Kybernetika

Ivan Kramosil

Mechanized experiment planning in automaton-environment systems

Kybernetika, Vol. 13 (1977), No. 4, (225)--244

Persistent URL: http://dml.cz/dmlcz/125063

KYBERNETIKA — VOLUME 13 (1977), NUMBER 4

Mechanized Experiment Planning in Automaton-Environment Systems

IVAN KRAMOSIL

The first part of this paper is devoted to a very intuitive and general discussion concerning the role of experiment planning in cognitive automaton-environment systems. Then we propose and ivestigate a method how to use the formulas, sampled at random during the statistical testing of the formula claiming the goal in question to be accessible, in order to propose an experiment verifying the formula under the condition it is statistically estimated to be valid. It is proved, under some more conditions, that the cost of the experiment proposed in this way is the smaller the better is, from the statistical point of view, the implemented statistical deducibility testing procedure.

1. THE ROLE OF EXPERIMENT PLANNING AND EXECUTING IN AUTOMATON-ENVIRONMENT SYSTEMS

There is a close connection between this paper and the papers [7], [8]; in a way, this paper can be considered as their continuation. It is caused by the fact that the theoretical model and framework of our present considerations will be the same as in the two papers mentioned above. Again, in the focus of our attention will be an automaton situated in the Euclidean plain, moving from one integer co-ordinated point to another one, observing the properties of such points or deriving them from the data being at its disposal and, last but not least, fulfilling some tasks in this environment. Because of a rather large autonomy of its behaviour and its goal-tending activity such an automaton is, in fact, more close to what "robot" is called in scientific literature than to the mathematical and theoretic interpretation of this expression, however, not wanting to involve some philosophical or anthropomorfic associations, or, on the other hand, some technically oriented problems or questions, we shall continue in our using of the expression "automaton" as in [7] and [8].

In those two papers our main attention was concerned to the way in which the automaton constructs its formal representation of the environment in order to classify somehow the quality of this formal representation and, if possible, to improve it. During the considerations leading to the notions and results contained in [7], [8] we have emphasized some "absolute" demands imposed to a formal representation of the environment (i.e. consistency, logical independence, semantical completeness, etc., see [8]) and we have neglected, to a degree, the goal-oriented character of all the automation activity including the process of formal representation formation. This point of view has given a number of interesting results and pourred a new light into the procedures of formal representation making, however, in the case of a more detailed study the goal-tending aspect of the automaton activity must be taken into consideration. In fact, an *internal* formal representation of the environment is always a secondary matter in the sense that it plays always the role of a mean, even if the most important one, never the role of the final goal. Considering the classical case when the goal consists in transforming the objects in the environment into an a priori defined configuration the submissive role of the formal representation with respect to this goal is quite clear. Nevertheless, even in the case of an automaton the goal of whose activity is to describe an unknown environment (the surface of Moon, the sea depths etc.) the goal consists, in fact, in an action offering some external formal representation to the user or designer, hence, the internal formal representation plays, again, an intermediate role.

At the level of "reasoning" and decision making the automaton activity is concentrated to the problem how to find or form a plan leading from the present situation to a goal situation, i.e. to a situation satisfying some a priori given goal conditions. In order to decide whether a goal is reachable from the present situation and, in the positive case, to find an appropriate plan some knowledge concerning the automaton abilities as well as some knowledge concerning the present state of the environment is necessary; in case the automaton is not the only active factor in the environment the knowledge concerning the environment must be up-dated from time to time.

Not penetrating into details as far as the way in which a plan is obtained is considered we must admit, in general, the possibility of a failure of this effort, i.e. the automaton is not able to obtain an appropriate plan to reach the goal. There are, in principle, two causes of such a failure; first, no goal situation is accesible from the present (initial) one by the mean of the operators being at the automaton disposal, second, there is a way leading to a goal situation, however, the data being at the automaton disposal together with its abilities to handle with these data are not sufficient enough to find a desired plan. In general, in case of a failure the automaton is not able to decide which of the two possibilities has occured. And it is just the instance when the experiment planning and executing begin to play a role in our considerations. *Experiment*, in what follows, is considered to be nothing else than a rather autonomous sub-goal tending automaton activity the aim of which is to enrich the knowledge about the environment being at the automaton disposal to a degree high enough to find a plan for the given goal – or to find that this goal is not accessible. The "autonomy" of automaton behaviour during the experiment planning and performing consists in fact that now, for a while, the priorities are reversed, an appropriate change of the formal representation of the environment becomes the sub-goal and some appropriate "physical" actions and their choice are subjected to this sub-goal. As late as after having terminated the experimenting the priorities take again their "usual" order and the (main) goal of the automaton dominates again its further activity.

