Title	Two-level planning for multi-objective systems
Sub Title	
Author	志水, 清孝(Shimizu, Kiyotaka)
Publisher	慶応義塾大学工学部
Publication year	1977
Jtitle	Keio engineering reports Vol.30, No.8 (1977. 3) ,p.69- 83
JaLC DOI	
Abstract	We study a two-level system having N local systems in the lower level subordinate to a central system in the higher one, such that both central and local systems have decision-making units. The central system is a coordinating agency and the local ones are semi-autonomous operating devisions. The basic principle of planning for this organization is that the central system allocates resources so as to optimize its own objective, while the local ones optimize their own objectives using the given resources. A local objective function, fn, is a function of the lower level decision variable vector x=(x1,・・・, xN) and the higher level one a=(a1,・・・, aN), where an is a resource vector allocated to the local system n. Since the functions ■ are mutually independent, the lower level composes a multi-objective system, in which the lower level decision-makers minimize a vector objective function f =(f1,・・・,fN) with respect to x in cooperation with each other. Thus, the lower level generates a set of noninferior (i.e. Pareto optimal) solutions ■(a) being parametric with respect to a. The central decision-maker, then, chooses the optimal resource allocation aº and the best noninferior solution ■º corresponding to aº from among a set of ■(a). The above problem becomes a decentralized two-level optimization, when the local system contains only its own variables (xn, an). Several theorems and iterative algorithms for the formulated problems are obtained by use of mathematical programming techniques.
Notes	
Genre	Departmental Bulletin Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO50001004-00300008-0069

KEIO ENGINEERING REPORTS
VOL. 30, NO. 8, pp. 69-83, 1977

TWO-LEVEL PLANNING FOR MULTI-OBJECTIVE SYSTEMS

KIYOTAKA SHIMIZU

Dept. of Instrumentation Engineering Keio University, Yokohama 223, Japan

(*Received January 6, 1977*)

ABSTRACT

We study a two-level system having N local systems in the lower level subordinate to a central system in the higher one, such that both central and local systems have decision-making units. The central system is a coordinating agency and the local ones are semi-autonomous operating devisions. The basic principle of planning for this organization is that the central system allocates resources so as to optimize its own objective, while the local ones optimize their own objectives using the given resources.

A local objective function, f_n, is a function of the lower level decision variable vector $\boldsymbol{x}=(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_N)$ and the higher level one $\boldsymbol{a}=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_N)$, where $\boldsymbol{a}_n$ is a resource vector allocated to the local system n. Since the functions $\{f_n\}_{n=1}^N$ are mutually independent, the lower level composes a multi-objective system, in which the lower level decision-makers minimize a vector objective function $\boldsymbol{f}=(f_1,\cdots,f_N)$ with respect to $\boldsymbol{x}$ in cooperation with each other. Thus, the lower level generates a set of noninferior (i.e. Pareto optimal) solutions $\hat{\boldsymbol{x}}(\boldsymbol{a})$ being parametric with respect to $\boldsymbol{a}$.

The central decision-maker, then, chooses the optimal resource allocation $\boldsymbol{a}^0$ and the best noninferior solution $\hat{\boldsymbol{x}}^0$ corresponding to $\boldsymbol{a}^0$ from among a set of $\hat{\boldsymbol{x}}(\boldsymbol{a})$.

The above problem becomes a decentralized two-level optimization, when the local system contains only its own variables $(\boldsymbol{x}_n, \boldsymbol{a}_n)$.

Several theorems and iterative algorithms for the formulated problems are obtained by use of mathematical programming techniques.

I. Introduction

In this paper, we are concerned with a class of organizations composed of a coordinating central system and plural semi-autonomous local systems.

In economic activities, for example, the central system distributes its available resources to the local ones and the local systems perform production activities utilizing the given resources. The central system determines resource allocation so as to optimize its objective consisting of values of products of the local systems and cost of the resources. That is, the central system pursues profit of the central level and governs the local systems through the way of resource allocations, while the local ones establish their autonomies by optimizing their own objectives.

Hence, we consider the hierarchical decision system in which the higher level determines values of decision variables peculiar to the central system (for instance, parameters such as the resources allocated to the local systems and/or policy coefficients defining objectives and constraints of the lower level) and the lower level determines values of decision variables peculiar to the local systems (for instance, process variables such as products, quality etc.). The basic principle of planning for this organization is that the central system allocates scarce resources so as to optimize its own objective, while the local ones optimize their own objectives under the restriction of the given resources. It is noted that a two-level system to be studied is such that both central and local systems possess independent decision-making units.

Among studies to solve mathematical programs by two-level decomposition techniques by resource allocation exist KORNAI's paper for LP (KORNAI, 1965), SILVERMAN's for convex program (SILVERMAN, 1972) and GEOFFRION's for decentralized two-level planning of a more general form of a resource allocation type (GEOFFRION, 1972). There are some others on optimization of hierarchical decentralized systems by parametric programming approach (SHIMIZU, 1969, 1974, 1976-a, b).

II. Two-level Optimization for Multi-objective Systems

In this section, we formulate optimization problems of the hierarchical system in which autonomy of the local systems is explicitly considered within the restriction imposed by the central one. The lower level consists of N local systems each of which possesses its own decision variable vector $\boldsymbol{x}_n$, an objective function f_n and a constraint vector function $\boldsymbol{g}_n \leq \mathbf{0}$. But there exist mutual interactions among the local systems, because f_n and/or $\boldsymbol{g}_n$ have not only $\boldsymbol{x}_n$ but $\boldsymbol{x}=(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N)$ as its argument generally. Therefore, the lower level is a group of local systems having mutual interaction and different goals.

The lower level decides the best value of the vector $\boldsymbol{x}$, which is dependent on a parameter vector $\boldsymbol{\alpha}$ assigned by the higher level. But the lower level composes a multi-objective system where local decision-makers try to minimize a vector objective function $\boldsymbol{f}=(f_1, \cdots, f_N)$ with respect to $\boldsymbol{x}$. Accordingly, one can not obtain a unique solution even if $\boldsymbol{\alpha}$ is fixed.

For the lower level problem, the following two cases are considered: (i) The local systems cooperate together and carry out vector minimization with respect to $\boldsymbol{x}=(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N)$. They generate a set of noninferior (Pareto optimal) solutions $\hat{\boldsymbol{x}}(\boldsymbol{\alpha})$ being parametric with respect to $\boldsymbol{\alpha}$. (ii) The local systems perform noncooperative game and generate a set of Nash solutions $\boldsymbol{x}^N(\boldsymbol{\alpha})$.

In this paper, we restrict ourselves to the first case.

Let $\boldsymbol{\alpha}_n$ be a resource vector allocated to the local system n, which is regarded as a parameter vector in the lower level problem. The central decision-maker determines an optimal resource allocation $\boldsymbol{\alpha}^0=(\boldsymbol{\alpha}_1^0,\cdots,\boldsymbol{\alpha}_N^0)$, based on the central objective function Φ. At the same time, however, it must choose the best noninferior solution $\hat{\boldsymbol{x}}^0$ corresponding to $\boldsymbol{\alpha}^0$ from among the set of noninferior solutions.

The two-level planning is formally stated as follows:

$$\min_{\alpha,\hat{x}(\alpha)} \Phi(\boldsymbol{\alpha},\hat{\boldsymbol{x}}(\boldsymbol{\alpha})) \tag{1.a}$$

$$\text{subj. to } \boldsymbol{G}(\boldsymbol{\alpha},\hat{\boldsymbol{x}}(\boldsymbol{\alpha}))\leqq \boldsymbol{0} \tag{1.b}$$

$$\begin{pmatrix} f_1(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}),\boldsymbol{\alpha}) \\ \vdots \\ f_N(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}),\boldsymbol{\alpha}) \end{pmatrix} = \min_x \begin{pmatrix} f_1(\boldsymbol{x},\boldsymbol{\alpha}) \\ \vdots \\ f_N(\boldsymbol{x},\boldsymbol{\alpha}) \end{pmatrix} \tag{1.c}$$

$$\text{subj. to } \begin{pmatrix} \boldsymbol{g}_1(\boldsymbol{x},\boldsymbol{\alpha})\leqq 0 \\ \vdots \\ \boldsymbol{g}_N(\boldsymbol{x},\boldsymbol{\alpha})\leqq 0 \end{pmatrix} \tag{1.d}$$

where $\hat{\boldsymbol{x}}(\boldsymbol{\alpha})$ represents a parametric noninferior solution. The eqn. (1.b) is a constraint imposed upon the parameter vector $\boldsymbol{\alpha}$ (resources and/or policy coefficients) and the decision of the lower level $\hat{\boldsymbol{x}}(\boldsymbol{\alpha})$. We assume existence of $(\boldsymbol{\alpha},\hat{\boldsymbol{x}}(\boldsymbol{\alpha}))$ satisfying (1.b). This equation is expressed separately as $\boldsymbol{G}_1(\boldsymbol{\alpha})\leqq \boldsymbol{0}$ and $\boldsymbol{G}_2(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}))\leqq \boldsymbol{0}$ in most cases.

