PREDECISIONS

Gunter Menges

Econometric Research Program
Research Memorandum No. 71
January 18, 1964

Princeton University
Econometric Research Program
92-A Nassau Street
Princeton, N. J.

ABSTRACT

This paper departs from the question as to in which phase of the solution of a decision problem the "tautological point" will be arrived at, after which the solution of the decision problem on hand becomes a matter of algorithm.

The first step is to formalize the general decision model in a way which makes its constituent stand out as clearly as possible. The constituent of a decision problem is defined, discussed, and is shown to provide a complete characterization of the decision problem.

Next the space of the constituent is introduced and the principle of accomodation is formulated, according to which the constituent and its space are to be selected in such a way that the discrepancy between the decision model and the decision problem to be solved is as small as possible. This principle is accomplished through a predecision rule.

Ways of fixing the constituent are discussed. It is demonstrated that numerous predecisions (of a more or less arbitrary type and according to more or less objective and formalizable criteria) must be made, and ways are indicated, how they can be made.

Foreword

During the months of September and October 1964 I was a Visiting Research Associate at the Econometric Research Program of Princeton University.

It was in this period of time that this contribution was written. It arose from a stimulating talk I had with Oskar Morgenstern.

The original German manuscript was translated in the Seminar for Theoretical Statistics and Econometrics of the University of the Saar.

Saarbrücken, December 1964

Günter Menges

Table of Contents

PREDECISIONS

I. Introduction

"These new machines make decisions, the old ones merely performed work", Oskar Morgenstern recently wrote referring to the "high-speed digital electronic computers". ([9] , P. 584). Certainly Oskar Morgenstern cannot be suspected of mystifying the "electronic brains". His statement, however, raises the question as to which phase of the process of solving a decision problem the "tautological point" will be arrived at, after which the solution of the problem is merely a matter of the use of an algorithm. Is the "tautological point" prior or posterior to, or does it coincide with, the handing over of the problem on hand to the machine? Or to put it this way: When is the decision model specified far enough as to allow its being handed over to the formal solution calculus, and how can it be specified, i.e., what is the specification process like?

This paper does not intend to give a final answer to this complex of questions. Nevertheless, I should like to contribute some considerations on these problems which are obviously very important for the practical application of decision theory. The fragments mainly concentrate upon the concept of predecision.

II. The Model

A decision problem generally consists in selecting, according to a rule which is designated here as $V_A^{(1)}$, an action a^* , which is in a certain sense optimal, from a given set A of possible actions a .

Let us write

(1) $$a^* = V_A^{(1)} \langle A \rangle ,$$

expressing that the action a^* has been selected from A according to the rule $V_A^{(1)}$. Let a σ-algebra Σ_1 be given over the space A of actions a, making A a measurable space (A,Σ_1). Probability measures δ can then be defined on (A,Σ_1). Let Δ stand for the set of all probability measures δ over (A,Σ_1). These measures assign a probability to every element of Σ_1.

The decision problem can now be formulated under stochastic aspects: From a given set Δ of possible probability distributions δ over A an optimal probability distribution δ^* is to be found according to the rule $V_\Delta^{(1)}$. By analogy with the above equation, let us write

$$\delta^* = V_\Delta^{(1)} \langle\Delta\rangle . \tag{2}$$

According to the probability distribution δ^* the individual actions $a \in A$ are to be selected with the probabilities assigned to them by δ^*.

(1) is a special case of (2) if δ^* assigns the probability 1 to some $a^* \in A$.

Let the combination of A with the set of all probability distributions on A be symbolized by $[A,\Delta]$. In order to make the rules $V_A^{(1)}$ and $V_\Delta^{(1)}$ more concrete we assume given in addition to A a certain set, Ω, of states of nature, θ.

By specifying a σ-algebra, Σ_2, over Ω it is possible to assign, by analogy with A, to the set of states Ω the class Λ of all probability distributions λ defined on Ω. These probability distributions are named <u>a priori</u> distributions on Ω.