There are two levels of automaton activity on which experimenting and planning of experiments plays a role. The first one is very close to the level being already technically accessible by the means of our days and is represented, e.g., by an automaton exploring the surface of Moon or Mars. The main goal of such an automaton is to communicate to the designer or to an organizing centre on the Earth as many pieces of interesting information concerning the unknown environment as possible. An experiment, here, may consist in submitting a small stone, met at random by the automaton, to the influence of some chemical substances in order to deduce some data concerning the chemical properties of the stone. In such a case the connection between the internal formal representation and the goal is the most simple and determined a priori – the goal is to enrich this representation as possible and to communicate this representation to the Earth.

The second level of automata experimenting is more sophisticated, however, the continually increasing complexity of problems which automata will have to solve as well as increasing and descriptively already not accessible complexity of the environment in which these problems will be solved require to investigate this case as soon as possible at least from the theoretical point of view. Here we mean automata "intelligent" enough to choose, in the case of a failure of the direct plan searching, an appropriate experiment which, having been executed, gives new data sufficient in order to find, after all, the desired plan or to say that such a plan does not exist. In other words, such an automaton is able, during its activity, to control the instantaneous priority of the goal itself or of the formal representation in order to achieve, after all, the main goal, if it is principally possible.

During all this paper we shall consider experiment planning and experiment executing as two principally different and one after another following automaton activities, i.e., first a plan is made or found, then the plan is executed. Such a viewpoint agrees with that most often used in papers dealing with automated planning and plan executing, however, it is far from being the only possible. In [9] we can find some interesting ideas concerning the possibility of mixed sequences of automaton actions, some actions being planning the next step, some being of executive character – and the result of actions influences immediately the next planning actions (there is a possibility of a feedback). Supposing the environment to be of non-deterministic (or random) character we must admit that some actions involving the environment can lead to unexpected consequences, hence, a feedback between the actual results of an action and the decision which action should be the next is necessary in order to ensure the final goal achieving. This idea leads immediately to another possibility how to understand experiment planning and executing in automaton-environment systems. We could understand by "experiment" any action or sequence of actions (or a more rich structure of actions) leading with a positive probability to the desired goal. Executing these actions we are experimenting – either the goal is reached, in other words saying, the experiment has been successful, or the goal has not been reached and we have to plan what to do in future. This is, besides other possible examples, also the case of plans consisting in repeatedly executed actions under the condition we know that, after all, such a repeating must lead to the desired goal. Driving a nail home we do not know, a priori, how many strokes are necessary to do so, so we can consider any stroke to be an experiment which, with a probability, can lead to the goal and, until the goal is reached we repeat this "experiment executing" as we know, that, after all, this will give the desired result. This way of understanding the notion "experiment" in automaton-environment systems is developed and investigated in more details in [2] and it is why we limit ourselves, here, to the usual way of reasoning when experiment planning proceeds experiment execution.

Before closing this introductory part of our work we would like to mention the fact that there is a book ([1]) dealing with the problems of experiment planning and executing in connection with the automata theory. The difference between our approach and that of [1] consists in the fact that in [1] "automaton" is understood in its purely theoretical sense known from the automata theory and without any cognitive and autonomous aspects ascribed to automata investigated in this paper or in [7] or [8].

2. STATISTICAL DEDUCIBILITY TESTING AND EXPERIMENT PLANNING

After these rather general conciderations concerning the role of automated experiment planning and executing in automaton-environment systems let us return to the concrete example of such systems introduced and investigated in [7] and [8]. Hence, environment means, in what follows, the set I^2 of all integer-coordinated points of the two-dimensional Euclidean space. These points may possess various properties in various situations and the environment is supposed to be of dynamical character in the sense that the properties possessed by a point in different situations may differ. Situations are identified with all the possible past-histories of the automaton-environment system, i.e. situations are identified with finite sequences of automaton actions; the set of all situations is denoted by S. The properties of points in I^2 are supposed to be enumerated (Gödelization) in such a way that to any $i \in \mathcal{N} = \{1, 2, \ldots\}$ a property corresponds. These assumptions give the possibility to introduce a ternary three-sorted predicate V such that $V(a, i, s)$, $a \in I^2$, $i \in \mathcal{N}$, $s \in S$ is interpreted as "the point a possesses the i-th property in the situation s".

