SIMULATION OF BEHAVIOR AND INTELLIGENCE *

SOME ASPECTS OF THE GENERAL THEORY OF BEST OPTION CHOICE **

M. A. Aizerman and A. V. Malishevski ***

UDC 62-505: 65.01

The general scheme of classical rules of selection of best options is discussed, in particular, optimization rules, where the structural principle of "binary comparability" of options is explicitly or implicitly used, as well as the specific logic of a "dominant" rule of selection. It is shown that it is possible and necessary to transcend the "pair-and-dominance" principle in the structures and selection rules, and to go over to non-conventional structures and rules. Such a transition can be realized by formalizing the concept of "multiple interrelations" in defining the best options.

1. Statement of Problem

Many problems of applied mathematics and control theory involve the selection of the "best" (in a certain sense) from a given set of options that are being compared. Thus, e.g., in models of optimal control, mathematical programming, and the like, the options can be in the form of plans, controls, strategies, etc., and the set of options that can be compared is specified by the constraints. In such problems, the choice of the best options is based on a (scalar or vector) optimality criterion using a certain extermination procedure. It is assumed that the scalar criterion that is being extremized (such as the objective function or the scale) or the set in the form of a "vector" of criteria (an n-tuple of scales) is assigned from outside; thus the aim of the theory is to establish the existence of a solution of the problem, to study its properties, and to develop a technique of determination of the optimum in specific problems. In control theory, the question of whether a sensible and purposeful choice of (in a certain sense) "best" options can be reduced to extremization based on an optimality criterion or on a set (vector) of criteria is usually not asked. Such problems belong to the

---

* 980, 0005-1179/81/4202-0184, $7.50 ©1981 Plenum Publishing Corporation.

** Moscow. Translated from Avtomatika i Telemekhanika, No. 2, pp. 65-83, February, 1981. Original article submitted September 5, 1980. Russian version: http://www.datalaundering.com/download/aspects-ru.pdf

*** This paper has been prepared by the authors for a special section of "IEEE Transactions on Automatic Control" dedicated to R. Bellman, and it is being published simultaneously in "Avtomatika i Telemekhanika." Its structure differs from the one usually adopted in "Avtomatika i Telemekhanika" by the absence of proofs of theorems. The authors intend to return in later publications to this subject and will then fill in the proofs omitted here, as well as other details.

subject matter of a related discipline, namely the theory of decision-making, which is mainly based on socioeconomic and psychological problems. Towards the middle of this century it appeared that the development of this theory had yielded exhaustive answers to this type of problem (see, e.g., review articles [1, 2], as well as monographs [3, 4]). Optimization choice based on one or several criteria had been represented in generalized terms of choice according to binary preference relations, and a system of axioms of "rational choice" had been formulated such that the fulfillment of these axioms for a given method of selection is equivalent to the existence of its optimization representation. As a result, it seemed that the "rational" methods of selection could be clearly distinguished from the "non-rational" or "pretentious" methods, and that the extremization approach to the choice of the best options was based on the reliable foundations of the theory of decision-making.

In the seventies, however, the situation became less clear. Articles were published more frequently in which well-founded doubts were expressed about the need that a rational choice must satisfy the axioms of rationality, and specific examples as well as whole classes of methods and procedures of selection were described, which although fully rational did not necessarily satisfy these axioms (see, e.g., [5, 6]). Therefore the types of questions raised above are again topical, i.e., which of the various common and natural methods of selection of the best options can be reduced to the classical optimization models, and which cannot be reduced, and how these models can be extended and generalized in such a way as to include rational (though not satisfying the classical axiomatics) methods of selection, To introduce the reader to this range of problems is the aim of this paper. It contains results known in the literature, but presented here in a slightly different formulation (Theorem 1) as well as new results (Theorem 2 and the following).

2. Formal Model and Selection Procedures

The choice of options is considered here on the basic of the following abstract model. We are given a set A of options $x, y, \ldots$. For simplicity we shall assume that A is finite and that any nonempty subset of options $X \subseteq A$ can be used for the selection.[1]

[1] In control theory, the set of options is in many cases infinite, but fundamental problems can be better studied by considering a finite A. The basic ideas of this article and many of our results can be extended (after making appropriate changes) to the case of infinite A, and also to the case that not just any subsets $X \subset A$ are used in the selection.

The act of selection consists in singling out from X some of the options $Y \subseteq X$ ($Y \neq \varnothing$) according to a certain rule.[2] The totality of all sorts of acts of selection generated by a given deterministic method of selection specifies the corresponding set of pairs $\{X,Y\}$. The latter is equivalent to specifying a choice function C, which assigns to every $X \subseteq A$ its subset $Y = C(X)$ [1,2,4,6-9].

Everywhere below, in referring to a certain method or type of selection, we shall have in mind not an individual act of selection, but, "selection in the large," i.e., the totality of acts of selection generated by the application of a given method of selection to all possible occurrences of $X \subseteq A$.

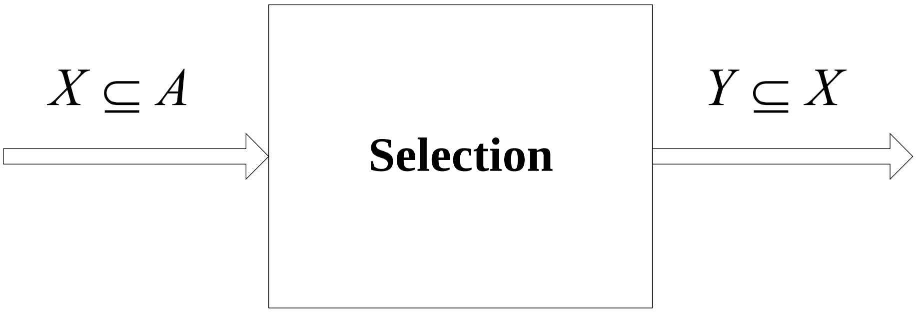

Fig. 1

According to conventional control theory, it is possible to represent the selection by the "converter" model shown in Fig. 1. To the input of the "selection" unit we apply the sets of options $X \subseteq A$, and at its output we obtain the subsets $Y \subseteq X$. Such a converter can have two descriptions: 1) an "internal" (or selection procedure) description which shows how to find on the basis of a given set X its "best part" Y, and 2) an "external" (input-output) description, which consists in indicating the set of pairs $\{X,Y\}$ generated by this procedure, i.e., the function $C(A)$.

Selection procedures that generate the same choice function C are said to be equivalent.

We shall use the following standard form of description of the selection procedure. Such a procedure can be specified by fixing a structure[3] σ on the set A and a rule π, which indicates how to find Y for each occurrence of X by using the structure σ. Let us illustrate these concepts by examples.

---

[2] The condition of nonemptiness of the choice ($Y \neq \varnothing$) is sometimes regarded as one of the axioms of "rationality" of choice. In this paper we retain it only for the sake of simplicity and brevity of the analysis. On the other hand the set Y may include a number of elements that are being selected simultaneously; for this reason we are using the term "options" instead of the more common term "alternatives" which usually refers to mutually exclusive objects of selection.

[3] Here the term "structure" is used in a sense close to the commonly accepted, i.e., not rigorously defined, mathematical concept (such as algebraic structure, ordering structure, etc.); in this introductory part we do not need general formal definitions.

Example 1. Scalar Optimization Selection. The structure σ is specified by a mapping φ of the set A onto the number axis, which in this case has the meaning of "worse-better" axis. Such a mapping $\varphi(x)$ is called a "scale," "criterion," etc.[4] In this case the rule π consists in extremization (more precisely, maximization) of $\varphi(x)$ on the set X, i.e., we must select from X a subset Y defined as

$$C(X) = Y = \{y \in X | \varphi(y) \geq \varphi(x) \text{ for any } x \in X\} \tag{1}$$

or, what amounts to the same thing,

$$C(X) = Y = \{y \in X | \text{ there exists no } x \in X \text{ such that } \varphi(x) > \varphi(y)\}. \tag{2}$$

The thus-defined scalar optimization procedure incorporates in the choice of Y [(1) or (2)] all the options with a maximum value of φ on X (and only those).

Example 2. Vector Optimization Selection. In this case the structure σ is defined in the form of $n > 1$ mappings $\varphi_j(x)$, $j = \overline{1,n}$ of the set A onto the number axis, i.e., we are given a vector function $\varphi(x) = (\varphi_1(x), \ldots \varphi_n(x))$. The rule π consists in "vector maximization" of the function $\varphi(x)$ on X, in the sense that we select from X the set of all Pareto-optimal options on the basis of a "vector criterion" (i.e., the Pareto set). Such a rule can be described by the above formula (2) if we interpret in it the inequality $\varphi(x) > \varphi(y)$ as a vector inequality [5] (let us note that with a similar vector interpretation of the inequality $\varphi(y) \geq \varphi(x)$ occurring in expression (1), the latter yields another rule [6]).

