AN APPROACH TO DISTRIBUTED MODELING

Alexander Umnov and Murat Albegov
International Institute for Applied Systems Analysis, Laxenburg, Austria

Research Reports, which record research conducted at IIASA, are independently reviewed before publication. However, the views and opinions they express are not necessarily those of the Institute or the National Member Organizations that support it.

FOREWORD

Much work aimed at analyzing the behavior of large complex systems is based on building large integrated models. However, such models are often difficult to calibrate, manipulate, and explain, and results from them are often hard to interpret.

Thus, there is a need to develop procedures for linking independent submodels into a larger system. There are two main approaches to meeting this need, each of which has its own difficulties. One is to make the large model decomposable, the other is to tie separate submodels together by analysts who exercise suitable judgments as the analysis proceeds.

This paper proposes a compromise between these two approaches that links submodels formally, but without building a large integrated model explicitly. It is based on the "smooth" version of the sequential unconstrained minimization technique (SUMT); from a mathematical point of view it can be viewed as a realization of a general composition scheme.

While this approach will be of value in many areas of systems analysis, at IIASA it has been used successfully in dealing with the complexities of regional analysis, where many different activities (and hence the variables and submodels that represent them) are brought together in a common geographical context.

BORIS ISSAEV
Leader
Regional Development Group

AN APPROACH TO DISTRIBUTED MODELING[1]

by Alexander Umnov and Murat Albegov

International Institute for Applied Systems Analysis

The problem of computer linkage of different mathematical models into a whole system in order to investigate their joint behavior with more common criteria and constraints seems to receive more and more consideration. Many works on analyzing the behavior of complex systems are based on building large-scale integrated models and sequentials using decomposition and aggregation procedures. In this article an approach is described which permits the investigation of a set of linked subsystems without explicitly building any integrated model.

This article represents a description of a particular approach which might be referred to as "distributed modeling." It deals with conceptual systems in general, which might be used to model concrete systems at any level. The use of the method is illustrated by the practical application of the development of a system of regional models. This approach, based on the smooth version of the sequential unconstrained minimization techniques (SUMT), can be considered from a mathematical point of view as a realization of the general decomposition scene.

∽

INTRODUCTION

IN INVESTIGATING an object of a complex structure, it is reasonable at the first stage of systems analysis to consider this object as a set of its independent parts. In this way we can build mathematical models of all these parts at a sufficiently high level of detail. At the second stage, we have to take into consideration all interactions between the subsystem parts when they operate under common criteria and constraints.

All this raises the necessity of developing procedures, both methodological and computational, which give us the possibility of linking independent submodels into a whole system. There are two main approaches for solving this problem: The first is to design a large-scale integrated mathematical model describing the behavior of the system as a whole and consequent decomposition of the model. This method is very convenient for use in computer analysis because it needs comparatively low hardward and software resources. On the other hand, appropriate transformations of all files of the submodels into a common form can be made using this approach. Besides, it is sometimes difficult to trace the process of optimization, which can give information of a significant practical value.

The second approach consists of using one or several analysts or decision makers to organize interactions between submodels. This method permits us to link mathematical models without any file transformations, but it is practically impossible to use optimization procedures because of the time required per iteration. Therefore, it seems very desirable for one to develop an approach that permits linkage of different submodels in a direct way, without building a large-scale model which is to be partitioned later. We would also like this approach to enable linkage of submodels prepared independently by different groups of specialists and to give us the possibility of using different mathematical methods for solving the subproblems, perhaps on different computers. Briefly speaking, this approach considers all the submodels which are linked to be black boxes, assuming that their input and output data are available

354

for use. The discussion of the approaches to the linkage of models and different economic applications is given by Bagrinovskii (1977). It should be noted that the second approach (which is generally more concerned with the linkage problem) has been investigated in less detail than the first approach (usually associated with the decomposition of the problem). The purpose of this paper is to describe a scheme of realization of the second approach and to discuss its applications.

STATEMENT OF THE PROBLEM

The easiest way to link different submodels into a whole system consists of using special variables to formalize interrelations between submodels. These variables, here called common variables or coupling variables, will be denoted as V, in contrast to inner variables of the submodels, here denoted as X.

There are different ways of introducing these common variables, but we will now consider the general case, omitting some details which will be discussed in other sections. Let us assume that each of the submodels can be formulated in terms of inner variables as follows:

Minimize with respect to X^k

$$F^k(X^k),$$

subject to

$$X^k \in \mathscr{E}^{n^k} \quad G_s{}^k(X^k) \geq 0, s = \overline{1, m^k}$$
$$k = \overline{1, N},$$

where N is the number of submodels to be linked.

It is important to emphasize that all functions of F^k and $G_s{}^k$ are not known, since we decided to consider the submodels as black boxes.

After introducing common variables in an appropriate way, we have the following statement for each subproblem:

(1)   Minimize with respect to X^k

$$F^k(X^k, V),$$

subject to

$$X^k \in \mathscr{E}^{n^k} \quad G_s{}^k(X^k, V) \geq 0, s = \overline{1, m^k},$$

where V is a vector of common variables, $V \in \mathscr{E}^L$, and is fixed in problem (1).

