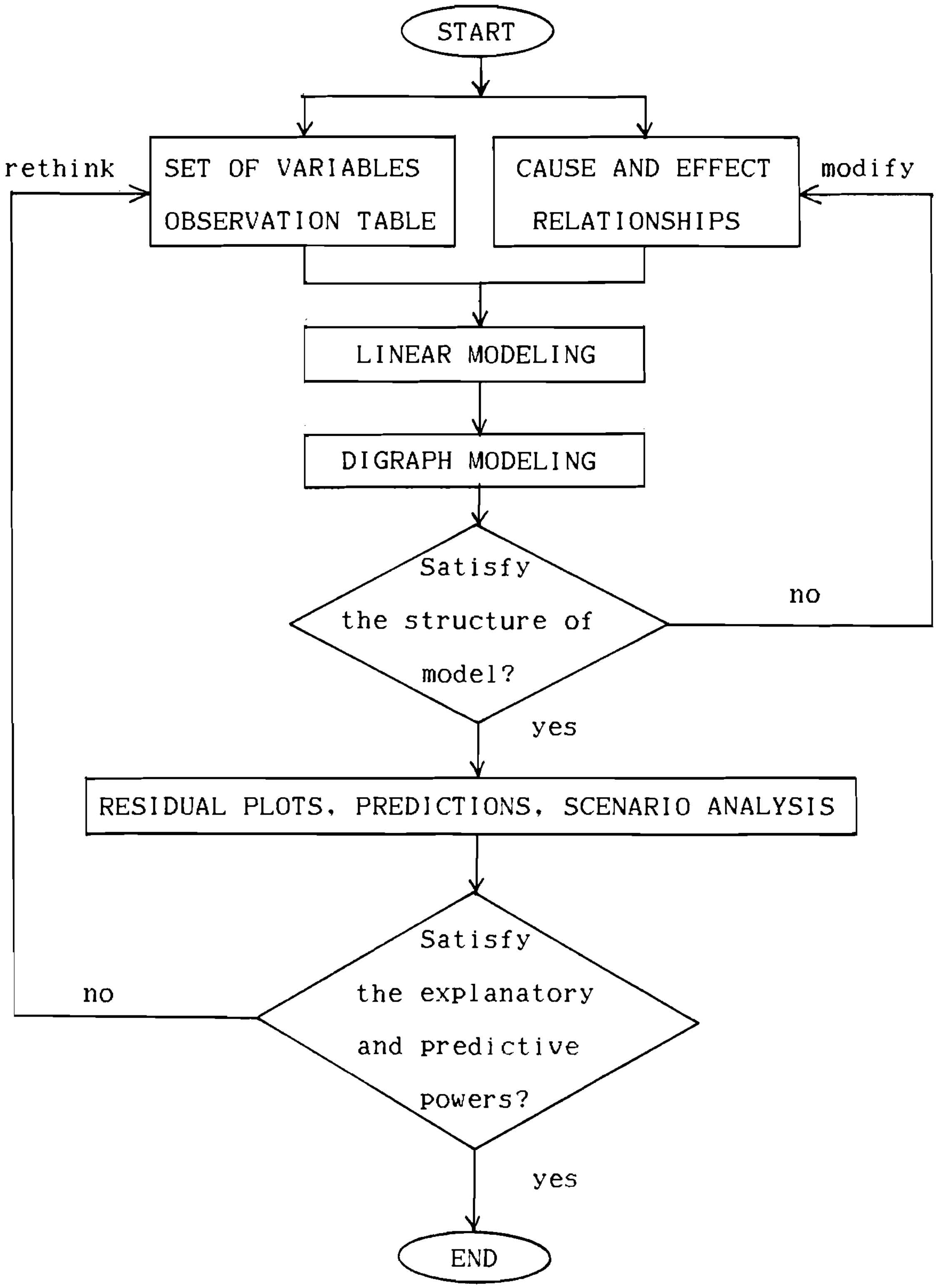

Figure 1. Structure of the interactive method of data handling.

Table 1. Selected variables in modeling.