---

[4] For simplicity we shall refer only to numerical scales. But all our analyses with regard to Examples 1-5 can be extended also to more general order scales.

[5] There are two common vector interpretations of this inequality: a "strict" interpretation, when $\varphi(x) > \varphi(y)$ is understood as a set of component inequalities $\varphi_j(x) > \varphi_j(y)$, $j = \overline{1,n}$ and a $\varphi(x) \geq \varphi(y)$ "semistrict" interpretation, when $\varphi(x) > \varphi(y)$ is defined as $\varphi(x) \geq \varphi(y)$ and $\varphi(x) \neq \varphi(y)$, where $\varphi(x) \geq \varphi(y)$ is understood as a set of component inequalities $\varphi_j(x) > \varphi_j(y)$, $j = \overline{1,n}$. Usually the definition of Pareto optimality is based on the "semistrict" definition of the inequality $\varphi(x) \geq \varphi(y)$, although it is also possible to consider $\varphi(x) \geq \varphi(y)$ "weak Pareto optimality" on the basis of the "strict" vector inequality. All the subsequent assertions with regard to the vector optimization procedure remain valid for either of these definitions of the vector inequality and of the Pareto set in (2).

[6] In the vector case, the choice (1) signifies that we select the options that are best in X according to each of the scales φ_j, $j = \overline{1,n}$ simultaneously; as a rule, such a choice is empty.

To avoid misunderstandings, let us note that the thus-defined vector optimization procedure includes in the choice all the Pareto optimal options without distinguishing between them (on principle). Sometimes we understand by "vector optimization" the selection of some special options from all the options that are Pareto optimal. According to the above definition this will be another procedure; some examples of such procedures, which single out "part" of the Pareto set will be considered below (see Examples 4 and 5). A very simple two-dimensional illustration is shown in Fig. 2. Here $\widetilde{X}$ is the Pareto set for an occurrence of X represented in the plane of two scales (criteria) φ_1 and φ_2. The set $\widetilde{X}$ contains the "extreme" elements a and c, which are the options that are optimal according to one of the scales φ_1 and φ_2. Moreover, $\widetilde{X}$ includes also an "intermediate" element b, which is a "compromise" option that is Pareto optimal according to the scale vector (φ_1, φ_2), but not optimal according to either of the scales separately. The procedure of finding the "extreme" options will be considered in Example 4 below, whereas another procedure, which in general singles out the "Compromise" options, will be considered in Example 5.

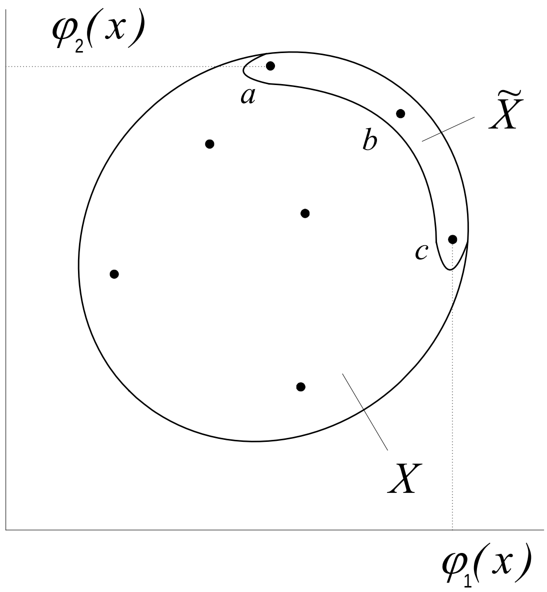

Fig. 2

The logic of the procedures of selection (1) and (2) is based on pair comparison of options. This logic is illustrated in "pure form" and in most general form in Example 3 that follows.

Example 3. Pair-Dominant Choice and its Graph-Dominant Realization. In this selection procedure, the structure σ [7] is defined by binary relation D on the set A (the notation xDy for $x, y \in A$ signifies that "x is in a relation D to y"). The rule π can be obtained by replacing in the definitions of the rules of extremal selection (1) and (2) the inequalities by the relations D:

$$Y = \{y \in X | yD_1x \text{ for any } x \in X\}, \tag{1'}$$

$$Y = \{y \in X | \text{ there exists no } x \in X \text{ such that } xD_2y\}. \tag{2'}$$

The definitions (1') and (2') are equivalent if $D_2 = (\overline{D}_1)^{-1}$, i.e., when the relations D_1 and D_2 are mutually "inversely complementary": $yD_2x \Leftrightarrow (\text{not } xD_1y)$. Next, assuming that this condition holds, we shall say that the inverse relation $(\overline{D}_1)^{-1}$ in (1') is a permitting relation, being denoted by α, whereas D_2 (2') is a prohibiting relation, denoted by β. By definition, the relations α and β are mutually complementary: $\beta = \overline{\alpha}$, $\alpha = \overline{\beta}$.

The relation $x\alpha y$ will be interpreted as "the option y is permitted by the option x," whereas the complementary relation $x\beta y$ is interpreted as "the option y is forbidden by the option x." An option $y \in X$ that is permitted by every (or, what amounts to the same, not forbidden by any) option $x \in X$ will be called a dominant in X. Thus a dominant is defined in terms of "pair comparisons" of a given option with all the other options in X. The procedure of selection of all the dominants in X ((1') or (2')) will be called a pair-dominant procedure.

In contrast to optimization procedures, a pair-dominant procedure does not ensure in the general case "automatically" that the choice is nonempty. In continuing to impose the condition of nonemptiness, we must therefore confine ourselves in this paper only to relations D_1 and D_2 in (1')-(2') that would ensure [8] nonemptiness of Y for any X.

Usually the investigated procedures of selection of the best options on the basis of "preference relations," i.e., strict preference P or nonstrict preference R [3, 4, 7-9], are perfectly fitting into the scheme of the procedure of pair-dominant choice. In this case P serves as D_2 in (2'), and R serves

---

[7] For simplicity we shall refer only to numerical scales. But all our analyses with regard to Examples 1-5 can be extended also to more general order scales.

[8] Thus in terms of the binary relation D_2, a necessary and sufficient condition of nonemptiness of Y for any nonempty $X \subseteq A$ in (2') is that D_2 must be acyclic [4, 8] (let us recall that the set A is finite).

as D_1 in (1'). Moreover, if the condition of nonemptiness of choice is satisfied, it would be possible from the very beginning to interpret [9] (and denote) D_1 in the definitions (1') and (2') as a relation of nonstrict preference (R), whereas D_2 can be interpreted as a relation of strict preference (P). However the conventional treatment of preference relations is strongly linked to the idea of "pair comparison" of options from which we intend to depart in the following. For the subsequent generalizations it is more convenient to use the "neutral" terms introduced by us above, i.e., the permitting relation α and the prohibiting relation β.

For greater clarity, it is convenient to represent the binary relations by graphs, by introducing the prohibition graph β or its complementary permission graph $\alpha = \beta$. The vertices of these graphs are the options $x \in A$, and the relation $x\alpha y$ (or $x\beta y$) is represented by a directed arc from x to y (in particular, $x\alpha x$ is a cycle). In the prohibition graph β the rule (2') amounts to selecting from the subgraph β_X corresponding to X the vertices $y \in X$ that do not receive arcs from other vertices of the subgraph β_X. In the permission graph α the rule (1') amounts to selecting the vertices $y \in X$ that receive arcs from all the vertices $x \in X$. The thus selected vertices (options y) are the dominants in β_X, and such a selection procedure will be called a graph dominant realization of a pair-dominant procedure (or, for brevity, a graph-dominant procedure).

The condition of nonemptiness of selection can be formulated very clearly with the aid of the graph β, i.e., this graph must, not have cycles.

The scheme of the pair-dominant procedure can be regarded as an abstract formulation of the concept of "optimization selection" (whereas the relations D_1 and D_2 can be regarded as abstract forms of the relations "not worse" and "better"). In particular, the class of pair-dominant selection procedures (example 3) "covers" in a certain sense the two classes of optimization procedures described above (examples 1 and 2). Indeed, to each optimization procedure of selection it is possible to assign an equivalent pair-dominant procedure with a prohibition relation

$$x\beta y \Leftrightarrow \varphi(x) > \varphi(y)$$

---

[9] In the general case, when an empty choice is allowed, such an interpretation may become invalid. Thus the dominance relation in games [10, 11] which is not necessarily asymmetric can serve as an e)<ample of a relation of type D_2, i.e., as a "prohibiting" relation β that cannot be always interpreted as "preference."