The second step of linking these submodels consists in formalizing relations between them. Let these relations be given as a set of constraints on common variables

$$R_s(V) \geq 0, \qquad s = \overline{1, M},$$

where M is the number of these constraints.

Finally, we have to formulate the common criterion of operating the whole system of submodels to be linked. There are several reasons why we should use this criterion as a linear combination of criteria from different submodels, subject to all weight coefficients of this combination which are nonnegative. Some of the foundations of this choice will be discussed later on.

Therefore, we have the following system of relations to find optimal values of inner variables X^k and common variables V:

(2)   Minimize with respect to X^k and V

$$\sum_{k=1}^{k=N} \lambda_k F^k(X^k, V)$$

subject to

$$G_s{}^k(X^k, V) \geq 0, s = \overline{1, m^k}; k = \overline{1, N},$$
$$R_s(V) \geq 0, \qquad s = \overline{1, M},$$

where λ_k are nonnegative weight coefficients.

Theoretically, a solution of this problem gives us all the desired data. But, in the first place, it seems to be impossible to solve it, since F^k and $G_s{}^k$ are not known to us and, second, this problem is a very large one. In order to overcome these difficulties it is advisable to employ the software of the submodels which provides us with optimal values of inner variables for fixed and perhaps nonoptimal values of common variables.

Let $X^{*k}(V)$ be a solution of problem (1), subject to the vector of common variables being fixed. Substituting this solution for each of the subproblems (1) to replace X^k in problem (2), we get a new problem

(3)   miminize with respect to V

$$\sum_{k=1}^{k=N} \lambda_k F^k(X^{*k}(V), V)$$

subject to

$V \in D$ and

$R_s(V) \geq 0, s = \overline{1, M},$

where D is the domain of definition of functions $X^{*k}(V)$.

Two explanations are necessary for the statement of problem (3) because it is the basic problem in our consideration. First, the domain of definition D must be taken into account, as the subproblems (1) do not have a feasible solution for any vector of common variables. Second, inner constraints $G_s^k \geq 0$ of problem (1) are omitted in problem (3) because they are satisfied by $X^{*k}(V)$ by definition. We shall further call problem (3) a master problem.

By some natural assumptions on conditions of the subproblems, we can find the optimal values of X^k as $X^{*k}(V^*)$, where V^* is the solution of the master problem (3). This gives us the possibility of finding solutions X^{*k} of the submodels independently after solving problem (3), which is a more preferable problem than problem (2) because of its lesser dimension. But from another point of view, there are two difficulties peculiar to the master problem: To solve problem (3) we have to know the domain of the definition D and the explicit form of the functions $X^{*k}(V)$. It is unlikely that this sort of data will be found in most practical cases, and we have to find an indirect way of solving the master problem.

To surmount the difficulties mentioned above, we can take into account the fact that any numerical algorithm for solving a mathematical programming problem needs only some numerical data associated with the current approximation of the solution, but not the explicit form of the condition of the problem. In other words, for operating these algorithms we must be able to calculate only some numerical characteristics of functions $X^{*k}(V)$, such as their values and maybe their derivatives, at some point V.

As to the domain of definition D, we can avoid the necessity of explicitly building this set by using special procedures checking the existence of $X^{*k}(V)$ at any given point V, or by such kinds of algorithms which give the pseudo-solution of the problem when it has no feasible points. There-

fore, the scheme of solving the whole problem can be formulated as follows.

For some current point V in the space of common variables, we find all the data needed for solving the master problem. It is likely that all the subproblems have to be solved for this fixed V. We then change the values of the common variables according to the procedure of minimizing the common criterion of the whole problem. Repeating these two steps, we eventually obtain the optimal value of V. It is necessary to emphasize that this scheme is also considered as a variant of a general decomposition approach described by W. Orchard-Hays (1968).

GENERAL DESCRIPTION OF THE APPROACH

There are many works in which master problems are used in different decompositional schemes (see, for example, Fiacco & McCormick, 1968; Geoffrion, 1970). In these works problems (1) and (3) are considered directly and that is the reason why we can not apply the standard algorithms of smooth optimization to solve problem (3). The main difficulty preventing this is that functions $X^{*k}(V)$ are not differentiable with respect to V. This forces us to use special delicate methods of analyzing their properties if only problems (1) and (3) are considered in the form given in the previous section, and suggests the idea of using nondifferentiable procedures of optimization to solve problem (3) (Lemarechal, 1978).

But there is a way in which we can make computer linkage of different submodels on the basis of smooth algorithms. The idea consists of a preliminary transformation of the problems (1) and (3), providing them with some desirable properties. This transformation is proposed according to the sequential unconstrained minimization techniques (SUMT), sometimes called the penalty functions method as well.

This method (its smooth exterior point version) (Fiacco & McCormick, 1968) consists of unconstrained minimizations of some auxiliary function associated with the mathematical programming problem to be solved.

Let $E^k(X^k, V, T)$ be this auxiliary func-

tion for the kth problem (1), and $\hat{X}^k(V, T)$ be an extremal point of this function. Then, under some natural assumptions, the following relation between $\hat{X}^k(V, T)$ and $X^{*k}(V)$ takes place

$$(4) \quad \lim_{T \to +0} \hat{X}^k(V, T) = X^{*k}(V),$$

where T is a positive fixed parameter defining the degree of penalty for violations of constraints. This means that the extremal point of this auxiliary function is the solution of the problem (1), with perhaps some small error.