I.	The population in Kyoto City
1.LA	Little age (age: 0~14)
2.PA	Productive age (age: 15~64)
3.OA	Old age (age: 65~)
4.BI	Birth
5.DP1	Daytime population of the primary industry
6.DP2	Daytime population of the secondary industry
7.DP3	Daytime population of the tertiary industry
8.UP1	Usual population of the primary industry
9.UP2	Usual population of the secondary industry
10.UP3	Usual population of the tertiary industry
II.	The population within Kyoto zone (except Kyoto City)
11.LAOU	Little age out of Kyoto City (age: 0~14)
12.PAOU	Productive age out of Kyoto City (age: 15~64)
13.OAOU	Old age out of Kyoto City (age: 65~)
14.BIOU	Birth out of Kyoto City
15.DPOU1	Daytime population of the primary industry out of Kyoto City
16.DPOU2	Daytime population of the secondary industry out of Kyoto City
17.DPOU3	Daytime population of the tertiary industry out of Kyoto City
18.UPOU1	Usual population of the primary industry out of Kyoto City
19.UPOU2	Usual population of the secondary industry out of Kyoto City
20.UPOU3	Usual population of the tertiary industry out of Kyoto City
III.	The industries
(Primary industry)	
21.PI	Primary industry
(Secondary industry)	
22.CON	Construction industry
23.TEX	Textile industry
24.MAC	Machine and metalworking industry
25.OTSE	Other industry
26.MIN	Mining industry
(Tertiary industry)	
27.WHO	Wholesale trade
28.RET	Retail trade
29.SER	Service
30.PUB	Public service
31.OTER	Others
IV.	The size of land
32.COL	Commercial
33.INL	Industry
34.HOUL	Housing
35.OTL	Others
V.	The others
36.CIN	Civil income
37.GAP	General accounts of Kyoto prefecture
38.SAP	Special accounts of Kyoto prefecture
39.GAC	General accounts
40.SAC	Special accounts
41.SIGH	Sightseer
42.ROAR	Road area

Table 2. The incidence matrix just before the final session.

1.LA	11100	00000	00000	00000	00000	00001	00011	00000	00
2.PA	11111	11111	00000	00000	11111	11111	11111	11111	11
3.OA	00101	11111	00000	00000	00000	01111	11111	11111	11
4.BI	11010	00000	00000	00000	00000	00000	00011	00000	00
5.DP1	11111	00100	00000	00000	10100	00111	11011	00000	11
6.DP2	11110	10010	00000	00000	01111	10111	11111	11111	11
7.DP3	11110	01001	00000	00000	00100	01111	11011	11111	11
8.UP1	11111	00100	00000	00000	10100	00111	11011	00000	11
9.UP2	11110	10010	00000	00000	01111	10111	11111	11111	11
10.UP3	11110	01001	00000	00000	00100	01111	11011	11111	11
11.LAOU	00000	00000	11100	00111	00000	00000	00000	00000	00
12.PAOU	00001	11000	11111	10111	00101	00000	01111	11111	11
13.OAOU	00000	00000	00100	01111	00000	00000	00000	00000	00
14.BIOU	00000	00000	11010	00111	00000	00000	00000	00000	00
15.DPOU1	00000	00000	11111	00111	00100	00000	00000	00000	00
16.DPOU2	00000	00000	11110	10111	00101	00000	00000	00000	00
17.DPOU3	00000	00000	11110	01111	00100	00000	00000	00000	00
18.UPOU1	00000	00000	11111	00111	00100	00000	00000	00000	00
19.UPOU2	00000	00000	11110	10111	00101	00000	00000	00000	00
20.UPOU3	00000	00000	11110	01111	00100	00000	00000	00000	00
21.PI	00001	00100	00000	00000	10000	00000	00011	11111	01
22.CON	00000	10010	00000	00000	01011	10001	11011	11111	01
23.TEX	00000	10010	00000	00000	00100	01001	11111	11111	01
24.MAC	00000	10010	00000	00000	00011	10001	11111	11111	01
25.OTSE	00000	10010	00000	00000	00011	10001	11111	11111	01
26.MIN	00000	10010	00000	00000	00000	10001	10011	11111	01
27.WHO	00000	01001	00000	00000	00100	01111	11011	11111	01
28.RET	00000	01001	00000	00000	00100	01111	11011	11111	11
29.SER	00000	01001	00000	00000	00100	01111	11011	11111	11
30.PUB	00000	01001	00000	00000	00000	00011	11011	11111	01
31.OTER	00000	01001	00000	00000	00000	01111	11011	11111	01
32.COL	00000	00000	00000	00000	01000	01101	11011	00000	11
33.INL	00000	00000	00000	00000	01011	00101	10111	00000	11
34.HOUL	00000	00000	00000	00000	11000	00101	10011	00000	01
35.OTL	00000	00000	00000	00000	01000	00101	11011	00000	01
36.CIN	00001	11111	00000	00000	11111	11111	11111	11111	01
37.GAP	00001	11111	00001	11111	11111	11001	11111	11111	11
38.SAP	00001	11111	00001	11111	11111	11001	11111	11111	11
39.GAC	00001	11111	00000	00000	11111	11001	11111	11111	11
40.SAC	00001	11111	00000	00000	11111	11001	11111	11111	11
41.SIGH	00001	11111	00000	11111	00000	00111	11111	11111	11
42.ROAR	00001	11111	00001	11111	11111	11111	11111	11111	11