Such a relation (graph) β is said to be transitive [10] if φ is a vector and, moreover, strongly transitive [11] if φ is a scalar. Thus we can find the equivalents of optimization procedures only in special subclasses of pair dominant selection procedures with a transitive (for vector optimization) or even strongly transitive (for scalar optimization) [12] structure, i.e., with a partial ordering or weak ordering structure, respectively.

The selection procedures described in Examples 1-3 are said to be classical, such types of procedures or their particular cases usually occur in the literature dealing with the theory of choice [1-4, 7-9]. Each of these examples considers in fact not just one selection procedure M , but also an entire class $\mathcal{M}$ of such procedures. A specific procedure M can be singled out from such a class by singling out the "structural parameter" σ (by assigning a specific scale, a specific n-tuple of scales, or a specific relation D).

It is evidently possible to construct other classes of selection procedures by using these same types of structures (scales, n-tuples of scales, relations, or graphs), but by changing the selection rule π, with the possibility that the new selection rule is "similar" to the old one. Thus, for example, on the basis of a given n-tuple of scales it is possible to select from each X only part of the Pareto set, instead of the entire set (see Examples 4 and 5 below). Another example: If the structure is assigned by a graph γ, which denotes "prohibitions" (or strict preferences), then instead of selecting from X the vertices which do not receive any arcs in γ_X, we shall select the vertices y from the subgraph γ_X that maximize the difference $N = n_2 - n_1$, where n_1 is the number of arcs that arrive at y from other $x \in X$, and n_2 is the number of arcs that depart from y towards other $x \in X$ (see Example 6 below).

There arises the question of whether the various "nonclassical" selection procedures of this sort can be reduced to equivalent classical procedures. In particular, does there exist a scale or an n-tuple of scales on which it is possible to map the set of options A in such a way that the optimization

[10] Let us recall that a relation (graph) β is said to be transitive if from $x\beta y$ and $y\beta z$ there follows $x\beta z$.

[11] A relation (graph) β is strongly transitive if both β and its complement $\overline{\beta}$ are transitive.

[12] The converse is also true, i.e. any pair-dominant procedure with a transitive (strongly transitive) structure β can be equivalently reduced to a vector optimization (scalar optimization) procedure of selection (see Theorem 1).

selection (Examples 1 and 2) coincides for any $x \in A$ with an assigned nonclassical selection? In cases in which there is no such equivalent reducibility, there appear strongly nonclassical selection procedures. In such procedures it is possible to use the same structures σ as in the classical procedures, and their rule π may include extremization procedures, but the class of such selection procedures is certainly not "covered" (in the sense of equivalent reducibility) by the class of classical optimization procedures.

3. The Space $\mathscr{B}$ and the Characteristic Properties of Choice Functions

Any specific selection procedure M generates a certain choice function $C(X)$ that can be regarded as a "point" of the space $\mathscr{B}$ of all sorts of choice functions on A; the class $\mathscr{M}$ of selection procedures generates a corresponding class of choice functions that form in the space $\mathscr{B}$ a "region" $\mathscr{B}_{\mathscr{M}}$.

Equivalent selection procedures generate the same point C^* in the space $\mathscr{B}$. Two classes of selection procedures are said to be equivalent if they generate the same region $\mathscr{B}^*$ in the space $\mathscr{B}$.

Equivalent classes $\mathscr{M}_1$ and $\mathscr{M}_2$ of selection procedures evidently have the same possibilities, i.e., for any procedure M_1 of the class $\mathscr{M}_1$ there exists an equivalent procedure M_2 in the other class $\mathscr{M}_2$ and conversely. If a class of selection procedures $\mathscr{M}_1$ generates a larger region $\mathscr{B}_1$ than the region $\mathscr{B}_2$ generated by the other class of procedures $\mathscr{M}_2$, i.e., $\mathscr{B}_1 \supset \mathscr{B}_2$ then it is natural to assume that the procedures belonging to the class $\mathscr{M}_1$ have "greater possibilities compared to $\mathscr{M}_2$. Indeed, in this case there exists for any procedure $M_2 \in \mathscr{M}_2$ a procedure $M_1 \in \mathscr{M}_1$ that "does the same" as $\mathscr{M}_2$, i.e., that is equivalent to M_2, whereas the converse is not true. Thus by considering the space $\mathscr{B}$ and by comparing the points and regions generated in this space by the selection procedures and their classes, we obtain a formal basis for a meaningful comparison of the "possibilities" of various selection procedures.

Together with the space $\mathscr{B}$ let us introduce the "structure space" Σ. All the scales, all the n-tuples of scales, and all the regions (graphs) form corresponding "regions" in this space; a specific scale, a specific n-tuple of scales, or a specific relation (graph) will be a point σ of such a region. In the structure space Σ it is evidently possible to include a priori also many other types of structures

σ. Then we could represent a "well-defined" class of selection procedures as a given set of structures $\tilde{\Sigma}$ with an assigned selection rule $\tilde{\pi}$. Here $\tilde{\pi}$ is an operator that maps the points σ of the structure space Σ into the points C of the space $\mathscr{B}$. Thus the region $\tilde{\Sigma}$ of the structure space will be mapped onto the region $\mathscr{B}_{\tilde{\Sigma},\tilde{\pi}}$ of the space $\mathscr{B}$ generated by the rule $\tilde{\pi}$. For establishing a correspondence between the regions $\mathscr{B}_{\tilde{\Sigma},\tilde{\pi}}$ generated by different classes of selection procedures, it is convenient to use independent concepts such as certain properties of choice functions henceforth called characteristic properties. To each characteristic property $\mathscr{P}$ we shall assign a region $\mathscr{B}^{\mathscr{P}}$ in the space $\mathscr{B}$ that consists of all the functions C having the property $\mathscr{P}$ (and only them). The regions $\mathscr{B}^{\mathscr{P}}$ will serve as "references" in the space $\mathscr{B}$.

The characteristic properties of choice functions have been introduced into the theory of decision making mainly with a special purpose, namely for outlining the region of "classical" choice functions. We shall present them here and give them special designations, having in mind a more general task, i.e., the description of fairly arbitrary (not necessarily classical) classes of procedures $\mathscr{M}_{\Sigma,\pi}$ and classes of choice functions $\mathscr{B}^{\mathscr{P}}$.

Definition. We shall say that a function $C(X)$ satisfies the condition of:

1. Succession $\mathbf{H}$ (which is the same as Postulate 4 of Chernoff [12], or condition α of Sen [13], or the axiom C2 of Arrow-Uzawa [14]) if

   it follows from $X' \subseteq X$ that $C(X') \supseteq C(X) \cap X'$.

2. Strict succession or constant residual choice $\mathbf{K}$ (it is the same as postulate 6 of Chernoff [12] and one of the forms of the "weak axiom of revealed preference" of Samuelson, i.e., the axiom C4 of Arrow [14]) if

   it follows from $X' \subseteq X$ and $X' \cap C(X) \neq \varnothing$ that $C(X') = C(X) \cap X'$.

3. Compatibility $\mathbf{C}$ (it is the same as Postulate 10 of Chernoff [12] and condition γ of Sen (13) if

   it follows from $X = X' \cup X''$ that $C(X) \supseteq C(X') \cap C(X'')$.

4. Independence with respect to dropping rejected options (or, for brevity, elimination of options) $\mathbf{O}$ (Postulate 5 [13] of Chernoff [12], or Axiom 2 in [15]) if

   it follows from $C(X) \subseteq X' \subseteq X$ that $C(X') = C(X)$.

---

[13] This theorem expresses in "synthetic" form a number of well-known particular assertions concerning the correspondence between various "classical" types of choice procedures and functions [3, 4, 15, 16].

All these properties $\mathbf{H}$, $\mathbf{K}$, $\mathbf{C}$, and $\mathbf{O}$ have a sufficiently clear meaning as being requirements towards the "logic" of the observed choice $C(X)$ that represent to a certain extent the qualities of the "better" [8, 12, 13].