After all, the problem (1) is to find the optimal resource allocation $\boldsymbol{\alpha}^0$ and the best noninferior solution $\hat{\boldsymbol{x}}^0=\hat{\boldsymbol{x}}_{\text{best}}(\boldsymbol{\alpha}^0)$ corresponding to $\boldsymbol{\alpha}^0$ so as to minimize Φ under the constraints (1.b) (1.c) (1.d). Therefore, Φ must provide both roles of determining the optimal resource allocations and at the same time choosing the best solution from among the noninferior solution set. Note that a value of Φ varies with which $\hat{\boldsymbol{x}}(\boldsymbol{\alpha})$ is choosen even if $\boldsymbol{\alpha}$ is fixed.

The problem (1) is regarded as the hierarchical multi-objective system in which there exist N mutually independent local objective functions $\{f_n\}$ and a super objective function Φ. Further, it is such a problem that parameterized constrained optimization problems of the lower level are contained in a part of the constraints of the higher level. Therefore, it is not of a type of usual mathematical programming problem.

Multilevel system is a parametric approach in principle in the sense that one defines a family of decision problems whose solution is attempted in a sequential manner from the higher to the lower level and the solution of the original problem is achieved when all sequential problems are solved. If a problem of certain level is solved, then its solution fixes some parameters in the problem of the subsequent level and one can attempt to solve the completely specified problem. In this way, we have a series of parametric constrained optimization problems, each of which is a simplified decision problem. The problem (1) can be solved in principle with such a parametric apprcach.

For simplicity, let $\boldsymbol{f}=(f_1,\cdots,f_N)$ and $\boldsymbol{g}=(\boldsymbol{g}_1,\cdots,\boldsymbol{g}_N)$. The problem (1.c) (1.d) is then written as follows.

$$\boldsymbol{f}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}),\boldsymbol{\alpha})=\min_x \boldsymbol{f}(\boldsymbol{x},\boldsymbol{\alpha}) \tag{1.c$'$}$$

$$\text{subj to } \boldsymbol{g}(\boldsymbol{x},\boldsymbol{\alpha}) \leqq \boldsymbol{0} \tag{1.d)$'$}$$

Now let us consider the following scalarization problem in relation to the problem (1.c) (1.d).

$$\min_{x} \ \boldsymbol{w}^T\boldsymbol{f}(\boldsymbol{x},\boldsymbol{\alpha})$$

$$\text{subj. to } \boldsymbol{g}(\boldsymbol{x},\boldsymbol{\alpha}) \leqq \boldsymbol{0} \tag{2}$$

$$\text{where } \boldsymbol{\alpha} \text{ is fixed and } \boldsymbol{w} \in \Omega = \left\{\boldsymbol{w} | \boldsymbol{w} \geqq \boldsymbol{0}, \sum_{n=1}^{N} w_n = 1\right\}$$

Then the following theorems are well known.

Theorem 1. If $\hat{\boldsymbol{x}}(\boldsymbol{\alpha})$ solves the scalarization problem (2) for some $\boldsymbol{w} > \boldsymbol{0}$, then $\hat{\boldsymbol{x}}(\boldsymbol{\alpha})$ is a noninferior solution to the problem (1.c) (1.d). Further, when $\boldsymbol{f}$ and $\boldsymbol{g}$ are convex functions in $\boldsymbol{x}$, if $\hat{\boldsymbol{x}}(\boldsymbol{\alpha})$ is a noninferior solution to the problem (1.c) (1.d), then $\hat{\boldsymbol{x}}(\boldsymbol{\alpha})$ solves the problem (2) for some $\boldsymbol{w} \in \Omega$.

Theorem 2. Assume that $\boldsymbol{f}$ and $\boldsymbol{g}$ are strictly convex and convex functions in $\boldsymbol{x}$, respectively. Then, in order that $\hat{\boldsymbol{x}}(\boldsymbol{\alpha})$ be a noninferior solution to the problem (1.c) (1.d), it is necessary and sufficient that $\hat{\boldsymbol{x}}(\boldsymbol{\alpha})$ is a solution to the scalarization problem (2) for some $\boldsymbol{w} \in \Omega$.

By use of Theorem 2 with the assumption on $\boldsymbol{f}, \boldsymbol{g}$ the problem (1) may be equivalently represented as follows.

$$\min_{\alpha, w \in \Omega} \ \Phi(\boldsymbol{\alpha}, \boldsymbol{x}^0(\boldsymbol{\alpha},\boldsymbol{w})) \tag{3.a}$$

$$\text{subj. to } \boldsymbol{G}(\boldsymbol{\alpha}, \boldsymbol{x}^0(\boldsymbol{\alpha},\boldsymbol{w})) \leqq \boldsymbol{0} \tag{3.b}$$

$$\boldsymbol{w}^T\boldsymbol{f}(\boldsymbol{x}^0(\boldsymbol{\alpha},\boldsymbol{w}),\boldsymbol{\alpha}) = \min_{x} \ \boldsymbol{w}^T\boldsymbol{f}(\boldsymbol{x},\boldsymbol{\alpha}) \tag{3.c}$$

$$\text{subj. to } \boldsymbol{g}(\boldsymbol{x},\boldsymbol{\alpha}) \leqq \boldsymbol{0} \tag{3.d}$$

where $\boldsymbol{x}^0(\boldsymbol{\alpha},\boldsymbol{w})$ is a parametric optimal solution to the scalarization problem given $\boldsymbol{\alpha}$ and $\boldsymbol{w}$.

A difficulty of the above approach is due to the fact that it is almost impossibe to obtain a parametric optimal solution to (3) in explicit form. Therefore, we need an iterative algorithm to search the better $(\boldsymbol{\alpha},\boldsymbol{w})$ based on the information of the lower level solutions under the fixed $(\boldsymbol{\alpha},\boldsymbol{w})$. Namely, the iterative procedure consists of the coordinator asking the local systems what would happen if the parameter vector were set at $(\boldsymbol{\alpha},\boldsymbol{w})$ to which the local systems respond by giving some local information concerning the optimal solution of the corresponding local problems. The center (coordinator) then uses this information in a prescribed manner to determine a revised trial setting for $(\boldsymbol{\alpha},\boldsymbol{w})$.

When a size of the problem is relatively small, direct search such as Constrained Simplex Method (Box method) seems to be useful. The method may be applied also when Φ is not stated explicitly. We will discuss on the Constrained Simplex Method in Section 4.

It is assumed throughout that all functions are differentiable.

Suppose that $\boldsymbol{f}$ and $\boldsymbol{g}$ are strictly convex and convex functions in $\boldsymbol{x}$, respectively. Then, it is necessary and sufficient for the problem (1.c) (1.d) that there exist $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$ satisfying the following equation.

$$\begin{aligned} &\boldsymbol{\mu}^T \nabla_x \boldsymbol{f}(\hat{\boldsymbol{x}}, \boldsymbol{\alpha}) + \boldsymbol{\lambda}^T \nabla_x \boldsymbol{g}(\hat{\boldsymbol{x}}, \boldsymbol{\alpha}) = \boldsymbol{0} \\ &\boldsymbol{g}(\hat{\boldsymbol{x}}, \boldsymbol{\alpha}) \leqq \boldsymbol{0},\ \boldsymbol{\lambda}^T \boldsymbol{g}(\hat{\boldsymbol{x}}, \boldsymbol{\alpha}) = 0,\ \boldsymbol{\lambda} \geqq \boldsymbol{0} \\ &\boldsymbol{\mu} \geqq \boldsymbol{0} (\boldsymbol{\mu} \neq \boldsymbol{0}) \end{aligned} \tag{4}$$

where $\boldsymbol{\mu}, \boldsymbol{\lambda}$ are Lagrange multipliers (refer to SHIMIZU, 1976-c). If we consider $\boldsymbol{\mu} \equiv \boldsymbol{w}$, the eqn. (4) is a necessary and sufficient condition for the scalarization problem of (1.c) (1.d) with respect to $\boldsymbol{x}$.

It is evident that the problem (1) is to solve $(\boldsymbol{\alpha}^0, \hat{\boldsymbol{x}}^0)$ such that $\Phi(\boldsymbol{\alpha}, \hat{\boldsymbol{x}}) \rightarrow \min$ under the constraints of (4) and (1.b).