In the approach under consideration there are two reasons why it is convenient to use the exterior point version of the SUMT. First, auxiliary functions for problem (1) will always have an extremal point independent of whether the problem has a feasible solution or not. Second, the smooth version of the SUMT gives us the possibility to find all necessary data associated with $\hat{X}^k(V, T)$ by using a well known implicit function theorem (if, of course, all required derivatives exist).

Let us choose the auxiliary function E^k in the following form,

$$(5) \quad E^k(X^k, V, T) = \lambda_K F^k(X^k, V) + \sum_{s=1}^{s=m^k} P(G_s{}^k(X^k, V), T),$$

where the used penalty function $P(A, T)$ is defined and has continuous partial derivatives of the second order for any $T > 0$ and any A, and it satisfies the following relation as well:

$$(6) \quad \lim_{T \to +0} P(A, T) = \begin{cases} 0, \text{ for any } A > 0 \\ +\infty, \text{ otherwise.} \end{cases}$$

An auxiliary function associated with the master problem (3) can be chosen as

$$(7) \quad E(\hat{X}(V, T), V, T)$$
$$= \sum_{k=1}^{k=N} \lambda_k F^k(\hat{X}^k(V, T), V)$$
$$+ \sum_{s=1}^{s=M} P(R_s(V), T)$$
$$+ \sum_{k=1}^{k=N} \sum_{s=1}^{s=m^k} P(G_s{}^k(\hat{X}^k(V, T), V), T).$$

The double sum in this formula represents the penalty term for violations of the domain of definition D.

After obvious transformations we see that

$$(8) \quad E(\hat{X}(V, T), V, T) = W(V, T)$$
$$+ \sum_{k=1}^{k=N} E^k(\hat{X}^k(V, T), V, T)$$

where

$$W(V, T) = \sum_{s=1}^{s=M} P(R_s(V), T).$$

Expression (8) is of great importance as it presents the auxiliary function (7) as a sum of the auxiliary functions associated with the problems (1) and a function W given in the explicit form.

Let $\hat{V}(T)$ be an extremal point of problem (7); then approximate values of the inner variables can be given as $\hat{X}^k(\hat{V}, T)$.

Our first problem is to give the procedure of finding $\hat{V}(T)$ and, second, to consider the problem of evaluating V^* and X^{*k}. The problem of accuracy will be discussed later, and the main attention will now be paid to describing the data needed for solving the master problem (3) by minimizing the auxiliary function (7).

By virtue of assumptions stated above, any standard scheme of unconstrained optimization may be used for finding V. As a rule, these schemes consist of building a sequence of points in the space of the common variables $V^{(i)}$, which converges to V and is defined by the following recurrent equation,

$$(9) \quad V^{(i+1)} = V^{(i)} + sZ^{(i)}, \qquad i = 0, 1, 2, \cdots,$$

where $Z^{(i)}$ is a direction of minimizing Eq. (7), and s is an appropriate stepsize along this direction. Hence, there are two problems to be solved: how to find $Z^{(i)}$ and how to evaluate s.

In the first place, we shall consider the problem of building the direction of minimization for the auxiliary function (7). Doing this, we have to know the value, the gradient, and perhaps the hessian matrix of the function to be minimized. We shall consider the case when all these data are needed in the chosen scheme of optimization.

Let GRAD and ∇_x denote conventional gradient operators in the spaces of common

and inner variables, respectively. In the same way, HESSIAN and $\nabla_x{}^2$ will be the hessian matrix with respect to common and inner variables.

It is very important to notice that all derivatives with respect to V have to take into account both explicit and implicit dependence of the function to be differentiated on common variables. By the chain rule

$$\text{GRAD } E = E_v{}' + \sum_{k=1}^{k=N} H_{xv}^k \nabla_x E^k,$$

where $E_v{}'$ is a vector of the partial derivatives of E with respect to V, and H_{xv}^k is the conventional matrix of sensitivity of X^k by V.

Analogously,

$$\begin{aligned}
\text{HESSIAN } E = {} & E_v'' + \sum_{k=1}^{k=N} H_{xv}^k E_{xv}^{k\prime\prime} \\
& + \sum_{k=1}^{k=N} H_{xvv}^k \nabla_x E^k \\
& + \sum_{k=1}^{k=N} H_{xv}^k (\nabla_x{}^2 E^k (H_{xv}^k)^t \\
& + E_{xv}''^k),
\end{aligned}$$

where E_v'' is the matrix of second partial derivatives of E with respect to V, and

$E_{xv}''^k$ is $L \times n^k$–dimensional matrix of partial derivatives of E^k with respect to X^k and V and, finally,

H_{xvv}^k is the matrix of sensitivity of the second order.