Now let us consider the formal relations between these properties. Let us begin with the conditions $\mathbf{H}$, $\mathbf{C}$ and $\mathbf{O}$. They are independent properties in the sense that there exist choice functions that satisfy one of the eight possible conjunctions of these conditions and their negations:

$$\mathbf{H} \cap \mathbf{C} \cap \mathbf{O},\ \overline{\mathbf{H}} \cap \mathbf{C} \cap \mathbf{O}, \ldots, \overline{\mathbf{H}} \cap \overline{\mathbf{C}} \cap \overline{\mathbf{O}}.$$

This means that in the space $\mathscr{B}$ the corresponding regions $\mathbf{H}$, $\mathbf{C}$, $\mathbf{O}$ and their complements $\overline{\mathbf{H}}$, $\overline{\mathbf{C}}$, $\overline{\mathbf{O}}$ have nonempty intersections (Fig. 3). With regard to the condition $\mathbf{K}$, it simply strengthens each of the conditions $\mathbf{H}$, $\mathbf{C}$ and $\mathbf{O}$, i.e., from $\mathbf{K}$ we obtain $\mathbf{H} \cap \mathbf{C} \cap \mathbf{O}$ but not conversely, so that the region $\mathbf{K}$ lies strictly inside the intersection of three regions $\mathbf{H} \cap \mathbf{C} \cap \mathbf{O}$ (Fig.3).

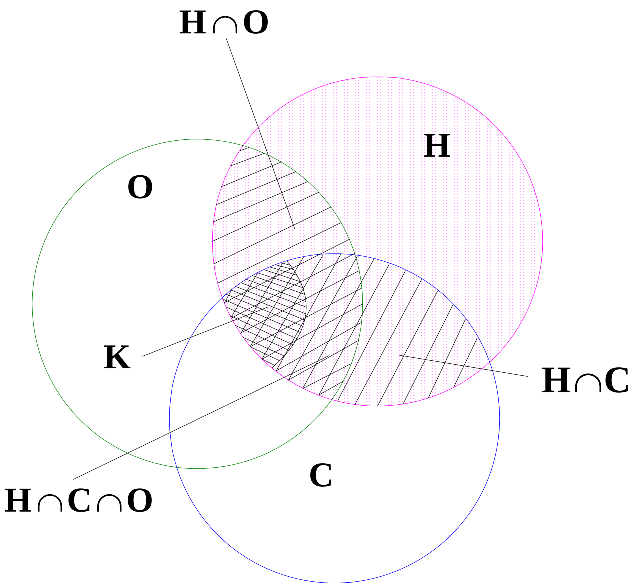

Fig. 3

It turns out that in such a partition of the space $\mathscr{B}$ it is convenient to locate the regions generated by classical selection procedures:

THEOREM 1.[14] For a choice function to be generated by the procedure of
1°) scalar-optimization (Example1),
2°) vector-optimization (Example 2),
3°) pair-dominant (Example 3)

selection, it is necessary and sufficient that it (respectively) satisfy the conditions

[14] This theorem expresses in "synthetic" form a number of well-known particular assertions concerning the correspondence between various "classical" types of choice procedures and functions [3, 4, 15, 16].

1) $\mathbf{K}$,
2) $\mathbf{H} \cap \mathbf{C} \cap \mathbf{O}$,
3) $\mathbf{H} \cap \mathbf{C}$.

The class of vector-optimization (scalar-optimization) procedures of selection is equivalent to the selection of pair-dominant procedures specified by transitive (strongly transitive) prohibition graphs or relations[15] β. By virtue of Theorem 1, for any scale, n-tuple of scales, or graph (relation), all the classical selection procedures (Examples 1-3) generate choice functions that lie at the intersection of the region $\mathbf{H}$ and $\mathbf{C}$ in the space $\mathscr{B}$. The regions generated in the space $\mathscr{B}$ by the classes of selection procedures 1°, 2° and 3° are represented in Fig. 3 by the "classical" regions $\mathbf{K}$, $\mathbf{H} \cap \mathbf{C} \cap \mathbf{O}$ and $\mathbf{H} \cap \mathbf{C}$ respectively, these regions being embedded in each other.

Now the above questions concerning the reducibility of selection procedures to classical (optimization) procedures can be formulated as follows: Does a given procedure or class of selection procedures generate a point or region in space $\mathscr{B}$ that lies inside the intersection $\mathbf{H} \cap \mathbf{C}$ or, more rigidly confined, inside $\mathbf{H} \cap \mathbf{C} \cap \mathbf{O}$ or even inside $\mathbf{K}$?

Negative answers to these questions lead to a problem that is of intrinsic interest, i.e., which procedures of generation of "nonclassical" regions are possible in the space $\mathscr{B}$?

As an example let us consider one such "nonclassical region," namely the intersection $\mathbf{H} \cap \mathbf{O}$ of the regions $\mathbf{H}$ and $\mathbf{O}$ represented in Fig. 3 by dashes (side by side with the "classical" regions $\mathbf{H} \cap \mathbf{C}$, $\mathbf{H} \cap \mathbf{C} \cap \mathbf{O}$ and $\mathbf{K}$). Let us begin with the description of a "nonclassical" selection procedure.

<u>Example 4. Aggregate-Extremal Selection.</u> The structure σ of scales is the same as in Example 2, i.e., an n-tuple (n > 1) of scales. The rule π is as follows: On the basis of each ("criterion") we separately perform the scalar-optimization selection (1) on the set X (as in Example 1), and the options thus selected are combined into a set $Y = C(X)$.

<u>THEOREM 2.</u> Suppose we are given a vector-optimization selection procedure based on an array of scales $\varphi_i(x)$, $i = \overline{1,n}$. Then this array of scales can be supplemented to a new array $\varphi_j(x)$, $j = \overline{1,m}$ $(m > n)$ in such a way that the aggregate-extremal selection procedure on the new array will be equivalent to the original vector-optimization procedure.

[15] Let us recall that β is always assumed to be acyclic, this being the condition of nonemptiness of pair-dominant choice.

Theorem 2 asserts that for any $X \subset A$ the aggregate-extremal selection on a "new" array will coincide with the Pareto set for the "old" array of scales (it is also possible to show that $C(X)$ is a Pareto set also for the "new" array of scales).

By virtue of this theorem the class of aggregate-extremal selection procedures will "incorporate" the class of vector-optimization selection procedures (Example 2) and hence also of scalar-optimization procedures (Example 1). On the other hand we have the following theorem.

THEOREM 3. For a function $C(X)$ to be generated by an aggregate-extremal selection procedure, it is necessary and sufficient that it satisfy the condition $\mathbf{H} \cap \mathbf{O}$ (Fig. 3).

Taking into account the assertion that the properties $\mathbf{H}$, $\mathbf{C}$ and $\mathbf{O}$ are independent, it follows from Theorem 3 that aggregate-extremal selection procedures can be equivalently reduced only in special cases (and not always) to classical optimization procedures or to just any pair-dominant procedures (although they involve extremization). Indeed, the region $\mathbf{H} \cap \mathbf{O}$ does not overlap with the "classical" region $\mathbf{H} \cap \mathbf{C}$ of pair-dominant choice functions, but only intersects it (Fig. 3).

By noting that the intersection of the regions $\mathbf{H} \cap \mathbf{O}$ and $\mathbf{H} \cap \mathbf{C}$ is the region $\mathbf{H} \cap \mathbf{C} \cap \mathbf{O}$, which occurs in part 2 of Theorem 1, we obtain from Theorem 3 the Corollary 1.

COROLLARY 1. Let the aggregate-extremal selection procedure be equivalent to a pair-dominant procedure. Then it will be equivalent to a vector optimization procedure.

This corollary shows that any given aggregate-extremal selection procedure is either equivalent to a vector optimization procedure or not equivalent to any "classical" pair-dominant procedure.

Plott [6] has considered a class of choice functions that satisfy the following condition of path independence (PI):

$$\text{if } X = X' \cup X'' \text{ then } C(X) = C\left[C(X') \cup C(X'')\right].$$

THEOREM 4. For a function $C(X)$ to be independent of the path in the sense of Plott, it is necessary and sufficient that it should be generated by an aggregate-extremal selection procedure.

By virtue of this theorem, all sorts of aggregate-extremal selection procedures will generate all the Plott functions, and only them. In conjunction with Theorem 3 this yields Corollary 2.

COROLLARY 2.[16] Property (7), i.e., "path independence," for choice functions is equivalent to a conjunction characteristic properties $\mathbf{H} \cap \mathbf{O}$.

[16] A similar result has been obtained in [17].

Example 5. Superposition of Classical Selection Procedures. The structure σ is the same as in Examples 2 and 4, i.e., an n-tuple of scales. The rule π is a two-step rule, i.e., at first we select from X the Pareto set (as in Example 2), and then we select from the latter the "best options" on the basis of an additional scalar optimality criterion defined on A (as in Example 1). [17]

It is possible to show that a choice function generated by such a two-step procedure will always satisfy the condition $\mathbf{C}$, but it may not satisfy the condition $\mathbf{H}$ (and also not the condition $\mathbf{O}$); thus it is not contained in he "classical" region $\mathbf{H} \cap \mathbf{C}$.