Next we consider the case when Φ is a function of $\boldsymbol{f}$ in the problem (1).

$$\min_{\alpha, \hat{x}(\alpha)} \Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}), \boldsymbol{\alpha}), \boldsymbol{\alpha}) \tag{5.a}$$

$$\text{subj. to } \boldsymbol{G}(\boldsymbol{\alpha}, \hat{\boldsymbol{x}}(\boldsymbol{\alpha})) \leqq \boldsymbol{0} \tag{5.b}$$

$$\boldsymbol{f}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}), \boldsymbol{\alpha}) = \min_{x} \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\alpha}) \tag{5.c}$$

$$\text{subj. to } \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{\alpha}) \leqq \boldsymbol{0} \tag{5.d}$$

where $\boldsymbol{G}(\boldsymbol{\alpha}, \hat{\boldsymbol{x}}(\boldsymbol{\alpha})) = \begin{pmatrix} \boldsymbol{G}_1(\boldsymbol{\alpha}) \\ \boldsymbol{f}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}), \boldsymbol{\alpha}) - \boldsymbol{c} \end{pmatrix}$ and $\boldsymbol{c}$ is a constant of satisfaction level. At the same time let us consider the following problem in connection with the problem (5).

$$\min_{\alpha, x} \Phi(\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\alpha}), \boldsymbol{\alpha}) \tag{6.a}$$

$$\text{subj. to } \boldsymbol{G}(\boldsymbol{\alpha}, \boldsymbol{x}) \leqq \boldsymbol{0} \tag{6.b}$$

$$\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{\alpha}) \leqq \boldsymbol{0} \tag{6.c}$$

where $\boldsymbol{G}(\boldsymbol{\alpha}, \boldsymbol{x}) = \begin{pmatrix} \boldsymbol{G}_1(\boldsymbol{\alpha}) \\ \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\alpha}) - \boldsymbol{c} \end{pmatrix}$ and $\boldsymbol{c}$ is a constant.

We have then the following properties.

Theorem 3. Assume that Φ is a strictly increasing function of $\boldsymbol{f}$. Then a solution to the problem (6), $(\boldsymbol{\alpha}^0, \boldsymbol{x}^0)$, solves the problem (5).

Proof First, we show that a solution $(\boldsymbol{\alpha}^0, \boldsymbol{x}^0)$ of the problem (6) is a feasible solution to the problem (5.c) (5.d) satisfying (5.b) as $\boldsymbol{\alpha}$ is fixed at $\boldsymbol{\alpha}^0$, then there exists $\check{\boldsymbol{x}}$ such that

$$\boldsymbol{f}(\check{\boldsymbol{x}}, \boldsymbol{\alpha}^0) \leqq \boldsymbol{f}(\boldsymbol{x}^0, \boldsymbol{\alpha}^0) \text{ and } \boldsymbol{f}(\check{\boldsymbol{x}}, \boldsymbol{\alpha}^0) \neq \boldsymbol{f}(\boldsymbol{x}^0, \boldsymbol{\alpha}^0) \tag{7.a}$$

$$\begin{aligned} &\boldsymbol{g}(\check{\boldsymbol{x}}, \boldsymbol{\alpha}^0) \leqq \boldsymbol{0} \\ &\boldsymbol{G}(\boldsymbol{\alpha}^0, \check{\boldsymbol{x}}) \leqq \boldsymbol{0} \end{aligned} \tag{7.b}$$

From the assumption and (7.a)

$$\Phi(\boldsymbol{f}(\check{\boldsymbol{x}}, \boldsymbol{\alpha}^0), \boldsymbol{\alpha}^0) < \Phi(\boldsymbol{f}(\boldsymbol{x}^0, \boldsymbol{\alpha}^0), \boldsymbol{\alpha}^0) \tag{8}$$

From (7.b) (7.c) (8) this contradicts to that $(\boldsymbol{x}^0, \boldsymbol{\alpha}^0)$ is a solution to (6). Therefore, since $(\boldsymbol{x}^0, \boldsymbol{\alpha}^0)$ is a noninferior solution of the lower level problem and satisfies $\boldsymbol{G}(\boldsymbol{\alpha}^0, \boldsymbol{x}^0) \leqq \boldsymbol{0}$, it is a feasible solution to the problem (5). Since a set of feasible solutions to the problem (5) is contained in that of the problem (6), it is evident that $(\boldsymbol{\alpha}^0, \boldsymbol{x}^0)$ minimizes the objective function of (5.a). This completes the proof.

We will consider the following vector minimization problem in relation to the problem (5):

$$\min_{\alpha, x} \begin{pmatrix} \Phi(\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\alpha}), \boldsymbol{\alpha}) \\ \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\alpha}) \end{pmatrix} \tag{9.a}$$

$$\text{subj. to } \boldsymbol{G}(\boldsymbol{\alpha}, \boldsymbol{x}) \leqq \boldsymbol{0} \tag{9.b}$$

$$\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{\alpha}) \leqq \boldsymbol{0} \tag{9.c}$$

Then we have:

Theorem 4. Assume that Φ is a strictly increasing function of $\boldsymbol{f}$. Then a solution $(\boldsymbol{\alpha}^0, \hat{\boldsymbol{x}}^0)$ of the problem (5) solves the problem (6) and further it is a weak noninferior solution of the problem (9).

Proof Assume that a solution $(\boldsymbol{\alpha}^0, \hat{\boldsymbol{x}}^0)$ of the problem (5) is not a solution to the problem (6). Then, there exists $(\check{\boldsymbol{\alpha}}, \check{\boldsymbol{x}})$ satisfying

$$\Phi(\boldsymbol{f}(\check{\boldsymbol{x}}, \check{\boldsymbol{\alpha}}), \check{\boldsymbol{\alpha}}) < \Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^0, \boldsymbol{\alpha}^0), \boldsymbol{\alpha}^0) \tag{10}$$

and

$$\boldsymbol{G}(\check{\boldsymbol{\alpha}}, \check{\boldsymbol{x}}) \leqq \boldsymbol{0}, \qquad \boldsymbol{g}(\check{\boldsymbol{x}}, \check{\boldsymbol{\alpha}}) \leqq \boldsymbol{0} \tag{11}$$

(i) In case that $\check{\boldsymbol{x}}$ is a noninferior solution of the lower level problem (5.c) (5.d) for fixed $\boldsymbol{\alpha} = \check{\boldsymbol{\alpha}}$, the eqns. (10) (11) contradict to the fact that $(\boldsymbol{\alpha}^0, \hat{\boldsymbol{x}}^0)$ solves the problem (5).

(ii) In case that $\check{\boldsymbol{x}}$ is not a noninferior solution to the lower level problem (5.c) (5.d), there exists a noninferior solution $\boldsymbol{x}^*$ to the lower level problem (5.c) (5.d) such that

$$\boldsymbol{f}(\boldsymbol{x}^*, \check{\boldsymbol{\alpha}}) \leqq \boldsymbol{f}(\check{\boldsymbol{x}}, \check{\boldsymbol{\alpha}}) \text{ and } \boldsymbol{f}(\boldsymbol{x}^*, \check{\boldsymbol{\alpha}}) \neq \boldsymbol{f}(\check{\boldsymbol{x}}, \check{\boldsymbol{\alpha}}) \tag{12.a}$$

$$\boldsymbol{g}(\boldsymbol{x}^*, \check{\boldsymbol{\alpha}}) \leqq \boldsymbol{0} \tag{12.b}$$

$$\boldsymbol{G}(\check{\boldsymbol{\alpha}}, \boldsymbol{x}^*) \leqq \boldsymbol{0} \tag{12.c}$$

against $(\check{\boldsymbol{\alpha}}, \check{\boldsymbol{x}})$ satisfying $\boldsymbol{G}(\check{\boldsymbol{\alpha}}, \check{\boldsymbol{x}}) \leqq \boldsymbol{0}, \boldsymbol{g}(\check{\boldsymbol{x}}, \check{\boldsymbol{\alpha}}) \leqq \boldsymbol{0}$.

Thus, by the assumption of the theorem

$$\Phi(\boldsymbol{f}(\boldsymbol{x}^*, \check{\boldsymbol{\alpha}}), \check{\boldsymbol{\alpha}}) < \Phi(\boldsymbol{f}(\check{\boldsymbol{x}}, \check{\boldsymbol{\alpha}}), \check{\boldsymbol{\alpha}}) \tag{13}$$

Hence, from (10) and (13)

$$\Phi(\boldsymbol{f}(\boldsymbol{x}^*, \check{\boldsymbol{\alpha}}), \check{\boldsymbol{\alpha}}) < \Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^0, \boldsymbol{\alpha}^0), \boldsymbol{\alpha}^0) \tag{14}$$

$\boldsymbol{x}^*$ is a noninferior solution to the lower level problem for $\boldsymbol{\alpha}=\check{\boldsymbol{\alpha}}$ and satisfies (5.b) from (12.c). Therefore, the relation (14) contradicts to the fact that $(\boldsymbol{\alpha}^0, \hat{\boldsymbol{x}}^0)$ solves the problem (5). Hence, $(\boldsymbol{\alpha}^0, \hat{\boldsymbol{x}}^0)$ solves the problem (6).