The combinations of Ω with a class $\Lambda^* \subset \Lambda$ of <u>a priori</u> distributions on Ω is denoted by $[\Omega,\Lambda^*]$.

Let a sample τ from a sample space $\mathfrak{T}$ provide the possibility of obtaining information about Λ^*. τ is distributed according to a fixed, but unknown, distribution $\bar{\mu}_\theta$. Information provided by τ can be utilized to estimate the unknown a priori distribution τ on Ω and - via the decision function - to reduce the degree of uncertainty. The combination of the sample space $\mathfrak{T}$ with one of its elements τ is denoted by $[\mathfrak{T},\tau]$.

A mapping of the sample space $\mathfrak{T}$ into the action space A is designated as e_A, a mapping of the sample space $\mathfrak{T}$ into the space Δ of probability measures δ on A as e_Δ:

$$e_A : \mathfrak{T} \to A , \quad e_\Delta : \mathfrak{T} \to \Delta .$$

The mappings e_A and e_Δ are called decision functions or strategies. The former are called nonrandomized, the latter randomized decision functions. The set of nonrandomized decision functions is denoted by D_A, the set of randomized decision functions by D_Δ. Each decision function $e_A \in D_A$ or $e_\Delta \in D_\Delta$ assigns uniquely a so-called final decision $a \in A$ or $\delta \in \Delta$ respectively, to each element $\tau \in \mathfrak{T}$:

$$e_A(\tau) = a \in A, \quad e_\Delta(\tau) = \delta \in \Delta .$$

Henceforth we consider as final decisions only the elements $\delta \in \Delta$, i.e., only randomized decision functions e_Δ and their space D_Δ. For the sake of briefness, let us use e and D for e_Δ and D_Δ, respectively.

Let a bounded real-valued function be given over the Cartesian product space $D \times \Omega$ (i.e., the set of all pairs (e,θ)):

$$r: D \times \Omega \to R \text{ (the set of real numbers).}$$

This function is the risk function; its values, the risks, are denoted by $r(e,\theta)$. Furthermore, let a so-called decision criterion, K, be given by which one can determine the e for which the risk $r(e,\theta)$, in a certain sense which is incorporated in K, is optimal.

With regard to the existence of such e, account must be taken of some mathematical requirements which, however, I am not going to treat here. See instead [2]. Concerning the possible decision criteria I refer to [8] and [10].

The rule $V^{(1)}$ (or, to be exact, $V_{\Delta}^{(1)}$), outlined above, can now be defined more precisely:

Let e^* be that decision function out of D for which the risk $r(e,\theta)$ becomes optimal in the sense of the decision criterion K. This decision function is called an optimal decision function or optimal strategy. The rule $V^{(1)}$ now consists in selecting as a final decision that $\delta^* \in \Delta$ which is assigned to the sample outcome τ by the optimal strategy e^*. With regard to K, δ^* is an optimal final decision. In this sense, the rule $V^{(1)}$ is an optimizing rule. The optimal final decision δ^* results from applying the optimal strategy e^* to the sample outcome τ, and e^*, in turn, results from the application of the decision criterion K to the risk function r. This is symbolized by:

(3)
$$\delta^* = e^* \quad (\tau)$$
$$e^* = K \quad [r] .$$

The optimal final decision δ^* is given if the five-dimensional vector

$$\zeta = (\xi_1, \ldots, \xi_5)$$

with its components

$$\xi_1 = [A,\Delta], \quad \xi_2 = [\Omega,\Lambda^*], \quad \xi_3 = [\mathfrak{T},\tau], \quad \xi_4 = r, \quad \xi_5 = K$$

is uniquely specified. Let us call the vector ζ the constituent of the decision problem; it characterizes the decision problem completely, and the final decision depends only on this constituent:

$$\delta^* = \delta^*(\zeta) = \delta^*(\xi_1,\ldots,\xi_5) \quad (4)$$

III. The Accomodation

We next introduce the space Π of the constituent ζ. Let

Ξ_1 be the set of all action spaces, combined with the respective class of all possible probability distributions, i.e., $\Xi_1 = \{[A,\Delta]\}$;