Table 3. A regional economic forecasting model using IMDH.

1. $LA=-0.4603+0.7514(LA)_{-1}+0.1361(DP1)_{-2}+0.5660(DP3)_{-1}$
2. $PA= 0.4551+0.4905(PA)_{-1}+0.0647(DP2)_{-2}$
3. $OA=-0,0654+0.9546(OA)_{-1}+0.1484(DP3)_{-1}$
4. $BI=-0.0118+0.8930(BT)_{-1}+0.1246(DP1)_{-2}$
5. $DP1= 0.0054+1.4648(DP1)_{-1}-0.4922(UP1)_{-2}$
6. $DP2= 0.2075+0.7972(UP2)_{-1}-0.0046(GAP)_{-1}$
7. $DP3= 0.1043+0.9219(DP3)_{-1}-0.0039(ROAR)_{-1}$
8. $UP1=-0.0090+1.6172(DP1)_{-1}-0.6211(UP1)_{-2}$
9. $UP2= 1.0767-0.0423(GAP)_{-1}-0.0017(SAP)_{-2}+0.0379(SAC)_{-1}$
10. $UP3= 0.1785+0.8013(UP3)_{-1}+0.0386(OA)_{-1}$
11. $LAOU= 0.2962-0.2403(DPOU1)_{-1}+0.6161(LAOU)_{-1}$
    $+0.0374(UPOU1)_{-2}+0.1003(PAOU)_{-2}+0.2361(DPOU3)_{-2}$
12. $PAOU= 0.3458+0.6741(UPOU3)_{-1}+0.0137(BIOU)_{-2}$
13. $OAOU= 0.0273+0.6680(OAOU)_{-1}+0.3477(DPOU3)_{-1}$
14. $BIOU= 0.0700+1.0914(BIOU)_{-1}+0.0960(UPOU2)_{-2}$
    $-0.2105'(DPOU2)_{-2}$
15. $DPOU1= 0.8589+0.4432(DPOU1)_{-1}-0.1116(SAP)_{-1}$
    $-0.2391(PAOU)_{-1}$
16. $DPOU2=-0.3231+1.5125(UPOU2)_{-1}-0.0092(DPOU2)_{-1}$
    $+1.7003(DPOU2)_{-2}-1.8641(DPOU2)_{-2}$
17. $DPOU3=-0.0434+0.5823(UPOU3)_{-1}-0.0549(DPOU3)_{-2}$
    $+0.284348(DPOU3)_{-1}+0.1079(DPOU3)_{-2}+0.1769(POAR)_{-2}$
18. $UPOU1= 0.5813-1.0062(DPOU1)_{-1}+1.7106(UPOU1)_{-1}$
    $-0.6542(DPOU3)_{-2}+0.2951(DPOU3)_{-1}$
19. $UPOU2=-0.2309+3.7767(UPOU2)_{-1}-1.9758(DPOU2)_{-1}$
    $-0.5618(UPOU2)_{-2}-0.0001(BIOU)_{-1}$
20. $UPOU3= 0.0869+1.2461(UPOU3)_{-1}-0.2813(OAOU)_{-1}$
21. $PI= 0.3526+0.4257(GAP)_{-1}+0.3160(PI)_{-1}+0.2839(GAP)_{-2}$
    $-0.2422(GAC)_{-1}$