Since the problem (6) is a scalarization problem of (9) for $\boldsymbol{w}=(1,0,0,\cdots,0)$, a solution $(\boldsymbol{\alpha}^0, \hat{\boldsymbol{x}}^0)$ to (6) is a weak noninferior solution to the problem (12).

Theorem 5. Suppose that $\boldsymbol{f}$ be a function of only $\boldsymbol{x}$, i.e. $\boldsymbol{f}(\boldsymbol{x})$, in the problem (5). As sufficient condition that $(\boldsymbol{\alpha}^0, \hat{\boldsymbol{x}}^0=\hat{\boldsymbol{x}}_{\text{best}}(\boldsymbol{\alpha}^0))$ be an optimal solution to the problem (5) is that $(\boldsymbol{\alpha}^0, \hat{\boldsymbol{x}}^0)$ satisfies the following:

(i) $\boldsymbol{f}, \boldsymbol{g}$ and $\boldsymbol{G}$, are convex with respect to $(\boldsymbol{x}, \boldsymbol{\alpha})$.

(ii) Φ is convex with respect to $\boldsymbol{f}$ and $\boldsymbol{\alpha}$.

(iii) $\nabla_f \Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^0), \boldsymbol{\alpha}^0) > 0$.

(iv) there exist $\boldsymbol{\mu}, \boldsymbol{\lambda}$ and $\boldsymbol{\gamma}$ such that

$$\begin{aligned}
&\boldsymbol{\mu}^T \nabla_x \boldsymbol{f}(\hat{\boldsymbol{x}}^0) + \boldsymbol{\lambda}^T \nabla_x \boldsymbol{g}(\hat{\boldsymbol{x}}^0, \boldsymbol{\alpha}^0) + \boldsymbol{\gamma}^T \nabla_x \boldsymbol{G}(\boldsymbol{\alpha}^0, \hat{\boldsymbol{x}}^0) = \boldsymbol{0}\\
&\nabla_\alpha \Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^0), \boldsymbol{\alpha}^0) + \boldsymbol{\lambda}^T \nabla_\alpha \boldsymbol{g}(\hat{\boldsymbol{x}}^0, \boldsymbol{\alpha}^0) + \boldsymbol{\gamma}^T \nabla_\alpha \boldsymbol{G}(\boldsymbol{\alpha}^0, \hat{\boldsymbol{x}}^0) = \boldsymbol{0}\\
&\boldsymbol{g}(\hat{\boldsymbol{x}}^0, \boldsymbol{\alpha}^0) \leqq \boldsymbol{0},\ \boldsymbol{\lambda}^T \boldsymbol{g}(\hat{\boldsymbol{x}}^0, \boldsymbol{\alpha}^0) = 0,\ \boldsymbol{\lambda} \geqq \boldsymbol{0}\\
&\boldsymbol{G}(\boldsymbol{\alpha}^0, \hat{\boldsymbol{x}}^0) \leqq \boldsymbol{0},\ \boldsymbol{\gamma}^T \boldsymbol{G}(\boldsymbol{\alpha}^0, \hat{\boldsymbol{x}}^0) = 0,\ \boldsymbol{\gamma} \geqq \boldsymbol{0}\\
&\boldsymbol{\mu} > \boldsymbol{0}
\end{aligned}$$

where $\boldsymbol{\mu}$ satisfies $\boldsymbol{\mu} = \nabla_f \Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^0), \boldsymbol{\alpha}^0)$.

Proof Let $I_a = \{i \mid g_i(\hat{\boldsymbol{x}}^0, \boldsymbol{\alpha}^0) = 0\}$ and $I_{\bar{a}} = \{i \mid g_i(\hat{\boldsymbol{x}}^0, \boldsymbol{\alpha}^0) < 0\}$ From the third equation of (iv), $\lambda_i g_i(\hat{\boldsymbol{x}}^0, \boldsymbol{\alpha}^0) = 0$, $\forall i \in I = I_a \cap I_{\bar{a}}$. Hence $\boldsymbol{\lambda}_{I_{\bar{a}}} = \boldsymbol{0}$.

It holds for an arbitrary feasible $(\boldsymbol{x}, \boldsymbol{\alpha})$ that $\boldsymbol{g}_{Ia}(\boldsymbol{x}, \boldsymbol{\alpha}) \leqq \boldsymbol{0} = \boldsymbol{g}_{I\bar{a}}(\hat{\boldsymbol{x}}^0, \boldsymbol{\alpha}^0)$. By this and convexity of $\boldsymbol{g}$

$$\begin{aligned}
&\nabla_x \boldsymbol{g}_{Ia}(\hat{\boldsymbol{x}}^0, \boldsymbol{\alpha}^0)(\boldsymbol{x} - \hat{\boldsymbol{x}}^0) + \nabla_\alpha \boldsymbol{g}_{Ia}(\hat{\boldsymbol{x}}^0, \boldsymbol{\alpha}^0)(\boldsymbol{\alpha} - \boldsymbol{\alpha}^0)\\
&\quad \leqq \boldsymbol{g}_{Ia}(\boldsymbol{x}, \boldsymbol{\alpha}) - \boldsymbol{g}_{Ia}(\hat{\boldsymbol{x}}^0, \boldsymbol{\alpha}^0) \leqq \boldsymbol{0}
\end{aligned}$$

Further, by the fact that $\boldsymbol{\lambda}_{Ia} \geqq \boldsymbol{0}$ and $\boldsymbol{\lambda}_{I\bar{a}} = \boldsymbol{0}$,

$$\boldsymbol{\lambda}^T (\nabla_x \boldsymbol{g}(\hat{\boldsymbol{x}}^0, \boldsymbol{\alpha}^0)(\boldsymbol{x} - \hat{\boldsymbol{x}}^0) + \nabla_\alpha \boldsymbol{g}(\hat{\boldsymbol{x}}^0, \boldsymbol{\alpha}^0)(\boldsymbol{\alpha} - \boldsymbol{\alpha}^0) \leqq \boldsymbol{0}$$

This relation holds for any arbitrary feasible solution of (5), $(\boldsymbol{\alpha}, \hat{\boldsymbol{x}}(\boldsymbol{\alpha}))$, thus

$$\boldsymbol{\lambda}^T \{\nabla_x \boldsymbol{g}(\hat{\boldsymbol{x}}^0, \boldsymbol{\alpha}^0)(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}) - \hat{\boldsymbol{x}}^0) + \nabla_\alpha \boldsymbol{g}(\hat{\boldsymbol{x}}^0, \boldsymbol{\alpha}^0)(\boldsymbol{\alpha} - \boldsymbol{\alpha}^0) \leqq \boldsymbol{0}\} \tag{15}$$

In similar way for an arbitrary feasible $(\boldsymbol{\alpha}, \hat{\boldsymbol{x}}(\boldsymbol{\alpha}))$

$$\boldsymbol{\gamma}^T (\nabla_x \boldsymbol{G}(\boldsymbol{\alpha}^0, \hat{\boldsymbol{x}}^0)(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}) - \hat{\boldsymbol{x}}^0) + \nabla_\alpha \boldsymbol{G}(\boldsymbol{\alpha}^0, \hat{\boldsymbol{x}}^0)(\boldsymbol{\alpha} - \boldsymbol{\alpha}^0)) \leqq \boldsymbol{0} \tag{16}$$

From the eqn. (15) plus the eqn. (16)

$$\begin{aligned}
&\{\boldsymbol{\lambda}^T \nabla_x \boldsymbol{g}(\hat{\boldsymbol{x}}^0, \boldsymbol{\alpha}^0) + \boldsymbol{\gamma}^T \nabla_x \boldsymbol{G}(\boldsymbol{\alpha}^0, \hat{\boldsymbol{x}}^0)\}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}) - \hat{\boldsymbol{x}}^0)\\
&\quad + \{\boldsymbol{\lambda}^T \nabla_\alpha \boldsymbol{g}(\hat{\boldsymbol{x}}^0, \boldsymbol{\alpha}^0) + \boldsymbol{\gamma}^T \nabla_\alpha \boldsymbol{G}(\boldsymbol{\alpha}^0, \hat{\boldsymbol{x}}^0)\}(\boldsymbol{\alpha} - \boldsymbol{\alpha}^0) \leqq \boldsymbol{0}
\end{aligned} \tag{17}$$