The automaton is supposed to be able to execute the four following types of actions: *moves* into any point neighbour to the present position of the automaton, *operations*, consisting in changing some properties of points in I^2, *observations*, i.e. the automaton has some capabilities (limited, as a rule), to observe properties of points and finally, *deductions*, including *actualizations* (updating), enabling to deduce some new assertions concerning the space I^2 from those already known. A more detailed description of the supposed automaton abilities can be found in [7] and is not repeated here as this informal description seems to be sufficient for the aims of this paper.

Let $\mathscr{L}$ be the formalized language built over the atomic formulas of the type $V(a, i, s)$ by the mean of usual propositional connectives and quantifiers (which can bound any type of indeterminates). Let $X = X(s)$, $Y = Y(s')$ be two formulas from $\mathscr{L}$ such that there is just one indeterminate of the sort "situation" occurring freely in X (Y, resp.), namely s (s', resp.). The pair $\langle X, Y\rangle$ of formulas is then called *problem* and this formal notion is connected with the following intuition: if the present situation s_0 satisfies X (i.e. if $X(s_0)$ is true) the problem is to transform, using the operators being at the automation disposal, the environment in such a way that the new situation, s', say, satisfies Y. The goal situation is, hence, defined as any situation in which an a priori given formula is valid.

Inside the systems investigated in papers dealing with artificial intelligence and automated problem solving problems are solved in such a way that, first, observations and deductions are used to obtain a number of data describing, to a degree "sufficient enough", the environment, at the second stage these data are used in order to find a plan solving the given problem. In the most simple case plan is a simple sequence of operators the execution of which assures the reaching of a goal situation, some more developed investigations in this field consider branching structures of plans with some further information necessary to decide which branch will be actually followed.

There is, within the framework of mathematical logic, a tool enabling to transform the problem of automated plan formation into a task of purely logical nature, namely the so called *situation calculus* ([11], [12]). Considering the problem $\langle X, Y\rangle$ we ascribe to it the formula $X(s_0) \rightarrow (\exists s')(Y(s'))$ and we ask whether this formula is deducible in the formalized theory based on the language $\mathscr{L}$ the role of axioms of this theory being played by those of an appropriate logical calculus and by the particular data being at the automaton disposal and describing the environment. If $X(s_0) \rightarrow (\exists s')(Y(s'))$ is provable in the theory and a resolution-based theorem-prover gives a formalized proof of this formula then the term (terms, resp.) substituted for s', gives (give, resp.) immediately a linear (branching, resp.) plan solving the problem $\langle X, Y\rangle$.

As a rule, the papers dealing with applications of mathematical logic in artificial intelligence are limited to transformation some AI problems into that of deducibility of some formula or formulas and neglect the questions connected with deducibility deciding. However, it is a well-known fact that, with the exception of some simplest theories, the set of all theorems of a formalized theory is recursively enumerable but not recursive. Hence, the question whether $X(s_0) \rightarrow (\exists s')(Y(s'))$ is a theorem or not is not algorithmically decidable, in general, and the procedure of plan making offered by the situation calculus is not completely defined supposing it is not enriched by a deducibility decision making procedure. On the other hand, the undecidability of formalized theories implies immediately that such a deducibility decision making procedure either is not recursive or it does not give the correct answer for all formulas of $\mathscr{L}$.

This way of reasoning justifies the idea of using some statistically based deducibility testing procedures. Submitting a formula to a statistical deducibility test we obtain the answer (concerning the question whether the tested formula is or is not a theorem) which is not always true, however, we request the probability of error to be "small enough". An appropriate statistical deducibility testing procedure is "better" than any deterministic decision procedure in the sense that the statistical procedure gives to any formula a positive probability to be decided correctly what no deterministic procedure is able to achieve effectively. Hence, the admitting of statistical deducibility tests as a tool used in automated theorem proving and automated plan formation seems to be a step naturally justified by real conditions in which automata are to act and which do not admit some non-recursive decision procedures of semi-algorithmical type. Moreover, in what follows we propose a way how to use an appropriate statistical deducibility testing procedure as a tool for automated experiment planning.

It is not the aim of this paper to investigate how an appropriate statistical deducibility testing procedure can be constructed, some information can be found in [10] or [5], some further references also in [6]. From the formally descriptive point of view which is sufficient for the sake of this paper we can say the following.