22. $CON=-4.531-1.7281(HOUL)_{-2}+12.0319(INL)_{-2}-5.0845(INL)_{-1}$
23. $TEX=-0.6415+3.8366(UPOU3)_{-1}+0.2085(PAOU)_{-1}+0.0157(UPOU3)_{-2}-2.6142(DPOU3)_{-1}+0.2901(TEX)_{-1}$
24. $MAC=-3.3813+2.9858(INL)_{-2}+0.2786(INL)_{-1}+0.0210(ROAR)_{-2}$
    $+0.2772(CON)_{-1}+0.9670(ROAR)_{-1}$
25. $OTSE= 0.1106+2.0803(CIN)_{-1}-1.4501(CIN)_{-2}+0.1145(OTSE)_{-1}$
    $+0.1225(CON)_{-2}$
26. $MIN=-6.9244-7.5986(DP2)_{-1}+0.5778(MIN)_{-1}+14.8971(UP2)_{-2}$
27. $WHO= 0.1514+0.5892(SAP)_{-2}+0.3515(RET)_{-1}+0.0642(CIN)_{-2}$
    $+0.0658(RET)_{-1}+0.0184(CIN)_{-1}$
28. $RET=-1.6999+2.3976(HOUL)_{-2}+0.5063(RET)_{-1}$
29. $SER=-1.7299-1.6583(OA)_{-2}+4.1885(OA)_{-1}+0.2927(OTER)_{-2}$
30. $PUB=-0.0161+0.4285(RET)_{-2}+0.3346(WHO)_{-1}+0.0254(GAP)_{-1}$
    $+0.4406(WHO)_{-1}+0.0166(OTER)_{-1}$
31. $OTER= 0.1001+0.2897(OTER)_{-1}+0.5438(SAC)_{-2}+0.0293(SAP)_{-1}$
    $+0.3044(CIN)_{-1}$
32. $COL= 0.4949+0.5054(COL)_{-2}+0.0129(UP3)_{-1}$
33. $INL= 0.0530+1.6367(INL)_{-1}-0.6836(INL)_{-2}$
34. $HOUL=-0.3133+0.5752(HOUL)_{-1}+0.3590(DP3)_{-1}+0.4197(DP3)_{-2}$
35. $OTL= 0.9500+0.0447(OTER)_{-1}+0.0393(CIN)_{-1}$
36. $CIN= 0.0648+1.4425(CIN)_{-1}-0.3557(CIN)_{-2}-0.0599(PAOU)_{-2}$
37. $GAP=-0.0173+0.5850(GAP)_{-1}+0.5898(GAP)_{-2}$
38. $SAP=-0.3596+1.2708(TEX)_{-1}+0.1026(CON)_{-1}+0.1166(PAOU)_{-2}$
39. $GAC=-0.1452+0.9229(GAP)_{-1}+0.4597(RET)_{-1}$
40. $SAC=-0.1855+0.1680(SAC)_{-1}-0.0419(GAP)_{-2}+0.6422(GAP)_{-1}$
    $+0.7287(PUP)_{-2}$
41. $SIGH=-0.1292+0.7998(SER)_{-2}+0.3800(SER)_{-1}+0.1546(SIGH)_{-1}$
42. $ROAR= 0.2813+0.7422(DP3)_{-2}+0.0117(DP1)_{-2}$