By the way, as $\hat{\boldsymbol{x}}^0$ satisfies the condition (iv), with the assumption on $\boldsymbol{G}, \hat{\boldsymbol{x}}^0$ is a noninferior solution to the lower level problem for $\boldsymbol{\alpha}=\boldsymbol{\alpha}^0$. Substituting the first and second relations of (iv) into (17) we have

$$\boldsymbol{\mu}^T\nabla_x \boldsymbol{f}(\hat{\boldsymbol{x}}^0)(\hat{\boldsymbol{x}}(\boldsymbol{\alpha})-\hat{\boldsymbol{x}}^0)+\nabla_\alpha\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^0),\boldsymbol{\alpha}^0)(\boldsymbol{\alpha}-\boldsymbol{\alpha}^0)\geqq \boldsymbol{0} \tag{18}$$

By convexity of $\boldsymbol{f}$ and $\boldsymbol{\mu}>\boldsymbol{0}$, $\boldsymbol{\mu}^T\boldsymbol{f}(\boldsymbol{x})$ is convex. Hence

$$\boldsymbol{\mu}^T\boldsymbol{f}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}))-\boldsymbol{\mu}^T\boldsymbol{f}(\hat{\boldsymbol{x}}^0)\geqq \nabla_x\boldsymbol{\mu}^T\boldsymbol{f}(\hat{\boldsymbol{x}}^0)(\hat{\boldsymbol{x}}(\boldsymbol{\alpha})-\hat{\boldsymbol{x}}^0)$$

Since $\boldsymbol{\mu}=\nabla_f\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^0),\boldsymbol{\alpha}^0)$,

$$\nabla_f\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^0),\boldsymbol{\alpha}^0)\{\boldsymbol{f}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}))-\boldsymbol{f}(\hat{\boldsymbol{x}}^0)\}\geqq \nabla_x\boldsymbol{\mu}^T\boldsymbol{f}(\hat{\boldsymbol{x}}^0)(\hat{\boldsymbol{x}}(\boldsymbol{\alpha})-\hat{\boldsymbol{x}}^0) \tag{19}$$

Substitute (19) into (18),

$$\nabla_f\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^0),\boldsymbol{\alpha}^0)\{\boldsymbol{f}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}))-\boldsymbol{f}(\hat{\boldsymbol{x}}^0)\}+\nabla_\alpha\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^0),\boldsymbol{\alpha}^0)(\boldsymbol{\alpha}-\boldsymbol{\alpha}^0)\geqq\boldsymbol{0}$$

By convexity of Φ with respect to $\boldsymbol{f}$ and $\boldsymbol{\alpha}$,

$$\begin{aligned}&\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha})),\boldsymbol{\alpha})-\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^0),\boldsymbol{\alpha}^0)\\&\quad\geqq\nabla_f\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^0),\boldsymbol{\alpha}^0)\{\boldsymbol{f}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}))-\boldsymbol{f}(\hat{\boldsymbol{x}}^0)\}+\nabla_\alpha\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^0),\boldsymbol{\alpha}^0)(\boldsymbol{\alpha}-\boldsymbol{\alpha}^0)\geqq\boldsymbol{0}\end{aligned} \tag{20}$$

Therefore, for any feasible solution $(\boldsymbol{\alpha},\hat{\boldsymbol{x}}(\boldsymbol{\alpha}))$ of the problem (5), it holds that $\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^0),\boldsymbol{\alpha}^0)\leqq\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha})),\boldsymbol{\alpha})$. This proves that $(\hat{\boldsymbol{x}}^0,\boldsymbol{\alpha}^0)$ solves the problem (5).

Theorem 6. Suppose that Φ is given as $\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}),\boldsymbol{\alpha}))$ in the problem (5). Then, a sufficient condition that $(\boldsymbol{\alpha}^0,\hat{\boldsymbol{x}}^0)$ be an optimal solution to the problem (5) is that $(\boldsymbol{\alpha}^0,\hat{\boldsymbol{x}}^0)$ satisfies the following:

(i) $\boldsymbol{f},\boldsymbol{g}$ and $\boldsymbol{G}_1$ are convex with respect to $(\boldsymbol{x},\boldsymbol{\alpha})$.
(ii) Φ is convex with respect to $\boldsymbol{f}$.
(iii) $\nabla_f\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^0,\boldsymbol{\alpha}^0))>\boldsymbol{0}$.
(iv) there exist $\boldsymbol{\mu},\boldsymbol{\lambda}$ and $\boldsymbol{\gamma}$ such that

$$\begin{aligned}&\boldsymbol{\mu}^T\nabla_x\boldsymbol{f}(\hat{\boldsymbol{x}}^0,\boldsymbol{\alpha}^0)+\boldsymbol{\lambda}^T\nabla_x\boldsymbol{g}(\hat{\boldsymbol{x}}^0,\boldsymbol{\alpha}^0)+\boldsymbol{\gamma}^T\nabla_x\boldsymbol{G}(\boldsymbol{\alpha}^0,\hat{\boldsymbol{x}}^0)=\boldsymbol{0}\\&\boldsymbol{\mu}^T\nabla_\alpha\boldsymbol{f}(\hat{\boldsymbol{x}}^0,\boldsymbol{\alpha}^0)+\boldsymbol{\lambda}^T\nabla_\alpha\boldsymbol{g}(\hat{\boldsymbol{x}}^0,\boldsymbol{\alpha}^0)+\boldsymbol{\gamma}^T\nabla_\alpha\boldsymbol{G}(\boldsymbol{\alpha}^0,\hat{\boldsymbol{x}}^0)=\boldsymbol{0}\\&\boldsymbol{g}(\hat{\boldsymbol{x}}^0,\boldsymbol{\alpha}^0)\leqq\boldsymbol{0},\ \boldsymbol{\lambda}^T\boldsymbol{g}(\hat{\boldsymbol{x}}^0,\boldsymbol{\alpha}^0)=0,\ \boldsymbol{\lambda}\geqq\boldsymbol{0}\\&\boldsymbol{G}(\boldsymbol{\alpha}^0,\hat{\boldsymbol{x}}^0)\leqq\boldsymbol{0},\ \boldsymbol{\gamma}^T\boldsymbol{G}(\boldsymbol{\alpha}^0,\hat{\boldsymbol{x}}^0)=0,\ \boldsymbol{\gamma}\geqq\boldsymbol{0}\\&\boldsymbol{\mu}>\boldsymbol{0}\end{aligned}$$

where μ satisfies $\boldsymbol{\mu}=\nabla_f\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^0,\boldsymbol{\alpha}^0))$.

Proof The first part before the eqn. (17) is exactly the same as in Theorem 5.

Now, since $\hat{\boldsymbol{x}}^0$ satisfies (iv), with the assumption on $\boldsymbol{G}$, it is a noninferior solution of the lower level problem for $\boldsymbol{\alpha}=\boldsymbol{\alpha}^0$.

Substituting the first and second relations of (iv) into (17), we obtain

$$\boldsymbol{\mu}^T\nabla_x\boldsymbol{f}(\hat{\boldsymbol{x}}^0,\boldsymbol{\alpha}^0)(\hat{\boldsymbol{x}}(\boldsymbol{\alpha})-\hat{\boldsymbol{x}}^0)+\boldsymbol{\mu}^T\nabla_\alpha\boldsymbol{f}(\hat{\boldsymbol{x}}^0,\boldsymbol{\alpha}^0)(\boldsymbol{\alpha}-\boldsymbol{\alpha}^0)\geqq 0.$$

By convexity of $\boldsymbol{f}$ and $\boldsymbol{\mu}>\boldsymbol{0}$, $\boldsymbol{\mu}^T\boldsymbol{f}(\boldsymbol{x},\boldsymbol{\alpha})$ is convex in $(\boldsymbol{x},\boldsymbol{\alpha})$. Hence,

$$\begin{aligned}&\boldsymbol{\mu}^T\boldsymbol{f}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}),\boldsymbol{\alpha})-\boldsymbol{\mu}^T\boldsymbol{f}(\hat{\boldsymbol{x}}^0,\boldsymbol{\alpha}^0)\\&\quad\geqq\nabla_x\boldsymbol{\mu}^T\boldsymbol{f}(\hat{\boldsymbol{x}}^0,\boldsymbol{\alpha}^0)(\hat{\boldsymbol{x}}(\boldsymbol{\alpha})-\hat{\boldsymbol{x}}^0)+\nabla_\alpha\boldsymbol{\mu}^T\boldsymbol{f}(\hat{\boldsymbol{x}}^0,\boldsymbol{\alpha}^0)(\boldsymbol{\alpha}-\boldsymbol{\alpha}^0)\geqq\boldsymbol{0}\end{aligned}$$

By $\boldsymbol{\mu}=\nabla_f\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^0,\boldsymbol{\alpha}^0))$,

$$\nabla_f\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^0,\boldsymbol{\alpha}^0))\{\boldsymbol{f}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}),\boldsymbol{\alpha})-\boldsymbol{f}(\hat{\boldsymbol{x}}^0,\boldsymbol{\alpha}^0)\}\geqq 0.$$

By convexity of Φ with respect to $\boldsymbol{f}$

$$\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}),\boldsymbol{\alpha}))-\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^0,\boldsymbol{\alpha}^0))$$
$$\geqq \nabla_f\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^0,\boldsymbol{\alpha}^0))\{\boldsymbol{f}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}),\boldsymbol{\alpha})-\boldsymbol{f}(\hat{\boldsymbol{x}}^0,\boldsymbol{\alpha}^0)\}\geqq 0$$

Thus, for any feasible solution $(\boldsymbol{\alpha},\hat{\boldsymbol{x}}(\boldsymbol{\alpha}))$ of the problem (5) it holds that $\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}^0,\boldsymbol{\alpha}^0))\leqq\Phi(\boldsymbol{f}(\hat{\boldsymbol{x}}(\boldsymbol{\alpha}),\boldsymbol{\alpha}))$.