Hence the superposition of two classical selection procedures may generate a "nonclassical" selection. We have a similar situation also for some other selection procedures in which the rule π "separates part" of the Pareto set.

Example 6. Tournament Selection. In fact this example has already been described above as a formal modification of the graph-dominant procedure, i.e., as selection of vertices on a directed graph according to the rule of maximization of the difference between the number of outgoing and incoming arcs. Let us show how such a "tournament" rule appears in an actual problem.

Let A be a set of chess players who can play in one- round tournaments with different participants, and suppose that in each tournament the winners are determined ("selected") by the ordinary rule, i.e., according to the maximum number of points (a win yields 1 point, a loss 0, and a draw half a point). We shall adopt the simplifying assumption of "constant results," according to which the result of a game between the players x and y ($x, y \in A$) is always the same. On the basis of this assumption let us construct a directed graph of wins [18] γ on the set of vertices A (x is linked to y by an arc if x wins against y; in the case of a draw, there is no arc between x and y). To each formation of the tournament $X \subset A$ there corresponds a subgraph of this graph on the set of vertices X, which by virtue of the assumption of constant results constitutes the graph of wins in the

[17] Such a selection rule can be regarded as a specific extension of a well-known lexicographic rule [3, 4], i.e., if we use one scale ($n = 1$) at the first stage, then we obtain precisely a lexicographic selection based on two successively used scalar criteria.

[18] Such a graph γ will be antireflexive and antisymmetric (i.e., without cycles and without return arcs). Let us note that in the theory of graphs a "tournament graph" usually presupposes also the fulfillment of the following condition: Any two vertexes $x \neq y$ are connected by an arc (in our terminology this signifies that there are no draws) [18]. We are not assuming this here.

tournament X. Ii we know the original graph γ on A, and hence also its subgraph γ_X on X it is thus possible to indicate the winner (or winners) of any tournament X. According to the above tournament rule (more precisely' according to the equivalent rule of maximum number of points with 0 for a draw and -1 for a loss), the winners will be represented by the vertices $y \in X$ that maximize the difference $N = n_2 - n_1$, where n_2 is the number of their outgoing arcs, and n_1 the number of their incoming arcs from other vertices $x \in X$.

Let us assume that the results of "everyone with everybody else" games are such that all the chess players (the entire set A) can be partitioned into an ordered family of disjoint subsets in such a way that the following two conditions are satisfied: 1) Within each subset all the games between all the chess players end in a draw; 2) each player belonging to a certain subset wins against all the players belonging to the "lower" subsets (according to their ordering). This case of "weak ordering" of the players in A according to the results of the games played with one another holds if and only if this graph of wins γ is strongly transitive. In this (and only in this) case the selection of the winners according to the tournament rule for any formation $X \subset A$ (which coincides with selection according to the graph-dominant rule on the graph of wins γ regarded as a prohibition graph β) will satisfy the condition $\mathbf{K}$, and it can be reduced to scalar-optimization [19] selection.

It is evident that such a graph of wins is a very particular case. It is easy to find graphs of that the wins such thus-generated function $C(X)$ does not satisfy not only the condition $\mathbf{K}$, but also the conditions $\mathbf{H}$, $\mathbf{C}$ and $\mathbf{O}$. An example of such a graph with four vertices (options) a, b, c, d is plotted in Fig.4-1).

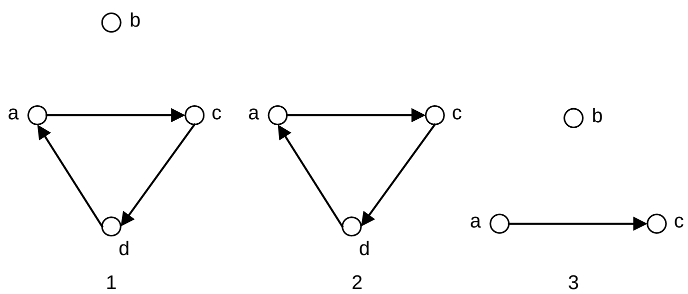

Fig.4. 1) Graph of wins γ on $A = \{a,b,c,d\}$, 2) subgraph γ_X of graph γ on $X = \{a,c,d\}$, 3) subgraph γ_X of graph γ on $X = \{a,b,c\}$.

[19] Here it is possible to obtain the scale φ by estimating $\varphi(x)$ by assigning to each chess player x the number of the subset to which he belongs (counted "from bottom to top").

For showing that condition H is violated for a tournament choice $C(X)$ on this graph of wins, let us consider its subgraph γ_X on the set of vertices $X = \{a, c, d\}$ (Fig.4-2). It is easy to see that according to the tournament rule we have $C(X) = \{a,c,d\}$. By then taking $X' = \{a,c\}$, we obtain $C(X') = \{a\}$. By noting that $c \in C(X)$ and $c \in X' \subset X$, but $c \notin C(X')$, we can see that $C(X) \cap X' \not\subset C(X')$, so that the succession condition is violated.

Thus even such a commonly used procedure of selection of "best players" as tournaments governed by the rule of "maximum number of points" cannot be reduced in the general case to classical (such as scalar-optimization) procedures of selection, although it involves extremization of a scalar quantity (the number of points).

4. Binary and Multiple Interrelations of Options and the Role of the Context of the Choice

The above examples have proved the inadequacy of the conventional optimization approach to the general problem of selection of best options. To explain the basic reason for this, let us return to Example 4 and consider the case that three options (a, b, and c) are located in the plane of two scales (criteria), as shown in Fig. 2. In the case of two-element occurrences $X' = \{a,b\}$ and $X'' = \{b,c\}$, the two options in each occurrence, with the use of the aggregate-extremal rule, will be subject to the choice $C(X') = \{a,b\}$, and $C(X'') = \{b,c\}$. But if all three options are displayed $\widetilde{X} = X' \cup X'' = \{a,b,c\}$, then the choice will have the form $C(\widetilde{X}) = \{a,c\}$, so that $C(X') \cap C(X'') = \{b\} \not\subset C(X)$, i.e., the classical condition of compatibility has been violated. Here it is natural to assume that the option b has not been included in the selection from $X = \{a,b,c\}$ for the reason that this option is "suppressed" ("prohibited") by the simultaneous presence of the options a and c in the occurrence of X. Such a joint "prohibition" is one of the examples of multiple "interrelations" between options.

Such multiple relations have the result that the selection of a given option is strongly affected by the overall "context of the choice," i.e., by the displayed set X as a whole. In contrast to the classical pair-dominant case, this influence cannot be in general divided into influences of "permission-prohibition" type due to individually considered "competing options" and it cannot be detected by the method of binary comparisons of options, Precisely these strong multiple influences due to the overall context of the choice make it impossible to reduce various natural and intrinsically logical procedures of selection to procedures using the conventional pair-dominant logic of selection (the "abstract optimization" logic). In other words, the "selection of best options" is no longer simple if the very concept of "best" has a complicated dependence on the context of the choice and cannot be reduced to the concept of "better" (or "not worse than") in a binary comparison.

We have a similar situation in Example 6, i.e., the selection of the winner in a tournament may depend in a complex manner on the entire "context," since the number of points that can be won by a given player will depend on the participation in the tournament. In particular, if a new player enters the tournament, this may cause a change in the person who wins, even if this new player is not the winner.[20] This example clearly illustrates also the difference between scalar-optimization selection as defined above and a more arbitrary selection based on maximization of a function. Indeed, in Example 6 the function to be maximized (the "number of points of the player x in the tournament X ") has the form $\varphi(x,X)$, and not the form $\varphi(x)$, which should have been possessed by the scale in the definition of a scalar optimality criterion. The dependence of the "estimate" φ of the option x not only on x itself, but also on the "context" X within which this option is compared with others, distinguishes such a selection procedure from the "narrow-sense" optimization procedure formulated in Example 1.

In many cases selection according to a scalar optimality criterion implies (explicitly or implicitly) a relaxation of the definition of a scalar-optimization procedure that would allow the extremization function to depend on the context of the choice. If in such a case the dependence of φ on X can be a6 desired, then it is easy to formally regard any preassigned choice $Y = C(X)$ as "optimal." For this purpose is suffices, e.g., to write $\varphi(x,X) = 1$ for $x \notin C(X)$ and $\varphi(x,X) = 0$ for $x \notin C(X)$. However this is in fact equivalent to assigning the estimate φ not a priori to the option x, but a posteriori to the "composite item" (x,X), which has the meaning of "the option x in the comparison context X." Let us note that in many cases the variational technique of solution of optimization problems not only provides for a possible dependence of the value of the extremization function φ on the "allowed set" X, but it may even deliberately introduce such a dependence.[21]

[20] The logic of such an effect of the context on the choice has been discussed, in particular, in the introduction of the characteristic properties of choice functions in [8] and [13].