Both these formulas are valid for any X^k and V, but $\hat{X}^k(V, T)$ are the minimum points of the auxiliary functions E^k. Therefore, by virtue of the fact that $\hat{X}^k$ satisfies the following equation:

$$(10) \qquad \nabla_x E^k(\hat{X}^k, V, T) = 0,$$

we simply have

$$\text{GRAD } E = E_v{}'.$$

Taking into consideration problem (10), and that the full derivative of this equation with respect to V is

$$(11) \qquad H_{xv}^k \nabla_x{}^2 E^k + E_{xv}''^k = 0,$$

we find

$$(12) \qquad \text{HESSIAN } E = E_v'' + \sum_{k=1}^{k=N} H_{xv}^k E_{xv}''^k.$$

To evaluate the quantity of information which needs to be transmitted from each of the submodels to the master problem, rewrite the expressions for GRAD and HESSIAN in new form

$$\text{GRAD } E = W_v{}' + \sum_{k=1}^{k=N} E_v^{k\prime}$$

and

$$\text{HESSIAN } E = W_v'' + \sum_{k=1}^{k=N} (E_v''^k + H_{xv}^k E_{xv}''^k),$$

by substituting Eq. (8).

It is easy to verify that the matrices $E_v''^k + H_{xv}^k E_{xv}''^k$ are symmetrical. Actually, if we multiply both sides of Eq. (11) by H_{xv}^k, we get

$$H_{xv}^k E_{xv}''^k = -H_{xv}^k \nabla_x{}^2 E^k (H_{xv}^k)^t,$$

and, by virtue of the symmetry of $\nabla_x{}^2 E^k$ and $E_v''^k$, we find the desirable result.

Therefore, the vector $E_v{}^k$, and the upper right half of the symmetrical matrix $E_v''^k + H_{xv}^k E_{xv}''^k$ are those which are to be calculated and transmitted by the submodels to the higher level of the whole system. Besides, there is no necessity of operating with any details of inner structure of the submodels to build the recurrent term of minimizing a sequence at the level of the master problem.

Let the number of components for the vector of common variables be used for joining the kth submodel and equal l^k. Then, considering that n^k is the dimension of this very submodel, we can show the calculation procedure of the matrix $E_v''^k + H_{xv}^k E_{xv}''^k$ in Fig. 1. It means that using the considered approach is only worthwhile if the dimensions of the subproblems are much greater than the numbers of components of V, belonging to the same subproblems, i.e.,

$$n^k \gg l^k.$$

It seems that this inequality may take place for many practical problems and, hence, the approach given above can be successfully used. To complete the general description of this method, it is necessary to note that the procedure of choosing the length of the step along the direction of minimization in the space of the common variables s may be done according to any standard scheme of one-dimension optimization or searching. Some foundations of the method are given in the appendix.
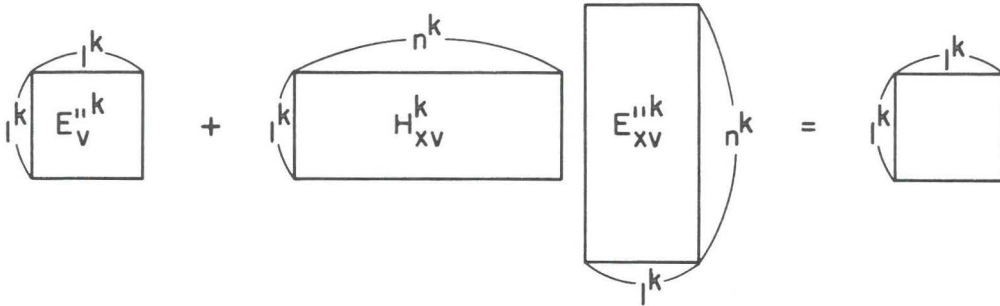
$$E''^k_v + H^k_{xv} \, E''^k_{xv} = $$

Fig 1.

PRACTICAL APPLICATIONS OF THE APPROACH

The "bottom-up" approach was used to model regional development. This method assumes that plans for regional economic growth can be based primarily on regional factors (i.e., available resources, regional demand, etc.), with only a minimal use of external information.

The starting point is the analysis of the regional specialization problem. At the second stage, intraregional location problems are solved, followed by analysis of labor and financial balance problems at the third stage. Finally, problems connected with environmental quality control and the provision of settlements and services are considered. It is assumed that the system of marginal costs for commodities produced and resources used and the data for determining regional flows of in- and out-migration are known.

Such a scheme allows the gain from industrial and agricultural activity to be maximized after a balance between different types of resources (including external investments) and production has been achieved. Three types of resources are specified in the analysis: capital investment, labor, and water.

Each block within the scheme is sufficiently detailed to describe practical sectoral problems. However, generally speaking, to make the system workable an appropriate level of aggregation must be chosen. The scheme is flexible and can include the following modifications:

The policy maker is able to vary the shares of the production and service sectors.

The objective function coefficients can be weighted in accordance with the priorities of the policy maker.

Resource consumption constraints can be introduced for some sectors.

Additional goods can be produced in different sectors.

The structure of the scheme and the coordination procedure may be changed to correspond to the set of problems characteristic of the region under analysis. Thus, the modules in the scheme represent a generally applicable description of a particular sector of the regional economy.