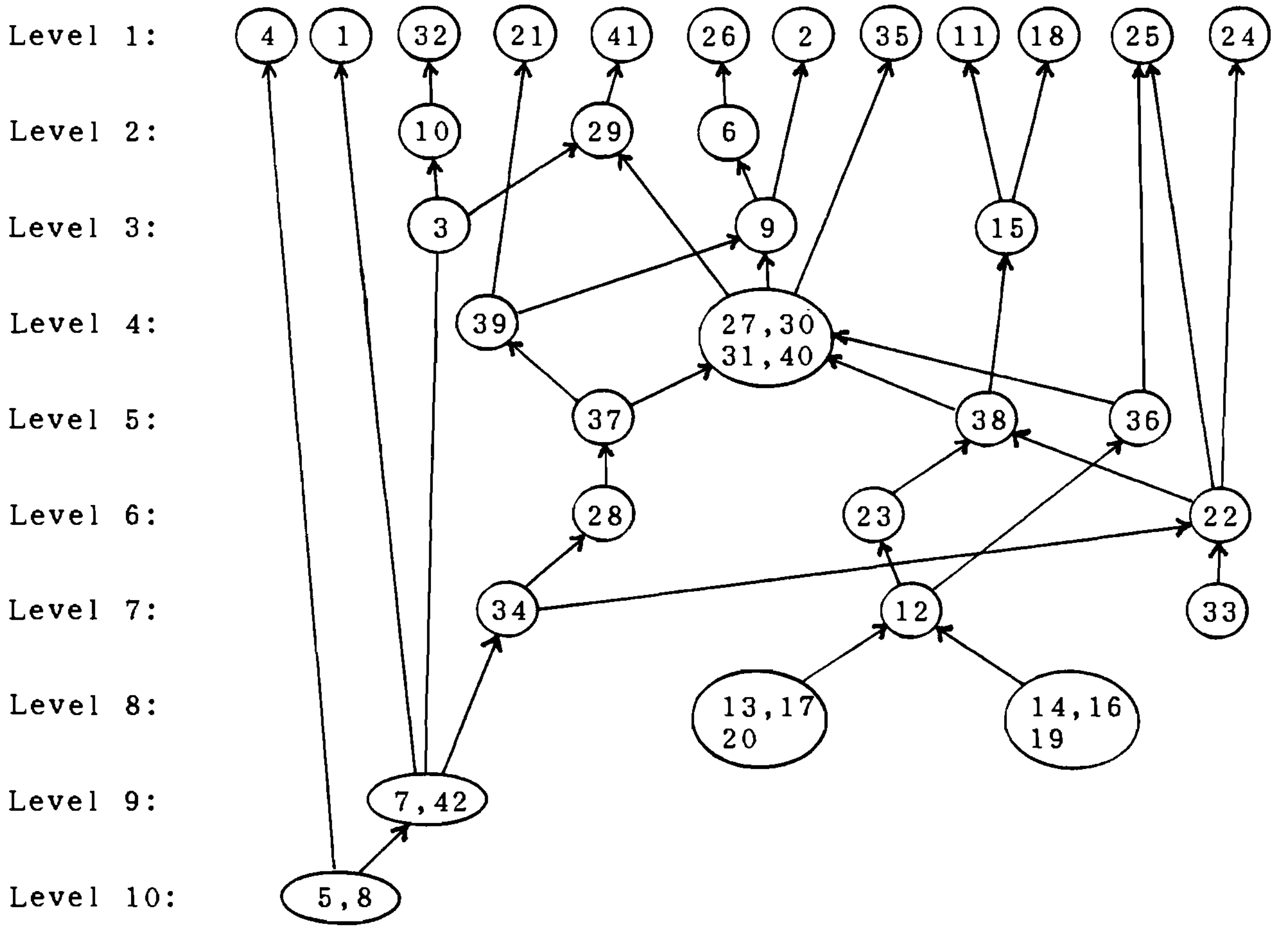

Figure 2. The skeleton digraph corresponding to the incidence matrix.

Table 4. Economic forecasting by the obtained model.