[21] This is done in fact in replacing the problem of constrained extremization (with the constraints being in the form of equalities or inequalities) by unconstrained extremization, by transferring the constraints into the function to be extremized (the method of Lagrange multipliers or of penalty functions).

Here we are interested, however, not in "technical" reducibility of selection problems to extremization procedures, but in the possibility in principle to define the "best" option as one that is "better" (not worse) than arty other of the available options in some unconditional sense that does not depend on the context of the comparison. Examples of selection procedures that cannot be reduced to pair-dominant procedures show that this is not always possible even in problems of well-founded selection of the "best" options, this being due to the effects of "multiple" interactions of options. Hence it is necessary to study the "logic of relations" not only between the individual options, but also between whole sets of options. Here we can take as a reference the logic of selection as it appears in the general characteristic properties of choice functions, more precisely, in the above-mentioned properties of succession, compatibility, and elimination ($\mathbf{H}$, $\mathbf{C}$ and $\mathbf{O}$). Below we shall show that natural generalizations of the classical "pair" structure of relations on the set of options and of the classical "dominant" rule of selection lead to nonclassical schemes of selection procedures that can generate any choice functions in the regions $\mathbf{H}$, $\mathbf{C}$ and $\mathbf{O}$ of the space $\mathscr{B}$.

5. Generalized Dominant Selection Procedures

Now let us consider some procedures of selection that are generalizations of the pair-dominant procedure (Example 3), The structure σ will be taken in the form of generalized binary relations or hyperrelations that are established not (or not only) for the individual options, but for certain subsets of options that are taken as an entity; these are relations $\mathfrak{D}$ of the form $y\mathfrak{D}Z$, $Z\mathfrak{D}y$ and $Y\mathfrak{D}Z$, where the y are individual options belonging to A ($y \in A$), and Y and Z are subsets belonging to A ($Y, Z \subseteq A$). The rules π are taken in the form of generalized dominant rules, three actual versions of which are formulated below.

Example 7. Strongly Dominant Selection. The structure σ is a generalized binary relation (hyperrelation) $\mathfrak{D}$ between the individual options $y \in A$ and the subsets of options [22] $Z \subseteq A$.

[22] Similarly to the ordinary binary relation D on the set A, it is possible to assign such a hyperrelation by specifying all the pairs (y, Z) for which this relation holds (and only them).

A strongly dominant selection rule π can be defined by the following expression for the choice function:

$$C(X) = \{y \in X | y\mathcal{D}_1 Z \text{ for any } Z \subseteq X\} \quad (8)$$

or, what amounts to the same,

$$C(X) = \{y \in X | \text{ there does not exists a } Z \subseteq X \text{ such that } Z\mathcal{D}_2 y\}, \quad (9)$$

where $\mathcal{D}_2 = \left(\overline{\mathcal{D}_1}\right)^{-1}$ is the inversely complementary relation of $\mathcal{D}_1$: $yD_2Z \Leftrightarrow (\text{ not } ZD_1y)$.

The formulations (8) and (9) are direct extensions of the formulations (1') and (2') of the "classical" pair-dominant selection. In continuing this analogy, it is possible to regard the hyperrelation $\mathcal{D}_1$ or, more precisely, the inverse hyperrelation $(\mathcal{D}_1)^{-1}$ as a generalized permission relation α, and $\mathcal{D}_2$ as a generalized prohibition relation β. In accordance with this interpretation, $y\mathcal{D}_1Z$ signifies that "the set Z permits the option y," and $Z\mathcal{D}_2y$ that "the set Z prohibits the option y." Then the formulations (8) and (9) will denote the selection .of options y that are "permitted by all" (or "not prohibited by any") sets of options Z "occurring" in the context X.

Just as in the classical model of the pair-dominant selection procedure (Example 3), in Example 7 the condition of nonemptiness of selection imposes constraints on the allowed form of hyperrelations [23] $\mathcal{D}_1$ and $\mathcal{D}_2$. We shall say that the relation $\mathcal{D}_1$ (or $\mathcal{D}_2$), in the definition of the strongly dominant selection procedure (8) (or (9)) is correct if the generated set $C(X)$ is not empty for any nonempty $X \subseteq A$.

<u>THEOREM 5</u>. For a choice function $C(X)$ to be generated by a strongly dominant selection procedure (8) (or (9)) for any correct hyperrelation $\mathcal{D}_1$ (or $\mathcal{D}_2$), it is necessary and sufficient that it should satisfy the characteristic succession condition ($\mathbf{H}$).

Thus the class of all strongly dominant selection procedures for all sorts of correct hyperrelations $\mathcal{D}_1$ (or $\mathcal{D}_2$) will generate precisely the region $\mathbf{H}$ in the space of choice functions $\mathcal{B}$**.** Each function C in this region can be generated by the procedure (8) for the hyperrelation $\mathcal{D}_1$ (or the procedure (9) for $\mathcal{D}_2$). However, in contrast to the classical pair-dominant procedure, the hyperrelation $\mathcal{D}_1$ (or $\mathcal{D}_2$), which generate a given function C, cannot be obtained in a unique manner. In general there exist several equivalent strongly dominant procedures that differ by their structures $\mathcal{D}_1$ (or $\mathcal{D}_2$) but which generate the same choice function C.

---

[23] In particular, for $\mathcal{D}_2$ (the generalized prohibition relation) such a constraint reduces to some sort of "generalized acyclic property" of this hyperrelation.

Example 8. Weakly Dominant Selection. As in Example 7, the structure σ is in the form of a hyperrelation between $y \in A$ and $Z \subseteq A$.

The weakly dominant selection rule n is defined by the formula

$$C(X) = \{ y \in X \mid \text{ for any } x \in X \text{ and for at least one } Z \subseteq X \text{ such that } x \in Z \text{ we have } y \mathcal{D}_3 Z \} \tag{10}$$

or, what amounts to the same,

$$C(X) = \{ y \in X \mid \text{ there does not exists an } x \in X \text{ such that for any } Z \text{ such that } x \in Z \subseteq X \text{ we have } Z \mathcal{D}_4 y \}, \tag{11}$$

where $\mathcal{D}_4 = \left(\overline{\mathcal{D}_3} \right)^{-1}$.

The formulations (10) and. (11) are different (compared to (8) and (9)) generalizations of the "classical" formulations (1') and (2'). The hyperrelations $\left(\overline{\mathcal{D}_3} \right)^{-1}$ in (10) and $\mathcal{D}_4$ in (11) can also be interpreted as generalized relations of permission α and prohibition β, respectively. However the weakly dominant rule "processes" such relations in (10) and (11) in a somewhat different manner from the strongly dominant rule in (8) and (9).

We shall say that the hyperrelation $\mathcal{D}_3$ in (10) and $\mathcal{D}_4$ in (11) is correct if $C(X) \neq \varnothing$ is generated for any nonempty $X \subseteq A$.

THEOREM 6. For a choice function $C(X)$ to be generated by a weakly dominant selection procedure (10) (or (11)) for any correct hyperrelation $\mathcal{D}_3$ (or $\mathcal{D}_4$) it is necessary and sufficient that it should satisfy the characteristic condition of compatibility ($\mathbf{C}$).

The generating hyperrelations $\mathcal{D}_3$ and $\mathcal{D}_4$ occurring in Theorem 6 for assigned choice functions C can be obtained in a nonunique manner, as was also the case for the hyperrelations $\mathcal{D}_1$ and $\mathcal{D}_2$ in Theorem 5.

The nonuniqueness of the structure of the hyperrelation $\mathcal{D}$ in constructing a procedure of strongly or weakly dominant selection that generates a given choice function in the region $\mathbf{H}$ or $\mathbf{C}$ poses the question of choosing among various equivalent structures a structure that is in a certain sense convenient (i.e., simple or economical). Here we shall confine ourselves only to indicating such a possibility, and to some remarks. First of all, although the definition of correct hyperrelations $D_1 - D_4$ does not exclude a comparison $y \mathcal{D} Z$ (or $Z \mathcal{D} y$) of the option y with the empty set $Z = \varnothing$, we can confine ourselves without loss of generality to nonempty sets Z only (in (10) and (11) this is done "automatically"). Next, a comparison of the option y with a one-element set of the form $Z = \{z\}$ can be regarded as an inclusion of the ordinary binary relation [24] $\mathcal{D}$ (between the options y, $z \in A$) in the corresponding hyperrelation $\mathcal{D}$. Finally, a comparison of the option y with subset $Z \in A$ of general form as provided by the relations $\mathcal{D}$ can be described "economically" by taking into account the logic of the subsequent processing of the relation $\mathcal{D}$ by the corresponding rule π in (8)-(11), Thus, e.g., for two sets Z' and Z'' let $Z'\mathcal{D}_2 y$ and $Z''\mathcal{D}_2 y$, and suppose that $Z' \subset Z''$. It then follows directly from (9) that the relation $\mathcal{D}_2$ can be "simplified" by retaining $Z'\mathcal{D}_2 y$ and dropping $Z''\mathcal{D}_2 y$. The possibilities of "economical" construction of structures that generate selection procedures will be illustrated below more precisely.