Completed modules

Work on the development of a set of widely applicable models of the most important sectors of a regional economy was initiated within the Regional Development Task at the International Institute for Applied Systems Analysis (IIASA) in 1977. However, scholars from other areas at IIASA and external institutes have also participated in this activity. Five models have already been completed:

Generalized regional agriculture model (GRAM);
Regional water supply model;
Migration model;
Population growth model; and
Generalized industrial model.

The models are supplied with the necessary computer software packages and are united to form a system. This system has a flexible structure. Therefore, according to the regional problems to be solved, some modules may be excluded and others added to the system.

Generalized regional agriculture models (GRAM). This model has been developed to analyze intraregional agriculture problems. It deals with:

Regional agricultural specialization;
Different types of production (crop, livestock, etc.) in disaggregated form;
Land-use problems, with reference to irrigation, drainage, etc.;
Choice of animal-feed compositions (protein, rough and green forage, etc.);
Choice of crop-rotation conditions;
Availability of regional supplies of labor, capital investment, fertilizers, water, etc.

A wide range of regional characteristics are described in the model; for example, land-type, technologies in use, type of market. Constraints on land use, the forage balance, human consumption, and production are included. Each group of constraints contains several inequalities.

Regional industry model. The model developed at the Central Institute for Economics and Mathematics in Moscow was used as a prototype for the industry model developed for the model system. The model describes transportation of different products, relations between resources and final products, and the dependence of costs on the scale of production.

Population and migration models. The intraregional character of the analysis requires sequential analysis of the labor force to be carried out on the regional as well as on subregional levels. Therefore, the future population, as well as in- and out-migration for the region as a whole, should be calculated. In addition, calculation of the future population and labor force in the multisubregional system should be made.

The results of the migration model for the region as a whole could be plugged into the regional population model, which is used to forecast the total regional population. The rate of regional migration can change from year to year, depending on the results of the migration model runs. The age and sex structure of migrants is assumed to be the same as for the previously observed period.

Water supply model. The water supply model is based on the following assumptions and constraints:

1. Water requirements are distributed over time (by seasons) and space, and are predetermined according to the location of industrial and agricultural activities.
2. All water users consume water resources irreversibly.
3. Within-year regulation of water resources only is considered.
4. The time delays for water transit are not taken into account.

The main goal of the model is to meet water requirements for a given period with minimal costs. Water quality problems are not considered. This model contains equations describing the mass balance for every node and reservoir, and the upper and lower bounds for nodes, reservoirs, pumping stations, and canals. The objective function is to minimize the sum of reduced costs for construction and operation.

Model linkage

Each of the models in the system has a common environment, which is represented by common variables and common constraints. Linkage is achieved through information flows between the models and the environment, rather than by direct flows of information between the models. Each model is described by means of two sets of numbers: variables, which define its state, and parameters, which define its environment. The models can be solved when the parameter values are fixed for the whole system, i.e., when additional linkage software is provided to perform this function.

The main idea behind the linkage procedure consists in searching for those parameter values with which all the models (solved independently) will obtain optimal solutions.

Theoretically, the optimal state of the environment of the whole system can be found by leveling the values of the Lagrangian multipliers for resources common to the different models. In other words, we are attempting to use the dependence of these multipliers on the state of the environment to estimate the optimality of the current distribution of resources among regional subsystems. However, in practical terms this idea is inconvenient for several reasons,

all of which should be taken into account during computer analysis.

1. It may not be possible to define the dependence of the optimal state of the models on the common environment for all environmental states.

2. This dependence may not be a functional relationship for any state.

3. Even if dependence exists and is functional, it will not be differentiable for different environmental states.

From the statements given above, we conclude that it is impossible to use any classical scheme based on Taylor's expansion theory to analyze this dependence.

In the approach we have described, a special algorithm (based on the smooth version of the penalty function method) was used to convert the dependence of the model state on the environmental state into a new function that exists and is differentiable for all states of the environment. It is important to note that, for all states of the environment, the newly determined dependence will be close to the initial dependence. This newly determined dependence was analyzed using a version of the modified Newton algorithm, which displays sufficiently good convergence properties.

Test sample

To test the proposed system of models a special sample problem was prepared for the Silistra (Bulgaria) case study using real data. The region under analysis was divided into three subregions. In the sample, the following sectors were included: agriculture, industry, water supply, and labor force.

In contrast to the population and migration models, in which employment is dependent on capital and labor allocation, in the labor force model employment is only dependent on the capital investment directed to service sector. Several dozen calculations were performed to prove that the model systems can successfully complement the results of changes to the following:

Objective function coefficients of all optimization models included;
Matrix of constraints within each optimization model in the system; and
Parameters of nonoptimization migration models.

The main series of calculations was oriented towards obtaining a picture of the changes in regional activities under conditions of changing external capital investments and employment.

CONCLUSIONS AND SOME FURTHER DEVELOPMENTS OF THE APPROACH

The presented approach permits us to link different models into the whole system without explicitly building a large-scale mathematical programming model. This approach is opposed, in a methodological sense, to usual decomposition schemes, but is rather close to them from a mathematical viewpoint. The main advantage of the approach is that all submodels are considered as black boxes and, therefore, can be built by different groups of specialists on the base of different software and hardware.