Rate of Growth(%) Year	Past records 1975	1976	Forecasts 1977	1978	1979
Pure production	4.2	10.9	3.3	8.9	9.7
Primary industry	7.4	7.9	11.2	9.4	7.2
Secondary industry	-4.1	7.2	2.4	6.6	7.4
Manufacturing	-6.2	8.3	6.5	7.6	8.3
T E X	0.9	7.1	6.7	8.0	8.2
M A C	-16.3	9.6	1.5	4.2	5.2
O T S E	3.6	8.3	10.2	9.8	10.4
C O N	7.3	2.3	-18.0	0.2	1.0
M I N	-19.8	-41.1	-91.3	—	—
Tertiary industry	9.2	12.8	3.7	10.0	10.9
W H O	9.6	6.5	2.1	7.3	11.4
R E T	10.6	10.9	2.1	6.1	7.1
S E R	8.7	16.6	4.8	11.9	11.3
Component Ratio(%) Year	**Past records 1975**	**1976**	**Forecasts 1977**	**1978**	**1979**
Pure production	100.0	100.0	100.0	100.0	100.0
Primary industry	0.4	0.4	0.5	0.5	0.4
Secondary industry	34.6	33.4	33.1	32.5	31.8
Manufacturing	28.6	27.9	28.8	28.4	28.1
T E X	8.8	8.5	8.8	8.7	8.6
M A C	8.6	8.5	8.3	8.0	7.7
O T S E	11.2	10.9	11.6	11.7	11.8
C O N	6.0	5.5	4.4	4.0	3.7
M I N	0.0	0.0	0.0	0.0	0.0
Tertiary industry	65.0	66.1	66.4	67.0	67.8
W H O	20.3	19.5	19.3	19.0	19.3
R E T	6.9	6.9	6.8	6.6	6.5
S E R	37.8	39.7	40.3	41.4	42.0
Amount(10^6yen) Year	**Past records 1975**	**1976**	**Forecasts 1977**	**1978**	**1979**
Pure production	1918107	2126707	2197521	2392499	2625625
Primary industry	8417	9056	10068	11011	11808
Secondary industry	663081	711011	728381	776711	834032
Manufacturing	548163	593648	632308	650458	736796
T E X	168793	180770	192944	208333	225416
M A C	164957	180770	183467	191211	201154
O T S E	214413	232108	255897	280914	310226
C O N	114529	117134	97053	96253	97236
M I N	389	229	20	—	—
Tertiary industry	1246609	1406640	1459072	1604777	1779785
W H O	389376	414708	423313	454329	506132
R E T	132349	146743	149794	158989	170289
S E R	724884	845189	885965	991459	1103364

6. CONCLUSION

IMDH starts with a belief in the prepared observation and, after iterative modeling sessions, it develops and refines both the computer models and the human mental models. Computer models can be obtained even when the amount of data is scarce, owing to the self-organization method, and easily modified with the assistance of graph-theoretic techniques.

Because the modeling can be done at low cost and in a short time and because this method intends to develop tentative models, a variety of applications is expected. Actually, we are now engaging in the development of regional economic forecasting models of Kyoto, Japan, as presented briefly in the previous section. Also, as a collaborative work with the IIASA Regional Water Policy Project (Project Leader: S.A. Orlovski) and its successive project (Decision Support Systems for Managing Large International Rivers), we are developing and elaborating a computer system to obtain water resources models usable in decision support systems.

REFERENCES

Akaike, H. (1976). Canonical Correlation Analysis of Time Series and the Use of an Information Criterion. *System Identification: Advances and Case Studies.* (eds. Mehra, R.K. and Lainiotis, D. G.). Academic Press, New York. pp. 27-96.

Altman, S. M. (1980). Pitfalls of Data Analysis. *Pitfalls of Analysis.* (eds. Majone, G. and Quade, E.). John Wiley, New York. pp. 44-56.

Beck, M. B. (1979). Model Structure Identification from Experimental Data. *Theoretical Systems Ecology: Advances and Case Studies.* (ed. Halfon, E.). Academic Press, New York. pp. 260-289.

Elzas, M. S. (1983). The Use of Structured Design Methodology to Improve Realism in National Economic Planning. *Adequate Modeling of Systems.* (ed. Wedde, H.). Springer-Verlag, Berlin. pp. 296-314.

Fedorov, V., Korostelev, A., and Leonov, S. (1984). *Analysis and Design of Simulation Experiments with Linear Approximation Models.* WP-84-74. IIASA, Laxenburg, Austria.

Gelovani, V. A. and Yurchenko, V. V. (1983). System for Interactive Modelling of Socio-Economic Development. *Adequate Modeling of Systems* (ed. Wedde, H.). Springer-Verlag, Berlin. pp. 267-273.

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

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. (1970). Heuristic Self-Organization in Problems of Engineering Cybernetics. *Automatica.* Vol. 6, pp. 207-219.

Ivakhnenko, A. G. (1971). Polynomial Theory of Complex System. *IEEE Trans. Syst. Man Cybern.* Vol. SMC-1, No. 4, pp. 364-378.

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: Advance and Case Studies.* (ed. Halfon, E.). Academic Press, New York. pp. 325-352.