For constructing a class of selection procedures that generate the last characteristic region in the space of choice functions, i.e., the region $\mathbf{O}$, we shall need an even more general type of structure, namely hyperrelations of the form $Y\mathcal{D}Z$, which connect pairs of sets of options.

<u>Example 9. Hyperdominant Selection</u>. The structure σ is a hyperrelation $\mathcal{D}$ between subsets of options [25] $Y, Z \subseteq A$.

The hyperdominant selection rule π on such a structure is defined by the formula

$$C(X) = Y, \text{ where } Y \subseteq X \text{ is such that } Y\mathcal{D}_5 Z \text{ for any } Z \subseteq X, \tag{12}$$

or, what amounts to the same, $\mathcal{D}_6 = (\mathcal{D}_5)^{-1}$

$$C(X) = Y, \text{ where } Y \subseteq X \text{ is such that there does not erxists a } Z \subseteq X \text{ such that } Z\mathcal{D}_6 Y\}. \tag{13}$$

We shall say that the hyperrelation $\mathcal{D}_5$ (or $\mathcal{D}_6$) in the definition of the hyperdominant selection procedure (12) (or (13)) is correct if for any nonempty $X \subseteq A$ the set Y defined by (12) (or (13)) is nonempty and unique.[26]

---

[24] It is possible to show that in hyperrelations $\mathcal{D}$ in the procedures (8)-(11) (adjusted in an equivalent manner if this is necessary for including in them the binary relation D) we retain the interpretation of D as relations of permissions (prohibition) of one option by another that appear in binary comparisons of options (the option y is allowed, i.e., not prohibited, by the option x if and only if $y \in C(\{z, y\})$).

[25] Such a hyperrelation on the set A is at the same time (and formally) a binary relation on the set 2^A, which is a family of subsets of A (for more details see the following).

[26] The requirement of uniqueness of definition of the set Y in (12) and (13) is due to the selection model adopted by us. In a more general formulation of the decision-making problem, it is sometimes allowed to make a nonunique choice of the subset of options, i.e., of the set of such subsets (see, e.g., [19]).

THEOREM 7. For a choice function $C(X)$ to be generated by a hyperdominant selection procedure (12) (or (13)) for any correct hyperrelation $\mathcal{D}_5$ (or $\mathcal{D}_6$) it is necessary and sufficient that it should satisfy the characteristic condition of elimination ($\mathbf{O}$).

Thus the class of hyperdominant selection procedures will generate, for all sorts of correct structures, all the choice functions in the region $\mathbf{O}$, and only them. The generating hyperrelation $\mathcal{D}_5$ (or $\mathcal{D}_6$) for such a choice function is in general not unique, and therefore we shall use here (as in the previous Examples 7 and 8) the criteria of convenience and economy in selecting it (in particular, we certainly can confine ourselves to comparing only nonempty subsets Y and Z).

6. Interpretation of Generalized Dominant Selection Procedures

The effects of "multiple" relations between options, which underlie the schemes of strongly, weakly, and hyperdominant procedures of selection become clearer in a slightly modified interpretation of the original model of selection. Let us recall that we regard the act of selection (X,Y) as a display of the set of options X from which we single out its subset ("best part") $Y \subseteq X$ ($Y = C(X)$. Until now e have interpreted X as a collection of elements (individual options $x, y, \ldots,$), and the choice Y as a collection of the "selected options" among them. But from a formal point of view we are just as entitled to interpret X as the simultaneous occurrence of all sorts of nonempty subsets $Z \subseteq X$ that influence (i.e., "permit" or "prohibit") the selection of certain options $y \in X$. Such an interpretation is suitable for strongly and weakly dominant procedures of selection (Examples 7 and 8).

Such an analysis could be enhanced by taking as the object of the selection not the individual options $y \in X$ but entire sets of options $Y \in X$. From this point of view the occurrence of X involves the simultaneous display of all the nonempty subsets $Z \subseteq X$, precisely one of which must be the "selected" one. Such an interpretation is suitable for the definition of a hyperdominant selection procedure (Example 9). Indeed, the expressions (12) and (13) can be regarded as a formal application of the definitions (1') and (2') of pair-dominant selection to the new situation; here the "options" are in the form of the sets Z of the original ("true") options $x, y, z, \ldots \in A$. With such an interpretation, an "occurrence" is represented not by the "true" set of displayed options $X \subseteq A$, but by the totality of nonempty subsets Z of this set X. Finally, the relations D_1 and D_2 in the definitions (1') and (2') are represented in this case, i.e., in (12) and (13), by the hyperrelations $\mathfrak{D}_5$ and $\mathfrak{D}_6$ respectively. The correctness of $\mathfrak{D}_5$ and $\mathfrak{D}_6$ signifies that precisely one (nonempty) set will always be selected from any such collection, By using the ordinary interpretation of the relations D_1 and D_2 in (1') and (2') as relations of preference nonstrict and strict respectively, and by extending it to the hyperrelations $\mathfrak{D}_5$ and $\mathfrak{D}_6$ in (12) and (13), we arrive at the interpretation of the hyperdominant procedure as a specific procedure of selection of the "best subset" from the displayed set of options.

Thus by applying the "better-worse" relation not to individual options, but to sets ("arrays") of options regarded as an entity, we shall obtain a hyperdominant selection procedure (on the assumption that the correctness condition is satisfied). Example 10 can serve as an illustration.

Example 10. Selection of Optimal Arrays. In this case the structure σ will be in the form of a scalar function $\Phi(Z)$ (hyperscale), which is defined on the set $2^A \setminus \varnothing$ (the family of all nonempty subsets of A) and which assigns to each nonempty $Z \subseteq A$ a number $\Phi(Z)$, which is the "estimate" of the array Z. The rule π consists in finding a nonempty set $Y \subseteq X$ with a maximum value of the estimate Φ:

$$C(X) = Y, \text{ where } \Phi(Y) = \max_{Z \subseteq A} \Phi(Z). \qquad (14)$$

The correctness of this definition is ensured by the single-valuedness of the function Φ: $\Phi(Z') \neq \Phi(Z'')$ for $Z' \neq Z''$ (and also by virtue of the fact that $Z = \varnothing$ in (14) is impossible by definition).

It is easy to see that such a selection procedure of optimal arrays reduces to a particular case of hyper-dominant procedure, i.e., in (12) it suffices to write $Y \mathfrak{D}_5 Z \Leftrightarrow \Phi(Y) \geq \Phi(Z)$.

It could be assumed that also all the effects of "multiple" relations between options, in any case for rational selection procedures, can be reduced to generalized binary relations (hyperrelations) of preference between sets of options, and the procedures themselves could be reduced to hyperdominant selection procedures. But, this assumption is contradicted by Theorem 7, according to which the hyperdominant procedures can be generated only by choice functions belonging to the region $\mathbf{O}$. Thus, e.g., the procedure of aggregate-extremal selection (Example 4) can indeed be always represented by an equivalent hyperdominant procedure (see below), whereas in general the procedures of two-step selection (Example 5) and of tournament selection (Example 6) cannot. Not even every classical pair-dominant selection procedure can be reduced to a hyperdominant procedure (compare Theorems 1 and 7). All the more so it is not always possible to reduce to a hyperdominant procedure the strongly and weakly dominant selection procedures in which multiple relations are formalized in a different manner.

7. Illustrative Examples

Now let us examine how to represent some specific procedures of selection with a certain type of multiple relations between options in the framework of generalized dominant selection procedures. For this purpose let us return to the above procedures of aggregate-extremal selection (Example 4) and of two-step selection (Example 5).

Analysis of Example 4. Aggregate-extremal selection on the basis of a set of scales $\varphi_1(x),...,\varphi_n(x)$ can be represented in the form

$$C(X) = \{y \in X \mid \text{there exists an } i \text{ such that } \varphi_i(y) \geq \varphi_i(x) \text{ for any } x \in X\} \tag{15}$$

or, what amounts to the same,

$$C(X) = \{y \in X \mid \text{there does not exists a set } x^1,...,x^n \text{, such that } \varphi_i(x^i) > \varphi_i(y), i = \overline{1,n}\} \tag{16}$$

(in (16), some of the options $x^1,...,x^n$ may coincide). According to Theorem 3, the procedure of aggregate-extremal selection generates a function $C(X)$ that satisfies both the condition $\mathbf{H}$ and the condition $\mathbf{O}$. It hence follows from Theorem 5 that this procedure can be represented by an equivalent strongly dominant procedure, and from Theorem 7 that it can be represented by an equivalent hyperdominant procedure.