Ξ_2 the set of all $[\Omega,\Lambda^*]$, i.e., every possible state space Ω, is combined with every possible subset Λ^* of the respective class of distributions Λ, and considered an element of Ξ_2, i.e., $\Xi_2 = \{[\Omega,\Lambda^*]\}$;

Ξ_3 the set of all sample spaces $\mathfrak{T}$, combined with each of its elements τ, i.e., $\Xi_3 = \{[\mathfrak{T},\tau]\}$;

Ξ_4 the set of all risk functions r over $D \times \Omega$, i.e., $\Xi_4 = \{r\}$;

Ξ_5 the set of all decision criteria K, i.e., $\Xi_5 = \{K\}$.

Finally, let $\Pi = (\Xi_1,\ldots,\Xi_5)$ denote the space of the constituent $\zeta = (\xi_1,\ldots,\xi_5)$.

As is seen easily the problem is to fix the constituent ζ in Π. The constituent ζ of the practical decision problem on hand being fixed and provided that the programming is appropriate and that the solution exists, the machine can find δ^* according to the rule $V^{(1)}$. Once the constituent has been fixed, the further, final solution of the problem is merely a matter of using the algorithm: Obviously the "tautological point" is marked by the fixation of ζ in Π ! But how can ζ be found?

We first write

$$\zeta = V^{(2)} < \Pi > = (V_1^{(2)} < \Xi_1 >, \ldots, V_5^{(2)} < \Xi_5 >), \tag{5}$$

in order to express that the constituent ζ is fixed in Π by a rule $V^{(2)}$. $V^{(2)}$ is called the predecision rule.

Generally spoken, its function is to bring about an accomodation between the general decision model as described in II. and the given practical decision situation. Hence the selection of ζ from Π obeys an accomodation principle: The discrepancy between the model and the concrete situation should be as small as possible!

In order to solve this approximation problem, one might, first of all, think of using a constituent η which, by the relation

$$\zeta = \zeta(\eta) ,$$

determines an "optimal" constituent ζ. Where this way is feasible, one will follow it, maybe even in such a way that η, for its part, is determined by a constituent η^*, and this one in turn, by η^{**}, etc. Unfortunately this way is in general not feasible (1) because the predecision rule $V^{(2)}$ is in most cases only partially objective and, consequently, to a large extent outside the realm of exact mathematical formulation and moreover, (2) because the discrepancy between model and reality, which is to be minimized, is in general not measurable.

How then can ζ be found? I am afraid that the answer to this question necessarily runs as follows: In general only with the aid of more or less arbitrary, more or less subjective propositions or - decisions! In fact, all our previous considerations lead to the conclusion that the "real" or final decision problem is merely a problem of secondary importance, a problem of arithmetic.

One could even go so far as to say: it is a spurious problem. As rule, the genuine problem will be to make decisions which lead to a constituent from Π to be fixed! Here a decision effort has to be made which the machine cannot make.

I shall now try to give an idea of these predecision problems which, in principle, everybody who wants to apply decision theory in practice is confronted with. The following considerations will also be of use for the (certainly not final) clarification of these questions:

(1) Of what sort is the discrepancy between model and reality?

(2) Of what sort are the (more or less) subjective propositions which fix the constituent?

(3) In which respect and in which parts can the predecision rule be made objective and formulated in exact mathematical terms?

(4) In which cases can the "tautological point" be advanced into the domain of predecisions? And in close connection with this question:

(5) In which cases can the "optimal" constituent ζ be determined by means of a constituent η ?

IV. The Predecisions in Detail

a) Predecisions about $[A,\Delta]$ and $[\Omega,\Lambda^*]$

As for the predecision about these two components, the main difficulty is the enumeration of all actions and states of nature which are essential for the decision problem. This enumeration first of all requires very exact knowledge about the environmental factors, the "milieu" of the decision problem on hand. Exact knowledge of the true aims of the decision maker are also required, as well as of the position of the decision situation before the calculus is applied. Consequently, these predecisions consist in making a relevance statement on the possible actions and states of nature.