The solution which may be found in this scheme is optimal in the sense of a new criterion which is a linear combination of the submodel's criteria with some nonnegative weight coefficients. It means that this approach can also be interpreted as a version of the multi-criteria optimization when a point of the Pareto set is a solution.

The approach allows different extensions and generalizations. It would be of interest to explore connections between linkage problems and multicriteria optimization because in both cases a man-machine procedure is involved. It is also interesting to apply this approach to analysis of dynamic multistage optimization problems considering each stage as some local static optimization problems which are to be linked when the whole planning horizon is considered.

REFERENCES

Bagrinovskii, K. A. *Foundations of planning decision coordinations.* Moscow: Nauka, 1977.

Fiacco, A. V., & McCormick, G. P., *Nonlinear programming: sequential unconstrained minimization techniques.* New York: John Wiley, 1968.

Geoffrion, A. M. Primal resource—Directive approaches for optimizing nonlinear decomposable systems. *Operations Research*, 1970, 18, 375–403.

Lemarechal, C. *Nonsmooth optimization and descent methods.* RR-78-4, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1978.

Orchard-Hays, W. *Advanced linear programming*

computing techniques. New York: McGraw-Hill, 1968.

Propoi, A. I. Problems of dynamic linear programming. RM-76-78, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1976.

Umnov, A. E., The interative linear extrapolation in the penalty functions method. Journal of Computational Mathematics and Mathematical Physics, 1974, 6.

APPENDIX

Foundation of the approach

In this appendix we shall study the conditions under which the considered scheme can be used. Let functions F^k, G_s^k, and R_s be smooth enough, and problem (2) have an isolated local solution. In other words, let functions F^k, G_s^k, and R_s have continuous partial derivatives of the second order and let there be a system of points V^* and $X^{*k}(k = \overline{1, N})$ so that

(13) $\quad G_s^k(X^{*k}, V^*) \geq 0, s = \overline{1, m^k}$;

$\quad\quad k = \overline{1, N}, R_s(V^*) \geq 0, s = \overline{1, M}$,

and, there is a system of nonnegative numbers P_s^k and q_s satisfying the following relations:

(14) $\quad p_s^k G_s^k(X^{*k}, V^*) = 0$ and p_s^k

$\quad\quad > 0$ if and only if $G_s^k(X^{*k}, V^*)$

$\quad\quad = 0, q_s R_s(V^*) = 0$, and q_s

$\quad\quad > 0$; if and only if $R_s(V^*)$

$\quad\quad = 0$ for all s and k.

Let U be the usual Lagrange function associated with the problem (2), i.e.,

(15) $\quad U = \sum_{k=1}^{k=N} (\lambda_k F^k(X^k, V)$

$\quad\quad - \sum_{s=1}^{s=m^k} p_s^k G_s^k(X^k, V))$

$\quad\quad - \sum_{s=1}^{s=M} q_s R_s(V)$,

and at the point V^*, X^{*k} we have $\nabla_x U = 0$ and $\nabla_v U = 0$.

For any nonzero vectors Δv and $\Delta x^k (k = 1, N)$ such that

$(\Delta v)^t \nabla_v G_s^k + \sum_{k=1}^{k=N} (\Delta x^k)^t \nabla_x G_s^k$

$\quad\quad\quad\quad = 0$, if $p_s^k > 0$

and

$(\Delta v)^t \nabla_v R_s = 0$, if $q_s > 0$ at the same point, the following inequality takes place

(16) $\quad (\Delta v)^t \nabla_v^2 U \Delta v$

$\quad\quad + 2 \sum_{k=1}^{k=N} (\Delta v)^t \nabla_{xv}^2 U x^k$

$\quad\quad + \sum_{k=1}^{k=N} (\Delta x^k)^t \nabla_x^2 U \Delta x^k > 0$,

where

$\nabla_v^2 U$ is the hessian matrix of U with respect to V,

$\nabla_x^2 U$ is the hessian matrix of U with respect to X^k,

and

$\nabla_{xv}^2 U$ is the matrix, elements of which equal to $\partial^2 U/\partial x_k \partial v_i$.

Then, by virtue of Theorem 4 (Fiacco & McCormick, 1968), the point V^* and X^{*k} $(k = \overline{1, N})$ is an isolated local solution of the problem (2).

Under assumptions given above the following theorems will be valid.

THEOREM 1. X^{*k} is an isolated local solution of the problem (1) for fixed $V = V^*$.

PROOF. We have to show that all conditions analogous to (13–16) are valid for the problem (1) at the point X^k.

At first, we have by virtue of problem (13),

$$G_s^k(X^{*k}, V^*) \geq 0, s = \overline{1, m^k}.$$

It is possible to use the numbers p_s^k defined in Eq. (14) as Lagrange multipliers associated with problem (1), then

$$p_s^k G_s^k(X^{*k}, V^*) = 0$$

and

$p_s^k > 0$ if and only if $G_s^k(X^{*k}, V^*)$

$\quad\quad\quad\quad = 0$, for all s, k.