For reduction to a strongly dominant procedure we shall use (16) and define the hyperrelation $\mathcal{D}_2$ in (9) as follows:

$$ZD_2y \Leftrightarrow Z = \{z^1,...,z^n\}, \text{ where } z^1,...,z^n \in A \text{ and } \varphi_i(z^i) > \varphi_i(y), \ i = \overline{1,n}. \tag{17}$$

It is easy to see that the thus-defined structure $\mathcal{D}_2$, yields a strongly dominant procedure (9) equivalent to (16). Let us note that the "prohibiting sets" Z in the definition (17) of this structure have been chosen "economically, " i.e., they contain not more than n elements each (and possibly less, since some of $z^1,\ldots,z^n$ may coincide).

For reduction to a hyperdominant procedure, it is possible, for example, to define the hyperrelation $\mathcal{D}_6$ in (13) only for one-element sets $Z=\{z\}$ and for $Z=Y$ as follows:

$$Z\mathcal{D}_6 Y = \begin{cases} Z=\{z\},\ z\in Y \text{ and there exists an } \mathrm{i}_0 \text{ such} \\ \text{that } \varphi_{\mathrm{i}_0}(z)\geq\varphi_{\mathrm{i}_0}(y) \text{ for any } y\in Y, \\ \text{or} \\ Z=Y \text{ and there exist } y^0,y^1,\ldots y^{\mathrm{n}}\in \mathrm{Y}, \\ \text{such that } \varphi_{\mathrm{i}}(y^0)<\varphi_{\mathrm{i}}(y^{\mathrm{i}}),\ \mathrm{i}=\overline{1,\mathrm{n}}. \end{cases} \tag{18}$$

It is easy to see that the structure $\mathcal{D}_6$ defined in (18) is correct and it yields a hyperdominant procedure (13) that is equivalent to the aggregate-extremal procedure (15). Let us note that here the "prohibition" relation $Z\mathcal{D}_6Y$ does not have the meaning of "preference" of Z to Y; this is particularly obvious in the case of an "economic" definition of the hyperrelation $\mathcal{D}_6$ in (18).

<u>Analysis of Example 5</u>. Two-step selection based on the set of scales $\varphi(x)=(\varphi_1(x),\ldots,\varphi_n(x))$ for the first step and the scale $\psi(x)$ for the second step can be expressed in the form

$$\begin{aligned} C(X)=\{y\in X \mid & \text{ there does not exist an } x\in X \text{ such that } \varphi(x)>\varphi(y), \\ & \text{and if } z\in Z \text{ is such that } \psi(z)>\psi(y), \text{ then there exists} \\ & \text{an } u\in X \text{ such that } \varphi(u)>\varphi(z)\}. \end{aligned} \tag{19}$$

Here $\varphi(x)>\varphi(y)$ is a "vector inequality" for the set of scales φ in the sense in which it occurs in the definition of vector optimization selection (Example 2). The nonfulfillment of this inequality will be written in the form $\varphi(x)>\varphi(y)$. Similarly, for the sake of uniformity, we shall write [27] the nonfulfillment of the inequality $\psi(z)>\psi(y)$ in the form $\psi(z)\not>\psi(y)$. As we noted above, such a two-step procedure generates a choice function that satisfies the condition $\mathbf{C}$. It hence follows from Theorem 6 that such a procedure can be represented by an equivalent weakly dominant selection procedure.

[27] In this way we make sure that instead of taking ψ as a scalar, we can take it as a vector of scales, without changing any of the results.

For this purpose let us define the hyperrelation $\mathcal{D}_3$ in (10) by writing $y\mathcal{D}_3 Z$ for sets $Z \subseteq A$ of two types, i.e., one-element sets $Z = \{z\}$ and two-element sets $Z = \{u, v\}$. Let us write

$$y\mathcal{D}_3 Z = \begin{cases} Z = \{z\}, \text{ where } \varphi(z) \not> \varphi(y) \text{ and } \psi(z) \not> \psi(y) \\ \qquad (\text{in particular, } z = y), \\ \text{or} \\ Z = \{u, v\}, \text{ where } \varphi(u) \not> \varphi(y),\ \varphi(v) \not> \varphi(y), \\ \qquad \psi(u) \not> \psi(y) \text{ and } \psi(u) \not> \psi(y) \end{cases} \tag{20}$$

(in the latter case it is assumed in (20) that $\psi(v) > \psi(y)$). By comparing (19) with (10) for $\mathcal{D}_3$ taken from (20), we find that the thus-defined weakly dominant procedure (10), (20) is equivalent to the two-step procedure (19). Let s note that the structure of $\mathcal{D}_3$ defined in (20) is "economical," i.e., the "permission" sets are not larger than two-element subsets in X .

8. Conclusion

The above analysis shows that in many cases the meaningful concept of "selection of best options," which in purest form can be formalized as "selection according to a scalar optimality criterion," may exceed the framework not only of scalar-optimization procedures, but also of any classical procedures of selection of options that are "best according to binary comparisons." Even the presence of a procedure of extremization of a scalar function in the definition of a selection process does not yet signify that the selection reduces to a comparison of options (according to their "preference" etc.) based on the strict classical scheme of pair-dominant selection (selection according to preferences). Indeed, on the one hand the options that maximize (or minimize) a certain scalar function can always be found (by virtue of the definition of maximization or minimization) by explicit or implicit binary comparison with the other displayed options. Moreover, the very concept of "variation," which lies at the basis of many extremization methods is nothing else but a specific binary comparison. On the other hand, however, we must take into account the character of the utilized estimating function as a basis for these comparisons, i.e., whether or not it depends on the entire sample (the "context" of the choice).

LITERATURE CITED

1. M. K. Richter, "Rational choice," in: J. S. Chipman et al. (editors), Preferences, Utility, and Demand, Harcourt Brace Jovanovich (1971), pp. 29-58.
2. C. R. Plott, "Axiomatic social choice theory: an overview and interpretation," Am. J. Pol. Sci., 20. No. 3, 511-595 (1976).
3. P. C. Fishburn, Utility Theory for Decision Making, Wiley, New York (1970).
4. B. C. Mirkin, The Problem of Group Choice [in Russian], Nauka (1974), translated by Yelena Oliker, Peter C. Fishburn, ed. Washington, D.C.: Winston & Sons, 1979.
5. P. C. Fishburn, "Should social choice be based on binary comparisons?" J. Math. Sociol., 1, No, 1, 133- 142 (197l).
6. C. R. Plott, "Path independence, rationality, and social choice," Econometrica, 41, No. 6, 1075-1091 (1973).
7. K. J, Arrow, Social Choice and Individual Values, Yale Univ. press, New Haven (1963).
8. A. K. Sen, Collective Choice and Social Welfare, Holden-Day, San Francisco (1970).
9. K. Suzumura, "Rational choice and revealed preference," Rev. Econ. Stud., 44, No. 1 (133), 149-158 (1976).
10. J. von Neumann and O. Morgenstern, Theory of Games and Economic Behavior, Princeton Univ. Press. Princeton (1947).
11. J. C. Harsanyi, "An equilibrium-point interpretation of stable sets and a proposed alternative definition," Manag. Sci., 20, No. 11, 1472-1495 (1974).
12. H. Chernoff, "Rational selection of decision functions," Econometrica, 22, No. 3, 422-443 (1954).
13. A, K, Sen, "Choice functions and revealed preference," Rev. Econ. Stud., 38, No, 3 (115), 307-317 (1971).
14. K. J. Arrow, "Rational Choice functions and orderings," Economica, 26, No. 102, 121-127 (1959).
15. D. T. Jamison and L. J. Lau, "Semiorders and the theory of choice," Econometrica, 41, No. 5, 901-912 (19?3).
16. F. S. Roberts, "What if utility functions do not exist?" Theory Decision, 3, No. 2, 126-139 (1972).
17. D. H. Blair, G. Bordes, J. S. Kelly, and K. Suzumura, "Impossibility theorems without collective rationality," Econ. Theory, 13, No.3, 361-379 (1976).
18. F. Harary, Theory of Graphs, Addison-Wesley, Reading, Mass. (1969).
19. J. Wilhelm, "Generalized solution principles and outranking relations in multicriteria decision making," European J. Oper. Res., 1, No. 6, 376-388 (1977).