The decision maker will take account of only those actions and states which are pairwise disjoint, and he will also take into account all a priori restrictions resulting from his technical equipment, e.g., from the computer at his disposal.

From a material point of view other restrictions will limit his action space. Perhaps some actions though possibly very favorable ones in other respects, cannot be realized for financial reasons. Further restrictions may arise from social or personal obligations and force one to sacrifice potentially favorable actions. The decision maker may also consider favorable actions as inopportune for they might bring him into conflict with the law. Among the states he will enter those into the calculus which are relevant in the sense that the decision might be different if they were missing. I imagine that methods of linear or nonlinear programming may contribute to the solution of these predecisions, e.g., by helping to reduce the sets A and Ω to a number of essential elements, capable of being more easily handled. In some cases Monte-Carlo-methods may provide good results, and frequently methods of statistical analysis, e.g., regression or correlation techniques, will also be successfully applied.

Another possibility has been pointed out to me by my collaborator Dr. Schneeweiss: Let two possible states θ_1 and θ_2 be given. The decision maker knows positively that two actions a_1 and a_2 are at his disposal, but he conjectures the existence of a third possible action a_3. Let the cost of obtaining information about the existence of a_3 amount to c. The possible actions to be defined now are

$\bar{a}_1$ = select a_1 without taking a_3 into account,

$\bar{a}_2$ = select a_2 without taking a_3 into account,

$\bar{a}_{ij}$ = take a_3 into account and apply a_i if it is actually found, apply a_j, if it is not found; $i = 1,2,3$ and $j = 1,2$.

The possible states now to be defined are

$\bar{\theta}_{1i} = \theta_i$ is given and a_3 is found,

$\bar{\theta}_{2i} = \theta_i$ is given and a_3 is not found.

Of course, the outcome matrix now contains 32 elements instead of the original four, eight of which, however, can immediately be eliminated, since they are related to actions dominated by others, as can easily be seen.

I shall not enter into the particulars of the extremely problematic predicision about the a priori distribution λ or its space Λ^*, since I have already dealt with the problem. See [6], [11] and [1] and IV b below.

b) Predecisions about $[\mathfrak{T},\tau]$

The decision maker has to make predecisions whether a sample should be drawn at all and if a sample is to be drawn, it must be decided beforehand which variable or variables are most characteristic of the states of nature. Consideration must, however, be given to the fact that in certain circumstances a very characteristic variable may not be measurable at all (such as opinions, standards of behavior, the political or social situation, etc.). Furthermore, a predecision must be made as to which sample function is most favorable, whether one should proceed sequentially, etc. What sample size should be chosen (and connected with that: how shall $\mathfrak{T}$ be determined)? Perhaps it is expedient to construct a statistical index of various variables (e.g., when measuring the price level or even some such phenomenon as "weather"). The question as to how far the sample elements are autocorrelated or obey a trend must be considered. Also important is the width of the inductive basis, i.e., to what degree are the conditions under which the sample elements are collected controllable. Fortunately, sampling theory offers numerous

instruments suited to deal with, and solve, these questions. See [3] and the literature quoted there. Moreover, it can be shown that some of the sampling decisions outlined above can be made according to objective criteria. See [3] and [5].

The following simple and expedient method of fixing the distribution λ over Ω or a space Λ^* by means of sampling has been pointed out to me by my collaborator, Mr. Diehl.

Let $(\theta_1, \ldots, \theta_k)$ be the space of states of nature, where the θ_i are possible distributions of a random variable x. The a priori distribution λ over $(\theta_1, \ldots, \theta_k)$ is to be estimated given observations x_i of the random variable x.