Let $\langle \Omega, \mathscr{S}, P \rangle$ be a probability space (i.e. Ω is a non-empty set, $\mathscr{S}$ is a σ-field of subsets of Ω and P is a probabilistic measure, defined on $\mathscr{S}$), let X be a parameter space (multidimensional, in general). Let $\{\mathfrak{t}, \mathfrak{f}\}$ be a two-elemented set the interpretation of whose elements are "provable" for $\mathfrak{t}$, "unprovable" for $\mathfrak{f}$, let T be a mapping of the Cartesian product $\mathscr{L} \times X \times \Omega$ into $\{\mathfrak{t}, \mathfrak{f}\}$ such that for any $a \in \mathscr{L}$, $x \in X$,

$$\{\omega : \omega \in \Omega,\ T(a, x, \omega) = \mathfrak{t}\} \in \mathscr{S}\,.$$

Then the mapping T is called *statistical deducibility testing procedure* and the random event $\{\omega : \omega \in \Omega,\ T(a, x, \omega) = \mathfrak{t}\}$ represents the result consisting in proclaiming the formula $a \in \mathscr{L}$ to be a theorem under the condition that the parameter of the statistical testing procedure equals x (and dually for $\{\omega : \omega \in \Omega,\ T(a, x, \omega) = \mathfrak{f}\}$).

As explained above, the possibility that a is a theorem ($a \in \mathscr{T} \subset \mathscr{L}$, in symbols) and, at the same time, $T(a, x, \omega) = \mathfrak{f}$ or vice versa cannot be excluded, in general, and the probabilities of these two possible errors serve as criteria measuring the qualities of such a testing procedure T. In order to be able to define the appropriate conditional probabilities (and to eliminate the dependence of our results on a particular formula a) we suppose to have at our disposal a random variable α, sampling the formula which is to be tested, i.e. α maps Ω into $\mathscr{L}$ in such a way that for any $a \in \mathscr{L}$

$$\{\omega : \omega \in \Omega,\ \alpha(\omega) = a\} \in \mathscr{S}\,.$$

Now, supposing $P(\{\omega : \omega \in \Omega,\ \alpha(\omega) \in \mathscr{T}\}) \neq 0$, $P(\{\omega : \omega \in \Omega,\ \alpha(\omega) \in \mathscr{L} - \mathscr{T}\}) \neq 0$, we can define, for any $x \in X$,

$$PE_1(x) = P(\{\omega : T(\alpha(\omega), x, \omega) = \mathfrak{t}\}/\{\omega : \alpha(\omega) \in \mathscr{L} - \mathscr{T}\})\,,$$

$$PE_2(x) = P(\{\omega : T(\alpha(\omega), x, \omega) = \mathfrak{f}\}/\{\omega : \alpha(\omega) \in \mathscr{T}\})\,.$$

The value $PE_1(x)$ is called the *first type probability of error* and consists in wrong proclaiming a non-theorem to be a theorem, $PE_2(x)$ is called the *second type probability of error* and consists in wrong proclaiming a theorem to be a non-theorem. As a rule, there are reasons for a separate treating and studying the two probabilities and this point of view has some good justification also if statistical deducibility testing procedures are investigated.

Let us limit ourselves to the statistical deducibility testing procedures based on at random sampled extensions, this idea occurred for the first time in [10], see also [5], [7]. Suppose to be given a sequence of mutually independent and equally distributed random variables $\alpha_1, \alpha_2, \ldots, \alpha_N$ (N given a priori, as well as $M \leqq N$) such that for any $a \in \mathscr{L}$, a closed

$$(1) \qquad P(\{\omega : \alpha_1(\omega) = a\}) > 0\,.$$

Having at our disposal an effectively modified theorem-prover, e.g. a resolution-based theorem-prover together with some time and space restrictions we shall try, for any $i \leqq N$, whether $\alpha_i(\omega) \to \alpha(\omega)$ is a theorem provable by the theorem-prover at our disposal or not. If there is at least one $i \leqq N$ such that $\alpha_i(\omega) \to \alpha(\omega)$ *is not* a theorem or if the number of cases in which it *is* a theorem is smaller than M, we proclaim $\alpha(\omega)$ to be a non-theorem. In the opposite case, i.e. if there are at least M formulas among $\alpha_1(\omega), \ldots, \alpha_N(\omega)$ enabling to prove $\alpha(\omega)$ and no of the formulas $\alpha_1(\omega), \ldots$ $\ldots, \alpha_N(\omega)$ enables to disprove $\alpha(\omega)$, we proclaim $\alpha(\omega)$ to be a theorem. The parameter space X, clearly, has the form

$$X = \{\langle M, N\rangle : N = 1, 2, \ldots, M = 0, 1, 2, \ldots, N\}\,.$$

Theorem 1. Let $\langle M_0, N_0\rangle \in X$, let $\langle \Omega, \mathscr{S}, P\rangle$ be a probability space on which random variables $\alpha, \alpha_1, \alpha_2, \ldots, \alpha_N$ are defined, mutually independent and, for $\alpha_1, \alpha_2, \ldots, \alpha_N$, equally distributed, such that