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

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

Kashyap, R. L. (1977). A Bayesian Comparison of Different Classes of Dynamic Models using Empirical Data. *IEEE Trans. Automatic Control*. Vol. AC-22, No. 5, pp. 715-727.

Klir, G. J. (1979). Computer-Aided Systems Modeling. *Theoretical Systems Ecology: Advance and Case Studies*. (ed. Halfon, E.). Academic Press, New York. pp. 260-289.

Klir, G. J. (1985). Complexity: Some General Observations. *Systems Research*. Vol. 2, No. 2, pp. 131-140.

Lady, G. M. (1981). On Organizing Analysis. *Computer-Assisted Analysis and Model Simplification*. (eds. Greenberg, H. J. and Maybee, J. S.). Academic Press, New York. pp. 1-15.

Lendaris, G. G. (1980). Structural Modeling - A Tutorial Guide. *IEEE Trans. Syst. Man Cybern*. Vol. SMC-10, No. 12, pp. 807-840.

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.

Majone, G. (1984). The Craft of Applied Systems Analysis. *Rethinking the Process of Operational Research and Systems Analysis*. (eds. Tomlinson, R. and Kiss, I.). Pergamon Press, Oxford. pp. 143-157.

Mehra, R. K. (1980). A Survey of Time Series Modeling and Forecasting Methodology. *Real-Time Forecasting/Control of Water Resource Systems*. (ed. Wood, E. F.). Pergamon Press, Oxford. pp. 7-33.

Nakamori, Y., Ryobu, M., Fukawa, H., and Sawaragi, Y. (1985). *An Interactive Modeling Support System*. WP-85-77. IIASA, Laxenburg, Austria.

Oren, T. I. (1979). Concepts for Advanced Computer Assisted Modelling. *Methodology in Systems Modelling and Simulation*. (eds. Zeigler, B. P., Elzas, M. S., Klir, G. J., and Oren, T. I.). North-Holland, Amsterdam. pp.29-55.

Reinschke, K. J. (1984). A Graph-Theoretic Criterion for Structural Completeness of Linear Systems. *Systems Analysis Modelling Simulation*. Vol. 1, No. 5, pp. 373-379.

Rissanen, J. (1976). Minimax Entropy Estimation of Models for Vector Processes. *System Identification: Advances and Case Studies*. (eds. Mehra, R. K. and Lainiotis, D. G.). Academic Press, New York. pp. 97-119.

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

Royer, D. (1980). Qualitative Hierarchies Among Variables of a Macroeconomic Decision Model. *Dynamic Modelling and Control of National Economics (eds. Janssen, J.M.L., Pau, L.F. and Straszak, A.J.). Proc. 3rd IFAC/IFORS Conference*. Warsaw, Poland, 16-19 June 1980. pp. 355-360.

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

Tao, K. M. and Hsia, P. (1982). On the Structure of Directed Graphs with Applications: A Rapproachement with Systems Theory, Part I. *Int. J. General Systems*. Vol. 8, No. 3, pp. 147-159. Part II. Vol. 8, No. 4, pp. 211-224.

Vansteenkiste, G. C., Spriet, J., and Bens, J. (1979). Structure Characterization for Systems Modeling in Uncertain Environments. *Methodology in Systems Modelling and Simulation*. (eds. Zeigler, B. P., Elzas, M. S., Klir, G. J. and Oren, T. I.). North-Holland, Amsterdam. pp. 289-303.

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

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

Warfield, J. N. (1981). Learning Through Model Building. *Computer-Assisted Analysis and Model Simplification*. (eds. Greenberg, H. J. and Maybee, J. S.). Academic Press, New York. pp. 69-78.

Warfield, J. N. (1982). Organization and Systems Learning. *General Systems*. Vol. XXVII, pp. 5-74.

Young, P. (1978). A General Theory of Modeling for Badly Defined Systems. *Proc. IFIP Working Conf. Modeling, Identificatin and Control in Environmental Systems* (ed. Vansteenkiste, G.C.). Ghent, Belgium, August 30-September 2, 1977. North-Holland, Amsterdam. pp. 103-136.

Zeigler, B. P. (1984). *Multifacetted Modeling and Discrete Event Simulation.* Academic Press, London.