Let k simple samples (i.e., samples, the elements of which are identically and independently distributed) of size n be given:

$$\tau_1^{(n)} = (x_{11}, \ldots, x_{1n})$$

$$\vdots$$

$$\tau_k^{(n)} = (x_{k1}, \ldots, x_{kn}) .$$

For each sample $\tau_i^{(n)}$ $(i = 1, \ldots, k)$ the empirical distribution as well as its χ^2-deviation from all states θ_j $(j=1, \ldots k)$ is determined. That state θ_j whose χ^2-deviation from the empirical distribution of $\tau_i^{(n)}$ is smallest is recorded. If several states show the same minimal deviation, all of them are recorded. In this way, we obtain a set of - let us say $m (m \leq k)$ - states θ_i :

$$\theta_{\kappa_1}, \ldots, \theta_{\kappa_m}, \kappa_\mu \in \{1, \ldots, k\} \text{ for } \mu = 1, \ldots, m .$$

From this set of states θ_i the empirical distribution over the space of states of nature is computed, each state being given the weight $\frac{1}{m}$. This empirical

distribution can subsequently be considered as a ("point") estimate of the distribution λ over the space of states.

In addition, a stochastic ϵ-range, covering the distribution λ over the space of states with a given probability ϵ, can be constructed on the basis of this estimate. (The limits of this range can be found using the well-known theorem by Kolmogoroff). The range itself may be considered an estimate of Λ^*; its width can be made to depend on the degree of uncertainty.

I shall not inquire here into the requirements that must be satisfied if one wishes to apply this method; namely, a sufficient sample size and a high degree of stability of λ or Λ^*, respectively. (See, however, [6]).

c) Predecisions about r

In order to formulate a risk function, the first step is to indicate a real-valued so-called outcome function

$$E: \quad A \times \Omega \to R \ ,$$

which is defined and bounded over the Cartesian product space $A \times \Omega$.

This outcome function assigns an outcome $E(a,\theta)$ to each coincidence of an action a and a state θ. The outcomes $E(a,\theta)$ can be interpreted as either gains or losses. In general, the risk function r is derived from the outcome function E in the following way:

The states $\theta \in \Omega$ are characterized by the stochastic variable x which will take values in a space of events X. Every θ induces a distribution μ_θ for x, or a distribution $\bar{\mu}_\theta$ for the sample τ, respectively. These distributions define the outcome expectations for the outcomes E; they are called risks.

If $E(a,\theta)$ are the outcomes over $A \times \Omega$, the risk function r is defined over $D \times \Omega$ by

$$r(e,\boldsymbol{\theta}) = \int_{T} \int_{A} E(a,\boldsymbol{\theta}) \; d\, e\, (\tau) \; d\bar{\mu}_{\boldsymbol{\theta}} \quad .$$

In case the distribution λ on Ω is known, Bayes' risk function r^* is set in addition. It is defined by

$$r^*(e,\lambda) = \int_{\Omega} r(e,\boldsymbol{\theta}) \; d\lambda \quad .$$

If the outcomes $E(a,\boldsymbol{\theta})$ are not numerically specified beforehand, a preference ordering must exist in order to enable the decision maker to make a choice between the competing a, that is to say, first, the decision maker must be able to say which of two outcomes he prefers, or that he is indifferent to them, and second, if he does not prefer one outcome to another and does not prefer the latter to a third one, the first outcome is not preferred to the third one.

The predecisions about E and r are mainly to decide what amount of effort, money, and time should be spent on collecting information about the outcome function E. Besides, instructions must be given (predecided!) for the case that the decision maker cannot set up a preference ordering of $E(a,\boldsymbol{\theta})$. This case is not at all unlikely to happen, since the decision maker must possess a strong, strictly speaking, an unlimited sensitivity with respect to the distances between the individual $E(a,\boldsymbol{\theta})$ in order to be able to rank them. If he does not possess this ability, his indifference is no longer transitive and the whole decision problem becomes absurd.