$$(2) \qquad 0 < P(\{\omega : \omega \in \Omega,\ \alpha(\omega) \in \mathscr{T}\}) < 1\,,$$

$$(3) \qquad 0 < P(\{\omega : \omega \in \Omega,\ \alpha_1(\omega) = a'\})$$

for any $a' = V(a, i, s) \in \mathscr{L}$. Let $T(\alpha, \langle M_0, N_0 \rangle, \cdot)$ be the statistical deducibility testing procedure defined, using X, $\alpha, \alpha_1, \alpha_2, \ldots, \alpha_N$ as above. Then there exists, for any $\varepsilon > 0$, a parameter value $\langle M_\varepsilon, N_\varepsilon \rangle$ such that

$$PE_1(\langle M_\varepsilon, N_\varepsilon \rangle) = PE_1(T(\alpha, \langle M_\varepsilon, N_\varepsilon \rangle, \cdot)) < \varepsilon .$$

Remark. Similar theorems are proved in [10], [5], [7] in the general case of statistical deducibility testing procedures, however, under a stronger condition, namely, $P(\{\omega : \alpha_1(\omega) = a'\}) > 0$ for any $a' \in \mathscr{L}$. Theorem 1 shows that in the more specified conditions investigated in this condition can be weakened. In other words, Theorem 1 shows the set of all formulas of the form $V(a, i, s)$ (this set will be denoted by $\mathscr{V} \subset \mathscr{L}$) to be "rich enough" to serve as a basis of a statistical deducibility testing procedure based on at random sampled extensions.

Proof. (2) gives that $\mathscr{T} \neq \mathscr{L}$, hence, the formalized theory $\langle \mathscr{L}, \mathscr{T} \rangle$ is consistent. Let $\alpha(\omega) \in \mathscr{L} - \mathscr{T}$, then a necessary (but not sufficient) condition for $\alpha_i(\omega) \to \alpha(\omega)$ to be a theorem is that $\alpha_i(\omega) \in \mathscr{L} - \mathscr{T}$. Taking into consideration the supposed statistical independence of the random variables $\alpha_1, \alpha_2, \ldots, \alpha_{N_0}$ and setting $M_0 = N_0$ we have

$$PE_1(T(\alpha, \langle N_0, N_0 \rangle, \cdot)) \leqq P(\{\omega : \omega \in \Omega, \alpha_1(\omega) \in \mathscr{L} - \mathscr{T}, \ldots, \alpha_{N_0}(\omega) \in \mathscr{L} - \mathscr{T}\}) =$$
$$= (P(\{\omega : \alpha_1(\omega) \in \mathscr{L} - \mathscr{T}\}))^{N_0} \to 0$$

for $N_0 \to \infty$, as $P(\{\omega : \alpha_1(\omega) \in \mathscr{L} - \mathscr{T}\}) < 1$ (e.g., if j_0 is a tautological property, then $V(a, j_0, s) \in \mathscr{T}$ for any $a \in I^2$, $s \in S$ and $P(\{\omega : \alpha_1(\omega) = V(a, j_0, s)\}) > 0$ according to (3).

So, if $M_\varepsilon = N_\varepsilon > (\log \varepsilon)\,(\log\,(P(\{\omega : \omega \in \Omega, \alpha_1(\omega) \in \mathscr{L} - \mathscr{T}\})))^{-1}$, then $PE_1(\langle M_\varepsilon, N_\varepsilon \rangle) < \varepsilon$. Q.E.D.

In all the rest of this paper we suppose to have chosen and fixed a statistical deducibility testing procedure $T(\alpha, \langle M, N \rangle, \cdot)$ such that $PE_1(\langle M, N \rangle) < \varepsilon$ and $PE_2(\langle M, N \rangle)$ is minimal or acceptable from some point of view. Abbreviately, this $T(\alpha, \langle M, N \rangle, \cdot)$ will be denoted by $T_\varepsilon(\alpha)$.

Let us return in our considerations to the moment when we transformed a question how to solve the problem $\langle X, Y \rangle$ into another question – to find a resolution based proof of the formula $X(s_0) \to (\exists s')\,(Y(s'))$. Suppose our attempt to prove this formula to fail and apply the statistical deducibility testing procedure $T_\varepsilon(\alpha(X(s_0) \to (\exists s')\,(Y(s'))))$, where $\alpha(a)$, $a \in \mathscr{L}$, is the random variable taking only one value from $\mathscr{L}$, namely a. Consider the case when $T_\varepsilon(\alpha(X(s_0) \to (\exists s')\,(Y(s')), \omega) = \mathrm{t}$, i.e. the tested formula is proclaimed to be a theorem, then we have m formulas, $M_\varepsilon \leqq m \leqq N_\varepsilon$, say $\alpha_{i_1}(\omega), \alpha_{i_2}(\omega), \ldots, \alpha_{i_m}(\omega)$, $i_m \leqq N_\varepsilon$, sampled at random and such that we are able to prove $\alpha_{i_j}(\omega) \to [X(s_0) \to (\exists s')\,(Y(s'))]$ for any $j \leqq m$. Hence, we are also able to prove