III. Hierarchical Optimization for Decentralized Systems

The two-level problem (1) becomes a hierarchical decentralized (semi-autonomous) system when an objective function f_n and a constraint $\boldsymbol{g}_n\leqq\boldsymbol{0}$ of each local system contain ony its own decision variable vector $\boldsymbol{x}_n$ and a parameter vector $\boldsymbol{\alpha}_n$ given from the higher level. Accordingly, the local systems are serated with respect to $\boldsymbol{x}$ each other. That is,

$$\min_{\alpha}\ \Phi(\boldsymbol{\alpha},\boldsymbol{x}^0(\boldsymbol{\alpha})) \tag{21.a}$$

$$\text{subj. to } \boldsymbol{G}(\boldsymbol{\alpha},\boldsymbol{x}^0(\boldsymbol{\alpha}))\leqq\boldsymbol{0} \tag{21.b}$$

$$f_n(\boldsymbol{x}_n^0(\boldsymbol{\alpha}_n),\boldsymbol{\alpha}_n)=\min_{x_n}\ f_n(\boldsymbol{x}_n,\boldsymbol{\alpha}_n) \tag{21.c}$$

$$\text{subj. to } \boldsymbol{g}_n(\boldsymbol{x}_n,\boldsymbol{\alpha}_n)\leqq\boldsymbol{0} \tag{21.d}$$

$$n=1,\cdots,N$$

where $\boldsymbol{\alpha}=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_N)$ and $\boldsymbol{x}^0(\boldsymbol{\alpha})=(\boldsymbol{x}_1^0(\boldsymbol{\alpha}_1),\cdots,\boldsymbol{x}_N^0(\boldsymbol{\alpha}_N))$. A vector $\boldsymbol{x}_n^0(\boldsymbol{\alpha}_n)$ means a usual parametric optimal solution. It is noted that the eqn. (21.a) achieves minimization with respect to only $\boldsymbol{\alpha}$ in contrast to the eqn. (1.a).

Hierarchical systems are characterized with concept of decomposition and coordination.

When the above problem is a typical resource allocation problem, it becomes as follows.

$$\min_{\alpha}\ \Phi(\boldsymbol{\alpha},\boldsymbol{x}^0(\boldsymbol{\alpha})) \tag{22.a}$$

$$\text{subj. to} \sum_{n=1}^{N}\boldsymbol{\alpha}_n\leqq\boldsymbol{b}\ (\boldsymbol{b}\text{: total resources}),\ \boldsymbol{G}_1(\boldsymbol{x}^0(\boldsymbol{\alpha}))\leqq\boldsymbol{0} \tag{22.b}$$

$$f_n(\boldsymbol{x}_n^0(\boldsymbol{\alpha}_n))=\min_{x_n}\ f_n(\boldsymbol{x}_n) \tag{22.c}$$

$$\text{subj. to } \boldsymbol{g}_n(\boldsymbol{x}_n)\leqq\boldsymbol{\alpha}_n,\ \boldsymbol{x}_n\varepsilon S_n \tag{22.d}$$

$$n=1,\cdots,N$$

The problems of the lower level consist of a set of N separated optimization subproblems, each of which is a usual parametric optimization problem with a scalar objective function. They are mutually interacting through (22.b). The first equation of (8.b) bounds a total amount of resources and the second constrains production activity of the lower level. The first equation of (8.d) represents an upper bound of the resources that is available to the local system n for production activity. The second is imposed by technologies of activity peculiar to the local system n.

Our aim is to develop algorithms to provide the optimal resource allocations and the problem (22) simply consists of the coordinating center and the decomposed (separated) local systems. But since the local systems are semi-automonous with its own objective, our problem is entirely different from the optimization for large-scale systems by the conventional decomposition principle for completely centralized planning.

Feasible Direction Method to Solve Problem (22)

Numerical method for the problem (22) is achieved by parametric approach in principle. A method using multiparametric LP was proposed in case of the lower level problem of linear programs (SHIMIZU, 1974). Generally speaking, however, it is very difficult to solve parametric constrained optimization problems. So we need consider to solve the resource allocation problem in the higher level by iteration without getting an explicit parametric solution in the lower level.

Feasible direction method seems to be useful for the coordinating center which updates the present allocations based on the solution of the local problems. In particular, when the central objective (22.a) is given as $\min_{\alpha} \Phi(\boldsymbol{f}(\boldsymbol{x}^0(\boldsymbol{\alpha}), \boldsymbol{\alpha}), \boldsymbol{\alpha})$, we can obtain numerical procedures using primal decomposition technique by resource allocation, one of which was derived by GEOFFRION (1972) elsewhere. But we present here our algorithm in a different manner.

Now let us consider the hierarchical decentralized system as follows.

$$\min_{\alpha} \Phi(\boldsymbol{f}(\boldsymbol{x}^0(\boldsymbol{\alpha})), \boldsymbol{\alpha}) \tag{23.a}$$

$$\text{subj. to} \sum_{n=1}^{N} \boldsymbol{\alpha}_n \leqq \boldsymbol{b} \tag{23.b}$$

$$f_n(\boldsymbol{x}_n^0(\boldsymbol{\alpha}_n)) = \min_{x_n} f_n(\boldsymbol{x}_n) \tag{23.c}$$

$$\text{subj. to } \boldsymbol{g}_n(\boldsymbol{x}_n) \leqq \boldsymbol{\alpha}_n, \boldsymbol{x}_n \varepsilon S_n \tag{23.d}$$

$$n = 1, \cdots, N$$

A common example of the eqn. (23) is $\Phi = F(\boldsymbol{f}(\boldsymbol{x}^0(\boldsymbol{\alpha})) + H(\boldsymbol{\alpha})$.

The above problem (23), to the higher level, may be viewed as one of optimally allocating the resource vector $\boldsymbol{b}$ to the local systems such that the central objective Φ is optimized. We attempt to solve it iteratively by choosing a feasible allocation testing it for optimality and improving it if it is not optimal. Then given the allocation $\boldsymbol{\alpha}_n$, the local system n must utilize it as well as possible, so

$$\text{Local Problem } P_n(\boldsymbol{\alpha}_n):\ \min_{x_n} f_n(\boldsymbol{x}_n) \tag{24}$$

$$\text{subj. to } \boldsymbol{g}_n(\boldsymbol{x}_n) \leqq \boldsymbol{\alpha}_n, \boldsymbol{\alpha}_n \varepsilon S_n$$

From the feasibility condition for the central problem

$$\sum_{n=1}^{N} \boldsymbol{\alpha}_n \leqq \boldsymbol{b} \tag{25}$$

$$\boldsymbol{\alpha}_n \varepsilon V_n = \{\boldsymbol{\alpha}_n | \text{there exists } \boldsymbol{x}_n \varepsilon S_n \text{ satisfying } \boldsymbol{g}_n(\boldsymbol{x}_n) \leqq \boldsymbol{\alpha}_n\} \tag{26}$$

We express the minimal objective value of $P_n(\boldsymbol{\alpha}_n)$ as

$$w_n(\boldsymbol{\alpha}_n) = \min_{x_n} \{f_n(\boldsymbol{x}_n) | \boldsymbol{g}_n(\boldsymbol{x}_n) \leqq \boldsymbol{\alpha}_n, \boldsymbol{x}_n \varepsilon S_n\} \tag{27}$$

Then the two-level decentralized optimization problem (23) is equivalent to the following master problem.

$$\min_{\{\alpha_n\}} \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha}) \tag{28.a}$$

$$\text{subj. to} \sum_{n=1}^{N} \boldsymbol{\alpha}_n \leqq \boldsymbol{b} \tag{28.b}$$

$$\boldsymbol{\alpha}_n \varepsilon V_n, n=1, \cdots, N \tag{28.c}$$

where $\boldsymbol{w}(\boldsymbol{\alpha}) = (w_1(\boldsymbol{\alpha}_1), \cdots, w_N(\boldsymbol{\alpha}_N))$

The equivalence between the master and original problems depends in no way upon convexity of their program. However, to make further theoretical progress, we impose the following conditions.