Further predecisions become necessary if, instead of the ordinary risk function r or r^*, the actual risk function $\tilde{r}$ or $\tilde{r}^*$, respectively, is subjected to the decision criterion. This should be done whenever samples are drawn, since sampling generally involves costs. The actual risk function is defined as follows:

$$\tilde{r} = r + Ex\,(C) \quad ,$$

C being the cost function, Ex (C) its expectation. The cost function mainly depends on the sample size, but on other factors too, e.g., on the number of stages into which the sampling process is divided, or, in unfavorable cases, on the outcome $E(a, \theta)$ itself. I do not intend to go into the details of all this. I only want to remark that one must explicitly know, of course, the cost function (if samples are drawn) in order to be able to solve the decision problem. And before it is known a series of predecisions will have to be made. If, finally, the cost function is known, another difficulty arises from the fact that some of its arguments are random elements which makes it necessary to determine the expectation Ex (C) . Its determination requires a rather intimate knowledge of the cost structure in question as well as of the stochastic behavior of the variables which determine C . Whether this knowledge should be gathered or not will also be a difficult decision to make. Only the fact that predecisions about the risk function can frequently be made according to measures which are (to a certain extent) objective, is somewhat comforting.

The space D of decision functions requires no predecision comparable to those about the other determinants. It is uniquely given by Δ and T and confined, by r , to the so-called complete class of admissible decision functions.

d) Predecision about K

The predecision as to which K should be selected from the class of decision criteria in a given problem depends essentially on the appraisal of the opponent. To some extent this is even an absolute criterion of selecting K . On the basis of this criterion, the establishment of a type of "quality ranking" of the opponent might be an objective method of making a predecision about K .

The pessimism inherent in Wald's minimax criterion is appropriate whenever incorrect decisions involve heavy losses in the struggle with a dangerous

opponent, and also whenever statistical decisions are made and an unfavorable distribution λ must be reckoned with. On the other hand, an optimistic criterion can be tolerated if the opponent is harmless and/or if it is possible to acquire information about his behavior.

If one has exact knowledge of the distribution λ, it will be most reasonable to apply Bayes' decision criterion, i.e., to select as optimal decision function e^* that one for which

$$r^*(e^*,\lambda) = \min_{e \in D} r^*(e,\lambda) .$$

In case Bayes' decision criterion is not applicable, either because λ is unknown or because its application seems inopportune for other reasons, e.g., because heavy losses must be expected in the case of incorrect decisions or because the conditions of the decision problem must be considered as unstable (a problem which I shall be satisfied to merely call attention to at this point), then the application of the adaptation criterion is recommendable (see [7], [8] and [10]). This criterion allows one to give explicit consideration to the given degree of uncertainty and instability, and to the decision maker's various possibilities of modifying nature. It is also suited to serve as a "quality ranking" for the appraisal of the opponent. Its extremes are Bayes' solution on the one hand and Wald's minimax solution on the other hand. In practice, it will always result in certain mixtures of Bayes' and Wald's solutions.

Of course, the application of the criteria proposed by Chernoff, Savage, Hurwicz and Hodges-Lehmann may also be considered. The Hodges-Lehmann criterion is a special case of the adaptation criterion. I consider, however, the application of the three "subjective" criteria proposed by Chernoff, Savage and Hurwicz as less recommendable. But this is perhaps a matter of personal taste.

Anyway, there can be no doubt that the selection of the decision criterion, too, requires a predecision in the sense of the accomodation principle.

Summary and Concluding Remarks

I. This paper departs from the question as to which phase of the solution of a decision problem the "tautological point" will be arrived at, after which the solution of the decision problem on hand becomes a matter using an algorithm.

II. In order to answer this question, the first step is to formalize the general decision model in a way which makes its constituent stand out as clearly as possible. The constituent of a decision problem is designated as the vector ζ given by:

$$\zeta = ([A,\Delta], [\Omega,\Lambda^*], [\Upsilon,\tau], r, K);$$

A = the set of actions a,

Δ = the set of all probability measures δ over (A,Σ_1),

Ω = the set of states of nature θ,

Λ^* = a subset of the set Λ of all probability distributions defined over Ω,

Υ = sample space,

τ = sample,

r = risk function,

K = decision criterion.

The constituent is discussed, and is shown to provide a complete characterization of the decision problem.

III. Next the space Π of the constituent is introduced and the principle of accomodation is formulated, according to which ζ and Π are to be selected in such a way that the discrepancy between the decision model and the decision

problem to be solved is as small as possible. This principle is accomplished through a predecision rule.

Obvious algorthmic ways of solving this problem turn out to be infeasible. Other ways of fixing the constituent are discussed subsequently.