By virtue of the separability of U with respect to X^k we have

$$\nabla_x U^k = \nabla_x U = 0,$$

where

$$U^k = \lambda_k F^k(X^k, V) - \sum_{s=1}^{s=m^k} p_s^k G_s^k(X^k, V).$$

Let Δv and all $\Delta x^1, \Delta x^2, \cdots, \Delta x^N$ be equal to the zero vector except Δx^k, and for any

nonzero Δx^k from Eq. (16) we get that by $(\Delta x^k)'\nabla_x G_s{}^k = 0$ when $p_s{}^k > 0$, the following inequality is valid

$$(\Delta x^k)'\nabla_x{}^2 U^k \Delta x^k = (\Delta x^k)'\nabla_x{}^2 U \Delta x^k > 0,$$

and then all sufficient conditions of the optimality of X^*k are proved.

By virtue of the assumptions given above and a new assumption that all gradients of active constraints at X^{*k} are linearly independent, we find from Theorem 6 (Fiacco & McCormick, 1968) that functions $X^{**k}(V)$ exist within nonempty vicinity of V^* and have at this point partial derivatives. It is necessary to notice that the existence of the matrix of sensitivity H_{xv}^k does not ensure the differentiability of $X^{k*}(V)$ at V^*.

Now we are able to prove:

THEOREM 2. The point V^* is an isolated local solution of the problem (3).

PROOF. By inequality (13) and Theorem 1 we have

$$G_s{}^k(X^{*k}(V^*), V^*) = G_s{}^k(X^{*k}, V^*) \geq 0,$$
$$s = \overline{1, m^k}; \quad k = \overline{1, N}$$

and

$$R_s(V^*) \geq 0, \quad s = \overline{1, M}.$$

Further, $p_s{}^k G_s{}^k(X^{*k}(V^*), V^*) = 0$, $p_s{}^k > 0$ if and only if $G_s{}^k(X^{*k}(V^*), V^*) = 0$ and $q_s R(V^*) = 0$, $q_s > 0$ if and only if $R_s(V^*) = 0$.

This means that nonnegative numbers $p_s{}^k$ and q_s can be used as Lagrange multipliers associated with the problem (3). Let

$$\bar{U} = \sum_{k=1}^{k=N} (\lambda_k F^k(X^{*k}(V), V),$$
$$- \sum_{s=1}^{s=m^k} p_s{}^k G_s{}^k(X^{*k}(V), V))$$
$$- \sum_{s=1}^{s=m} q_s R_s(V)$$

be the Lagrange function for problem (3). By the chain rule and Theorem 1 we have

$$\nabla_v \bar{U} = \nabla_v U + \sum_{k=1}^{k=N} H_{xv}^k \nabla_x U,$$

but granting problem (15), we find $\nabla_v U = 0$ at the point V^*.

Finally, let Δv be any vector satisfying the following relations

$$(\Delta v)'\nabla_v G_s{}^k(X^{*k}(V^*), V^*)$$
$$= 0 \quad \text{if only} \quad p_s{}^k > 0,$$

and

$$(\Delta v)'\nabla_v R_s(V^*) = 0 \quad \text{if only} \quad q_s > 0$$

at the point V^*.

Then, by virtue of the chain rule:

$$(\Delta v)'\nabla_v{}^2 \bar{U}\Delta v = (\Delta v)'\nabla_v{}^2 U \Delta v$$
$$+ 2 \sum_{k=1}^{k=N} (\Delta v)'\nabla_{xv}^2 U H_{xv}^k \Delta v$$
$$+ \sum_{k=1}^{k=N} (H_{xv}^k)'(\Delta v)'\nabla_x{}^2 U H_{xv}^k \Delta v.$$

Denoting $\Delta x^k = H_{xv}^k \Delta v$, we get

$$(\Delta v)'\nabla_v{}^2 \bar{U}\Delta v = (\Delta v)'\nabla_v{}^2 U \Delta v$$
$$+ 2 \sum_{k=1}^{k=N} (\Delta v)'\nabla_{xv}^2 U \Delta x^k$$
$$+ \sum_{k=1}^{k=N}(\Delta x^k)'\nabla_x{}^2 U \Delta x^k.$$

The new vector of local variations Δx^k satisfies the following relations

$$(\Delta x^k)'\nabla_x G_s{}^k(X^k, V) = 0, \quad \text{if only} \quad p_s{}^k > 0.$$

Really,

$$(\Delta x^k)'\nabla_x G_s{}^k = (\Delta v)'(H_{xv}^k)'\nabla_x G_s{}^k$$
$$= (\Delta v)'\nabla_v G_s{}^k = 0,$$

by virtue of our assumption.

Hence, $(\Delta v)'\nabla_v{}^2 \bar{U}\Delta v > 0$ and V^* is an isolated local solution of the master problem (3).

To finish the foundation of the approach we have to consider the properties of the penalty function $P(A, T)$. Except for the conditions stated above, this function will satisfy the following relations at any point from its domain of definition

$$\partial P/\partial A < 0 \quad \text{and} \quad \partial^2 P/\partial A2 > 0.$$

Then Theorem 10 (Fiacco & McCormick, 1968) holds and we have

$$(17) \qquad \lim_{T \to +0} \hat{V}(T) = V^*.$$

The possibility of using some of classical optimization procedures for minimizing Eq. (7) arises from Theorem 2 and the assumptions about the existence of continuous partial derivatives of the second order for F^k, $G_s{}^k$, and R_s.