$$[\alpha_{i_1}(\omega) \vee \alpha_{i_2}(\omega) \vee \ldots \vee \alpha_{i_m}(\omega)] \to [X(s_0) \to (\exists s')\,(Y(s'))] .$$

It follows immediately, that if we were able to verify at least one among the formulas $\alpha_{i_j}(\omega)$, $j \leqq m$, as being valid in the environment, we can enrich our formalized theory by the new axiom $\bigvee_{j=1}^{m} a_{i_j}(\omega)$ and then we shall be able to prove $X(s_0) \rightarrow (\exists s')(Y(s'))$ in the enriched theory using the resolution-based theorem-prover supposed to be at our disposal. Clearly, this proof can be then used in order to derive a plan solving the problem $\langle X, Y\rangle$. In other words, the statistical deducibility testing procedure serves as a random generator of auxiliary hypotheses; the verification at least one of them is a sufficient condition for obtaining a plan which solves $\langle X, Y\rangle$.

Using Theorem 1 we may suppose and shall suppose in all the rest of this paper, that the random variables $\alpha_1, \alpha_2, \ldots, \alpha_N$ take their values in the set $\mathscr{V} \subset \mathscr{L}$, $\mathscr{V} = \{V(a, i, s) : a \in I^2,\ i \in \mathscr{N}, s \in S\}$ under the condition that (3) holds. It means, that every $\alpha_i(\omega)$ is a formula $V(a, i, s)$, $a \in I^2$, $i \in \mathscr{N}$, $s \in S$, and we may associate with any statistical deducibility testing procedure $T(\alpha, \langle M, N\rangle, \cdot)$ and any $a \in \mathscr{L}$ a set-valued random variable $T^*(a)$, defined on $\langle \Omega, \mathscr{S}, P\rangle$ and taking its values in the set of all the most N-elemented subsets of $\mathscr{V}$, namely:

$$T^*(a)(\omega) = \{V(a_1, j_1, s_1, \omega),\ V(a_2, j_2, s_2, \omega), \ldots, V(a_m, j_m, s_m, \omega)\} \subset \mathscr{V},$$

for any $k \leqq m$ there is $l_k \leqq N$ such that $\alpha_{l_k}(\omega) = V(a_k, j_k, s_k, \omega)$ and $\alpha_{l_k}(\omega) \rightarrow a$ has been proved to be a theorem and for no $\alpha_{l'}(\omega)$ not belonging to $T^*(a)(\omega)$ the implication $\alpha_{l'}(\omega) \rightarrow a$ has been proved to be a theorem. This formal definition of $T^*(\alpha)(\omega)$ seems to be very complicated, however, the intuitive sense is quite simple: $T^*(a)(\omega)$ is the set of all formulas sampled at random during the process of statistical deducibility testing of a which "helped" us to prove a. From this also the following assertion immediately follows:

Theorem 2. For any $a \in \mathscr{L}$, if $T(\alpha(a), \langle M, N\rangle, \omega) = \mathfrak{t}$, then $T^*(a)(\omega)$ contains at least M elements (supposing that possible multiple samples of the same formula are counted separately).

3. AUTOMATON EXPERIMENTING POSSIBILITIES AND THEIR CLASSIFICATION

We have to consider, now, our abilities (or automaton abilities) when a formula $V(a, i, s) \in \mathscr{V}$ is to be verified. These abilities are determined by the technical and program equipment of the automaton and by the real physical state of the environment and they cannot be, in such a case, an object of a direct mathematical investigation. The situation is similar to that of automaton operators and conditions associated with them; such model is investigated in [11] and other papers dealing with robotics and artificial intelligence. We have accepted this point of view also in [7],

[8] and among the four types of automaton actions (moves, observations, deductions and operations) operations were the only connected with some conditions of applicability the other actions being supposed to be applicable in any case. In order to develop further our model of automated experiment planning we shall suppose, in what follows, that also observations are connected with some conditions and validity at least one of these conditions is necessary and sufficient for the observation to be successful (i.e. giving a definite answer).