Assumption(i) S_n is compact and convex.

(ii) $F(\boldsymbol{f}(\boldsymbol{x}), \boldsymbol{\alpha})$ is convex and differentiable on $S_1 \times \cdots \times S_N$.

(iii) $f_n(\boldsymbol{x}_n), \boldsymbol{g}_n(\boldsymbol{x}_n)$ are convex and differentiable on S_n.

(iv) $F(\boldsymbol{f}(\boldsymbol{x}), \boldsymbol{\alpha})$ is convex and differentiable in $\boldsymbol{\alpha}$.

(v) the problem (23) is feasible.

Two useful sufficient conditions for the assumption (ii) to hold are:

(a) $F(\boldsymbol{f}(\boldsymbol{x}), \boldsymbol{\alpha})$ is convex and differentiable with respect to $\boldsymbol{f}$ and each $f_n(\boldsymbol{x}_n)$ is linear in $\boldsymbol{x}_n$.

(b) $F(\boldsymbol{f}(\boldsymbol{x}), \boldsymbol{\alpha})$ is monotone increasing convex and differentiable with respect to $\boldsymbol{f}$ and each $f_n(\boldsymbol{x}_n)$ is convex and differentiable in $\boldsymbol{x}_n$.

The convexity assumptions on $f_n, \boldsymbol{g}_n, S_n$ ensure that V_n is convex and that w_n is convex (with respect to $\boldsymbol{\alpha}_n$) over V_n, so the master problem is a convex program.

It is very hard to obtain an explicit expression of the objective function (28.a). Despite all idealized assumptions w_n may not be differentiable everywhere on V_n. This is the principal difficulty in solving the master problem. Fortunately, however, since each w_n is convex, it has directional derivatives at all points in int. V_n.

Feasible Direction Method using directional derivatives was first applied by GEOFFRION (1970) and SILVERMAN (1972) to separable mathematical programs. Although our problem is not quite the same that they studied, we can still apply the similar technique as far as calculation is concerned.

By solving a set of $P_n(\boldsymbol{\alpha}_n)$, we obtain value of directional derivatives of $\Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})$ at $\boldsymbol{\alpha}, D\Phi(\boldsymbol{\alpha}; \boldsymbol{s})$, and information of the minimal value of $\Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})$. Then a direction-finding problem is constructed which, by minimizing the directional derivative of the master objective function (at the allocation $\boldsymbol{\alpha}$) subject to feasibility restriction (25) (26), finds a usable feasible direction $\boldsymbol{s}$ in which the allocation $\boldsymbol{\alpha}$ can be improved. Such direction is a "locally best" usable feasible direction. If $\boldsymbol{\alpha}$ is not optimal, a step is taken in the direction $\boldsymbol{s}$, a new allocation $\boldsymbol{\alpha}$ is determined and the process is repeated.

In order that we may take a small step in any direction $\boldsymbol{s}_n$ from $\boldsymbol{\alpha}_n$ and keep $\boldsymbol{\alpha}_n + \beta \boldsymbol{s}_n$ in V_n, we assume that each vector $\boldsymbol{\alpha}_n$ is in the interior of V_n.

Now, a directional derivative* of Φ is given as follows.

Theorem 7. A directional derivative of $\Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})$ is given as

$$D\Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha}; \boldsymbol{s}) = \sum_{n=1}^{N} \left[\frac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial w_n} Dw_n(\boldsymbol{\alpha}_n; \boldsymbol{s}_n) + \frac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}_n} \boldsymbol{s}_n \right] \tag{29}$$

where $Dw_n(\boldsymbol{\alpha}_n; \boldsymbol{s}_n)$ is a directional derivative of $w_n(\boldsymbol{\alpha}_n)$.

Proof By definition of the directional derivative

$$D\Phi(\boldsymbol{\alpha}; \boldsymbol{s}) = \lim_{\beta \to 0+} \frac{\Phi(w(\boldsymbol{\alpha}+\beta \boldsymbol{s}), \boldsymbol{\alpha}+\beta \boldsymbol{s}) - \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\beta}$$

Let $\boldsymbol{w}(\boldsymbol{\alpha}+\beta\boldsymbol{s}) = \boldsymbol{w}(\boldsymbol{\alpha}) + \boldsymbol{h}$ and $\boldsymbol{\alpha}+\beta\boldsymbol{s} = \boldsymbol{\alpha}+\boldsymbol{k}$. Since Φ is differentiable with respect to $(\boldsymbol{w}, \boldsymbol{\alpha})$,

$$\Phi(\boldsymbol{w}(\boldsymbol{\alpha}+\beta\boldsymbol{s}), \boldsymbol{\alpha}+\beta\boldsymbol{s}) = \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha}) + \frac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial \boldsymbol{w}} \boldsymbol{h} + \frac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \beta \boldsymbol{s} + 0(\boldsymbol{h}, \beta\boldsymbol{s})$$

where

$$\frac{0(\boldsymbol{h}, \beta\boldsymbol{s})}{\|\boldsymbol{h}, \beta\boldsymbol{s}\|} \to 0, \qquad \text{as } \|\boldsymbol{h}, \beta\boldsymbol{s}\| \to 0.$$

Hence, it holds that

$$\begin{aligned} &\Phi(\boldsymbol{w}(\boldsymbol{\alpha}+\beta\boldsymbol{s}), \boldsymbol{\alpha}+\beta\boldsymbol{s}) - \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha}) \\ &= \frac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial \boldsymbol{w}} \boldsymbol{h} + \frac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \beta\boldsymbol{s} + 0(\boldsymbol{h}, \beta\boldsymbol{s}) \end{aligned}$$

Deviding both sides by β and achieving limit operation, we have

$$D\Phi(\boldsymbol{\alpha}; \boldsymbol{s}) = \lim_{\beta \to 0+} \frac{\dfrac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial \boldsymbol{w}} \boldsymbol{h} + \dfrac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \beta\boldsymbol{s} + 0(\boldsymbol{h}, \beta\boldsymbol{s})}{\beta}$$

where $\|\boldsymbol{h}, \beta\boldsymbol{s}\| \to 0$ as $\beta \to 0$, thus

$$\lim_{\beta \to 0+} \frac{0(\boldsymbol{h}, \beta\boldsymbol{s})}{\beta} = 0.$$

Thus,

$$\begin{aligned} D\Phi(\boldsymbol{\alpha}; \boldsymbol{s}) &= \lim_{\beta \to 0+} \frac{1}{\beta} \left[\frac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial \boldsymbol{w}} \{\boldsymbol{w}(\boldsymbol{\alpha}+\beta\boldsymbol{s}) - \boldsymbol{w}(\boldsymbol{\alpha})\} + \frac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \beta\boldsymbol{s} \right] \\ &= \sum_{n=1}^{N} \left[\frac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial w_n} \lim_{\beta \to 0+} \frac{1}{\beta} \{w_n(\boldsymbol{\alpha}_n + \beta\boldsymbol{s}_n) - w_n(\boldsymbol{\alpha}_n)\} + \frac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}_n} \boldsymbol{s}_n \right] \end{aligned}$$

---

* A directional derivative of f in $\boldsymbol{s}$ direction with respect to $\boldsymbol{x}$ is defined:

$$Df(\boldsymbol{x}; \boldsymbol{s}) = \lim_{\beta \to 0+} \frac{f(\boldsymbol{x}+\beta\boldsymbol{s}) - f(\boldsymbol{x})}{\beta}$$

$$= \sum_{n=1}^{N} \left[\frac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial w_n} Dw_n(\boldsymbol{\alpha}_n; \boldsymbol{s}_n) + \frac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}_n} \boldsymbol{s}_n \right]$$

This completes the proof.

By use of Theorem 7, the direction-finding problem to find a "locally best" usable feasible direction is given as

$$\min_{\{s_n\}} \sum_{n=1}^{N} \left[\frac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial w_n} Dw_n(\boldsymbol{\alpha}_n; \boldsymbol{s}_n) + \frac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}_n} \boldsymbol{s}_n \right]$$

$$\text{subj. to} \sum_{n=1}^{N} s_{ni} \leqq 0, \ i \varepsilon B \tag{30}$$

$$-1 \leqq s_{ni} \leqq 1, n=1, \cdots, N, i=1, \cdots, \dim \boldsymbol{\alpha}_n$$

where B is an index set of the binding constraints at $\boldsymbol{\alpha}$, *i. e.* $B=\left\{ i | b_i - \sum_{n=1}^{N} \alpha_{ni} = 0 \right\}$.

Then the following properties hold (see Chap. 9. of LASDON (1969)).