IV. Numerous predecisions (of a more or less arbitrary type and according to more or less objective and formal criteria) must be made, for example:

a) Concerning $[A,\Delta]$ and $[\Omega,\Lambda^*]$:

Which are the possible and which are the "essential" actions? What are the "true" aims of the decision maker? Which are the "relevant" states of nature? - A relatively simple and expedient method of making predecisions is pointed out. (Predecisions about λ and Λ^*, respectively are not discussed).

b) Concerning $[\Upsilon,\tau]$:

Is a sample to be drawn? Which variable(s) is (are) characteristic of the states of nature? Which sample function is most favorable? Etc. - A convenient method of fixing λ and Λ^*, respectively with the aid of τ, is indicated.

c) Concerning r:

What amount of time and money should be spent on gathering information about the outcome function? How should one proceed if no preference ordering at all is given among the outcomes? How exact and how detailed should the information about the cost function be? How, if necessary, should this information be obtained? Etc.

d) Concerning K:

Which decision criterion is to be applied: Bayes' criterion; Wald's minimax criterion, or one of its variants; or the adaptation criterion? (The author

considers the latter to be the most objective criterion).

This paper tries to show the "true", the genuine, decision problem is hardly to find a final decision. Machines can find this final decision once the constituent of the problem on hand has been fixed. To fix the constituent is the "true" decision problem! This requires a number of decisions (which we have called predecisions), some of which are very complicated, and which can at best be handled by machines only partially. The true effort of decision making is prior to the decision problem as commonly defined and conceived of. It is the predecisions which assemble the determinants of the problem on hand.

References

[1] Tschuprow, A., 'Ziele und Wege der stochastischen Grundlegung der statistischen Theorie', (Aims and Ways of the Stochastic Foundation of Statistical Theory), Nordisk Statistik Tidsskrift, 1924, 3. Pp.433 - 493.

[2] Wald, A., 'Statistical Decision Functions', (J. Wiley and Sons, Inc., New York, 1950).

[3] Menges, G. and Behara, M., 'Das Bayessche Risiko bei sequentiellen Stichprobenentscheidungen' (Bayes Risk of Sequential Sampling Decisions, Statistische Hefte, 1962, 3, pp. 39-61).

[4] Savage, L.J., et al., 'The Foundations of Statistical Inference', (Methuen and Co., Ltd., London, 1962).

[5] Menges, G., and Behara, M., 'Einige grundsätzliche Betrachtungen über prozessuale Entscheidungen unter Ungewissheit' (Some Fundamental Considerations About Processual Decisions under Uncertainty), Zeitschrift für handelswissenschaftliche Forschung, 1962, 14, pp. 483-494.

[6] Menges, G., 'Gedanken zur Frage der Stabilität statistischer Entscheidungen' (Thoughts on the Problem of Stability of Statistical Decisions), Metrika, 1963, 6, pp. 84-94.

[7] Menges, G., and Behara, M., 'On Decision Criteria under Various Degrees of Stability', Journal of the Indian Statistical Association, 1963, 1, pp. 185-195.

[8] Menges, G., 'Kriterien optimaler Entscheidungen unter Ungewißheit', (Decision Criteria Under Uncertainty), Statistische Hefte, 1963, 4, pp. 151-171.

[9] Morgenstern, O., 'Pareto Optimum and Economic Organization' in: Systeme und Methoden in den Wirtschaftsund Sozialwissenschaften, Erwin von Beckerath zum 75. Geburtstag; herausgegeben von N. Kloten, W. Krelle, H. Müller, F. Neumark. Tübingen 1964, pp. 573-586. Das Zitat findet sich auf Pp.584.

[10] Menges, G., "The Adaptation of Decision Criteria and Application Patterns", Proceedings of the 3rd International Conference On Operational Research, Oslo 1963, Paris 1964, pp. 585-594.

[11] Menges, G., 'Über einige Fundamentalprobleme der statistischen Entscheidungstheorie', (On Some Fundamental Problems of Statistical Decision Theory), to be published.