The observational conditions are described, formally, by a subset $EXP \subset I^2 \times \times \mathcal{N} \times I^2 \times \mathcal{N} \times S$ it means, EXP contains quintuples of the form $\langle b, j, a, i, s\rangle$, $a, b \in I^2$, $i, j \in \mathcal{N}$, $s \in S$. The intuitive sense is as follows: let $\langle b, j, a, i, s\rangle \in EXP$, then, in order to be able to decide, by an observation, whether $V(a, i, s)$ holds or does not hold we must change the situation of the environment into such a new situation s' that $V(b, j, s')$ holds. In other words, to be able to observe $V(a, i, s)$ or its negation the point $b \in I^2$ must possess the j-th property. Of course, the possibility that there are different b's and j's such that $\langle b, j, a, i, s\rangle \in EXP$ is not excluded. For a fixed $V(a, i, s)$ we denote

$$EXP_V(a, i, s) = \{\langle b, j\rangle : \langle b, j, a, i, s\rangle \in EXP\}$$

and any $\langle b, j\rangle \in EXP_V(a, i, s)$ can be called *an experiment for deciding* $V(a, i, s)$.

In case the set $EXP_V(a, i, s)$ contains more than one pair we have a possibility to choose among them when desiring to verify $V(a, i, s)$. As the most natural criterion according to which the pairs from $EXP_V(a, i, s)$ can be classified and ordered we take the minimal expenses connected with the realization of a situation enabling to verify $V(a, i, s)$. The notion of expenses is supposed to be general enough to cover not only the demands of financial kind but also those of processing time, computer storage space etc.

Let us suppose that to any operation φ which the automaton has, in principle, at its disposal a cost $c(\varphi)$ of the application of this operation is ascribed; for the sake of simplicity we suppose this cost to be the same in all the cases when φ applied. To any finite sequence $\langle \varphi_1, \varphi_2, \ldots, \varphi_n\rangle$ of operations we may ascribe the cost $\sum_{i=1}^{n} c(\varphi_i)$. Denote, now, for any $\langle b, j, a, i, s\rangle \in EXP$:

$$g(b, j, a, i, s, s_0) = \{s' \in S : s' = s_0 \varphi_1 \varphi_2 \ldots \varphi_n \text{ for some } n \text{ and some } \varphi_1, \varphi_2, \ldots, \varphi_n,\ V(b, j, s') \text{ holds}\},$$

i.e. $g(b, j, a, i, s, s_0)$ is the set of all situations accessible from the present (or initial) situation s_0 by a finite operator sequence and enabling to verify or disprove $V(a, i, s)$. The case $g(b, j, a, i, s, s_0) = \emptyset$ can be interpreted in such a way that, in the actual situation s_0, no experiment exists which would give the possibility to verify $V(a, i, s)$.

The cost of transforming the initial situation s_0 into an $s' \in g(b, j, a, i, s, s_0)$ can be simply defined as $\sum_{i=1}^{n} c(\varphi_i)$, where $s' = s_0 \varphi_1 \varphi_2 \dots \varphi_n$ and denoted by $c(s'/s_0)$, define $c(s'/s_0)$ as ∞ for any $s' \in S - g(b, j, a, i, s, s_0)$. For any $\langle b, j, a, i, s\rangle \in EXP$ we set

$$C(b, j, a, i, s) = \min \{c(s'/s_0) : s' \in g(b, j, a, i, s, s_0)\} ,$$
$$C(b, j, a, i, s) = \infty , \quad \text{if} \quad g(b, j, a, i, s, s_0) = \emptyset .$$

The value $C(b, j, a, i, s)$ has a simple interpretation: it is the minimal cost necessary to verify or disprove $V(a, i, s)$ by realizing $V(b, j, s')$ for an appropriate s'. Setting

$$C(EXP_V(a, i, s)) = \min \{C(b, j, a, i, s) : \langle b, j\rangle \in EXP_V(a, i, s)\} ,$$
$$C(EXP_V(a, i, s)) = \infty , \quad \text{if} \quad EXP_V(a, i, s) = \emptyset ,$$

we can easily see that the value $C(EXP_V(a, i, s))$ expresses the minimal cost of an experiment enabling to verify or disprove $V(a, i, s)$. Hence, we may immediately define

Definition 1. Let $T(\alpha, \langle M, N\rangle, \cdot)$ be a statistical deducibility testing procedure such that $PE_1(\langle M, N\rangle) < \varepsilon, \varepsilon > 0$ given a priori, let $Y \in \mathscr{L}$ be a formula. Define

$$C(Y, \varepsilon, \omega) = \min \{C(EXP_V(a, i, s)(\omega)) : V(a, i, s, \omega) \in T^*(Y)(\omega)\} ,$$
$$\text{if} \quad T^*(Y)(\omega) \neq \emptyset ,$$
$$C(Y, \varepsilon, \omega) = \infty , \quad \text{if} \quad T^*(Y)(\omega) = \emptyset .$$

The value $C(Y, \varepsilon, \omega)$ will be called the *cost of verifying or disproving of Y on the ground of an experiment sampled by the statistical deducibility testing procedure* $T(\alpha(Y), \langle M, N\rangle, \cdot)$.