(A) Let $\boldsymbol{\alpha}$ be feasible for the master problem. If $\boldsymbol{s}=\boldsymbol{0}$ is optimal for the direction-finding problem, then $\boldsymbol{\alpha}$ solves the master problem.

(B) Let $\partial w_n(\boldsymbol{\alpha}_n)$ be a set of subgradients of w_n at $\boldsymbol{\alpha}_n$.

Then the directional derivative $Dw_n(\boldsymbol{\alpha}_n; \boldsymbol{s}_n)$ is expressed as

$$Dw_n(\boldsymbol{\alpha}_n; \boldsymbol{s}_n) = \max_{x^* \varepsilon \partial w_n(\alpha_n)} \boldsymbol{x}^{*T} \boldsymbol{s}_n = \max_{\lambda_n \varepsilon \Lambda_n} (-\boldsymbol{\lambda}_n^T \boldsymbol{s}_n)$$

where Λ_n is a set of all optimal Lagrange multiplier vectors for $P_n(\boldsymbol{\alpha}_n)$.

Let each set S_n be expressed as

$$S_n = \{ \boldsymbol{x}_n | \boldsymbol{q}_n(\boldsymbol{x}_n) \leqq \boldsymbol{0} \}$$

We assume that each component of $\boldsymbol{q}_n$ is convex and differentiable. Let $\boldsymbol{x}_n^0$ is an optimal solution to $P_n(\boldsymbol{\alpha}_n)$, and $\boldsymbol{\lambda}_n$ is the Lagrange multiplier vector associated with $(\boldsymbol{g}_n \leqq \boldsymbol{\alpha}_n, \boldsymbol{q}_n \leqq \boldsymbol{0})$. Then $\boldsymbol{\lambda}_n \varepsilon \Lambda_n$ if and only if $\boldsymbol{x}_n^0$ and some $\boldsymbol{\lambda}_n$ satisfy the Kuhn-Tucker conditions for $P_n(\boldsymbol{\alpha}_n)$. Accordingly, under the assumption that Φ be monotone increasing function with respect to $\boldsymbol{f}$, from the eqn. (29) of Theorem 7, we have the direction-finding problem in a final form

$$\min_{x,s} \left[-\sum_{n=1}^{N} \frac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial w_n} \frac{\partial f_n(\boldsymbol{x}_n^0)}{\partial \boldsymbol{x}_n} \boldsymbol{z}_n + \sum_{n=1}^{N} \frac{\partial \Phi(\boldsymbol{w}(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}_n} \boldsymbol{s}_n \right] \tag{31.a}$$

$$\text{sub. to} \sum_{n=1}^{N} s_{ni} \leqq 0, \ i \varepsilon B \tag{31.b}$$

$$-1 \leqq s_{ni} \leqq 1, n=1, \cdots, N, i=1, \cdots, \dim \boldsymbol{\alpha}_n \tag{31.c}$$

$$\frac{\partial g_{ni}(\boldsymbol{x}_n^0)}{\partial \boldsymbol{x}_n} z_n \geqq -s_{ni}, \ i \varepsilon C_n = \{ i | g_{ni}(\boldsymbol{x}_n^0) - \alpha_{ni} = 0 \}, n=1, \cdots, N \tag{31.d}$$

$$\frac{\partial q_{ni}(\boldsymbol{x}_n^0)}{\partial \boldsymbol{x}_n} \boldsymbol{z}_n \geqq 0, \ i \varepsilon D_n = \{ i | q_{ni}(\boldsymbol{x}_n^0) = 0 \}, n=1, \cdots, N \tag{31.e}$$

$$-1 \leqq z_{ni} \leqq 1, n=1, \cdots, N, i=1, \cdots, \dim \boldsymbol{x}_n \tag{31.f}$$

We obtain the locally best $\boldsymbol{s}$ by solving the above LP. Given a usable feasible direction $\boldsymbol{s}^k$, a new point $\boldsymbol{\alpha}^{k+1}$ is generated by solving a convex one-dimensional problem to determine a step size:

$$\min_{\beta} \{\Phi(w(\boldsymbol{\alpha}^k+\beta \boldsymbol{s}^k), \boldsymbol{\alpha}^k+\beta \boldsymbol{s}^k) | \boldsymbol{\alpha}^k+\beta \boldsymbol{s}^k \text{ satisfies (25), (26)}\}$$

where k is an iteration number.

IV. Application of Constrained Simplex Method

The feasible direction algorithm is analytical method to some extent, but it requires the assumptions of separability with respect to $\boldsymbol{x}_n$, monotone increasing of Φ with respect to $\boldsymbol{f}$, differentiability of all functions and the algorithm itself is fairly complicated.

Among direct search method exists Simplex Method which uses no derivatives. A simplex in the n demensional parameter space is a geometric figure having more than $(n+1)$ vertex points. The basic method is just to replace the worst vertex in the simplex by its reflexion in the centroid of the others, thereby producing a new vertex that is expected to improve the objective value so that the procedure of generating a new simplex can be repeated. The basic idea was improved with various modifications and Simplex Method for constrained optimization problem was developed (see in detail, SHIMIZU, 1976-c).

The Constrained Simplex Method is of simple principle and assumes almost no assumptions, thus is appreciated useful to problems with relatively small number of decision variables.

The procedure can be applied to the problems (3), (21), (23) to decide the optimal allocations. First, generate a simplex with a trial ($K=2$ dim $\boldsymbol{\alpha}$) points $\{\boldsymbol{\alpha}^k\}_{k=1}^K$. Since N local problems with the allocations $\{\boldsymbol{\alpha}_n^k\}$ may be solved by appropriate nonlinear programming, one can calculate $\{\boldsymbol{x}_n^0(\boldsymbol{\alpha}_n^k)\}_{n=1}^N$ and $\{f_n(\boldsymbol{x}_n^0(\boldsymbol{\alpha}_n^k), \boldsymbol{\alpha}_n^k)\}_{n=1}^N$ given $\boldsymbol{\alpha}_n^k$.

Then, as one can evaluate values of Φ and $\boldsymbol{G}$, the process of Constrained Simplex Method can be achieved.

Since this method does not need derivatives etc., it can be applied to any complicated structured problems numerically. In particular, it may be applicable to the case when the central objective is not defined explicitly by introducing man-machine interactive decision systems.

V. Conclusion

We have studied the hierarchical optimization problems for multi-objective systems in which there exist plural semi-autonomous local systems subordinate to a central system and in which the lower level composes a vector minimization problem. The coordinating center allocates scarce resources so that the central objective function is optimized, while the local systems optimize their own independent local objective with use of the given resources.

Two-level optimization problems for multi-objective systems were formulated and various theories were established, which were essentially parametric approach.

In perticular, decentralized optimization was studied for the hierarchical systems, in which the local systems were separated with respect to local decision variables. A primal method by resource allocation was proposed, which was a feasible derection method using directional derivatives.

Besides, Constrained Simplex Method was proposed to apply to the more general structured problem which necessitates to solve parametric programming.

The systems as newly formulated here exist in many real and important organizations.

REFERENCES

GEOFFRION, A.M. (1970): "Primal Resource-directive Approaches for Optimizing Non-linear Decomposable Systems," Opns. Res., Vol. 18, No. 3, pp. 375/403.

GEOFFRION, A.M., and W.W. HOGEN (1972): "Coordination of Two-level Organizations with Multiple Objectives," in A.V. Balakrishnan (ed.), Techniques of Optimization, Academic Press, pp. 455/468.

KORNAI, J. (1965), "Two-level Planning," Econometrica, Vol. 33, No. 1, pp. 141/169.

LASDON, L.S. (1970): Optimization Theory for Large Systems, McMillan, New York, pp. 396/459.

SHIMIZU, K. (1969): "Overall Optimization by Coordinating Center and Local Optimizations," Keio Engineering Report, Vol. 22, No. 9, Keio Univ.

SHIMIZU, K., and ANZAI, Y. (1974): "Optimization for Hierarchical System with independent Local Objectives" Trans. of the Society of Instrument and Control Engineers, Vol. 10, No. 1, (in Japanese).

SHIMIZU, K. et al, (1976-a): "Decentralized Two-level Optimization Systems and Multi-commodity Network Flow Problems," Preprint of Fifth SICE Control Theory Symposium, pp. 137/142, (in Japanese).

SHIMIZU, K. (1976-b): "Optimization Algorithms for Multiple Objective programs and Decentralized Hierarchical Systems," Proc. of IFAC-LSSTA, Italy (Udine), pp. 489/497.

SHIMIZU, K. (1976-c): Systems Optimization Theory, Corona, pp. 92/98.

SILVERMAN, G.J. (1972): "Primal Decomposition of Mathematical Programs by Resource Allocation: II," Opns. Res., Vol. 20, No. 1, pp. 58/74.