The problem of accuracy

Since the smooth version of the exterior point unconstrained minimization techniques gives us only the approximate solution of the problem to be solved, we have

to consider the problem or find the exact solution or, at least, reduce the error of the approximation.

There are two aspects of the problem of accuracy in the given approach. First, we have no troubles because the solutions of problem (1) are approximate during the minimizational process of the auxiliary function (7), as this process is an iterative one. Second, we have to study the problem of approximation at the final point of the algorithm used. The first aspect is not significant, but we shall consider the second.

The simplest way to evaluate the exact solution of the master problem is to use the relation (17) and standard Taylor approximation of the function $V(T)$.

Granting that

$$\hat{V}(T + \Delta T) = \hat{V}(T) + \Delta T \hat{V}'_T + o(T),$$

where

$$\lim_{\Delta T \to 0} \frac{o(\Delta T)}{\Delta T} = 0.$$

Going over to the limit when $\Delta T \to -T$, we find

$$(18) \qquad V^* = \hat{V}(T) - T\hat{V}'_T + o(T).$$

This means that to eliminate the linear part of the error we have to find the derivative $\hat{V}'_T$.

Before going into the details of this procedure it is necessary to notice that the assumptions being made in the previous section guarantee the existence of the trajectory of local minima of the SUMT and the validity of Eq. (18). Moreover, Umnov (1974) shows that if we choose the penalty function $P(A, T)$ as a function of the single argument A/T, then $\hat{V}'_T$ will have a bounded limit value by $T \to +0$.

Taking into consideration that the function $\hat{V}(T)$ is implicitly defined by the equation

$$(19) \qquad \nabla_v E(\hat{X}(\hat{V}, T), \hat{V}, T) = 0,$$

and by virtue of the implicit function theorem we get

$$\hat{V}'_T = (\text{HESSIAN } E)^{-1} \mathscr{E}''_{vT},$$

where HESSIAN E is given by Eq. (12) and $\mathscr{E}''_{vT}$ is the derivative of the left part of Eq. (19) with respect to T.

As far as there are dependencies of $\Delta_v E$ on T both in explicit and implicit ways we have

$$\mathscr{E}''_{vT} = E''_{vT}$$
$$+ \sum_{k=1}^{k=N} \frac{\partial \hat{X}^{kt}}{\partial T} E''^k_{xv}$$
$$+ \sum_{k=1}^{k=N} H^k_{xvv} \nabla_x E^k$$
$$+ \sum_{k=1}^{k=N} H^k_{xv} \left(\Delta_x{}^2 E^k \frac{\partial X^k}{\partial T} + E''^k_{xT} \right),$$

where

E''_{vT} is a vector of partial derivatives of E with respect to V and T,

E''^k_{xT} are vectors of partial derivatives of E^k with respect to X^k and T.

This formula can be rewritten as

$$\mathscr{E}''_{vT} = E''_{vT}$$
$$+ \sum_{k=1}^{k=N} H^k_{xv} E''^k_{xT}$$
$$+ \sum_{k=1}^{k=N} H^k_{xvv} \nabla_x E^k$$
$$+ \sum_{k=1}^{k=N} \frac{\partial \hat{X}^{kt}}{\partial T} (\Delta_x{}^2 E^k (H^k_{xv})^t + E''^k_{xv}),$$

but, by virtue of Eqs. (10) and (11), we finally can find

$$\mathscr{E}''_{vT} = E''_{vT} + \sum_{k=1}^{k=N} H^k_{xv} E''^k_{xv}.$$

It is followed by the fact that $\mathscr{E}''_{vT}$ can be calculated separately by different submodels. In fact, we have

$$\mathscr{E}''_{vT} = W''_{vT} + \sum_{k=1}^{k=N} (E''^k_{vT} + H^k_{xv} E''^k_{xT}).$$

Now we have to consider the problem of finding X^{*k}. Since this point is a limit for $\hat{X}^k(\hat{V}, T)$ when $T \to +0$, we can use the Taylor approximation again.

$$\hat{X}^k(\hat{V} + \Delta V, T + \Delta T)$$
$$= \hat{X}^k(\hat{V}, T) + \frac{\partial X^k}{\partial T} \Delta T$$
$$+ (H^k_{xv})^t \Delta V + o(\Delta V, \Delta T),$$

where $\partial \hat{X}^k / \partial T$ can be found by means of the implicit function theorem from the Eq. (10).

Taking into consideration that $\Delta V = \hat{V}'_T \Delta T$ and going over to the limit when ΔT

$\rightarrow -T$, we get

$$(20) X^{*k} = \hat{X}^k(\hat{V}, T) - T \frac{\partial \hat{X}^k}{\partial T}$$

$$- T H_{xv}^k \hat{V}_{T}' + o(T).$$

This also means that the correction of the approximate solutions can be made independently by different submodels, but only after finding $\hat{V}_T'$ in the master problem.

It can happen sometimes that one step of the procedure (18–20) does not provide us with the desirable level of accuracy. It is possible in this case to repeat all these calculations. The conditions of the convergence of the process which can be called iterative linear extrapolation are given by Umnov (1974).