This definition closes the part of this paper devoted to developing a model of automated experiment planning based on an appropriate statistical deducibility testing procedure and it is why we feel the need to add two remarks concerning some open points in our notions introduced above.

First, we have to mention the problem of reversibility of experiments. Among the demands which any experiment is usually supposed to satisfy is that of the *possibility to return back* to the situation before starting the experiment. This demand may involve many questions and problems going up to a rather high philosophical level, but such considerations would be outside the subject of this paper. In our model this demand is not satisfiable because of our identifying "situations" with the sequences of actions having been applied since the initial situation till the present one, hence, the situation *after* the experiment can never be the same as *before* the experiment. To be able to cope with the demand of reversibility within our model we would need an equivalence relation on the set S of situations defining which two situations are considered to be "the same", however, to formalize it in an appropriate way is far from being so simple.

Second, considering the value $C(Y, \varepsilon, \omega)$ as the cost of verifying or disproving of Y on the ground of $T(\alpha(Y), \langle M, N\rangle, \cdot)$ we must admit that it covers only partially the actual situation. The procedure T offers a set of hypotheses $T^*(Y)(\omega)$ the verification any of them being, at once, also the verification of Y and to this set another set, namely

$$(4) \qquad \bigcup \{EXP_V(a, i, s)(\omega) : V(a, i, s, \omega) \in T^*(Y)(\omega)\}$$

of adequate experiments is ascribed. The positive answer to any of these experiments verifies also Y and the cost of the cheapest among them is just $C(Y, \varepsilon, \omega)$. This value does not express, however, the expenses connected with finding this cheapest experiment among all the others contained in (4). The close and interesting connections between the properties of the statistical deducibility testing procedure $T(\alpha, \langle M, N\rangle, \cdot)$ and those of $C(Y, \varepsilon, \omega)$ justifies, nevertheless, the introduced definition of $C(Y, \varepsilon, \omega)$.

The value $C(Y, \varepsilon, \omega)$ describes the expenses necessary in order to verify or disprove at least one among the formulas in $T^*(Y)(\omega)$. However, these two possible answers are of different value for the further use, namely, verifying at least one formula from $T^*(Y)(\omega)$ to be valid is sufficient for proclaiming Y to be valid as well without any doubts. On the other hand, even in case all the formulas from $T^*(Y)(\omega)$ are disproved on the ground of some experiments we cannot be sure whether Y holds or does not hold. It is why we define the notion of *verifying conditions* $EXP^+ \subset EXP$ as the set of all quintuples $\langle b, j, a, i, s\rangle \in EXP$ such that the formula $V(a, i, s)$ is verified supposing $V(b, j, s')$ to be valid in a situation s'. Using EXP^+ instead of EXP we define $EXP_V^+(a, i, s)$ (experiment for verifying $V(a, i, s)$), $g^+(b, j, a, i, s, s_0)$, $C^+(b, j, a, i, s)$, $C^+(EXP_V^+(a, i, s))$ and, finally, $C^+(Y, \varepsilon, \omega)$.

Theorem 3. If Y is a non-valid formula, then $C^+(Y, \varepsilon, \omega) = \infty$ for any $\varepsilon > 0$, $\omega \in \Omega$.

Proof. If Y is not valid, then $Y \in \mathscr{L} - \mathscr{T}$, hence, if $V(a, i, s) \rightarrow Y$ is provable, then $V(a, i, s) \in \mathscr{L} - \mathscr{T}$, so $T^*(Y) \subset \mathscr{L} - \mathscr{T}$. However, experiments are supposed to give correct answers, so no experiment exists verifying a non-valid formula. This gives, that $V(a, i, s) \in \mathscr{L} - \mathscr{T}$ implies $EXP_V^+(a, i, s) = \emptyset$, $C^+(EXP_V^+(a, i, s)) = \infty$ and, finally, $C^+(Y, \varepsilon, \omega) = \infty$. Q.E.D.

Set, for any $Y \in \mathscr{L}$,

$$C_0(Y) = \min \{C(EXP_V(a, i, s)) : V(a, i, s) \vdash_{ef} Y\},$$
$$C_0^+(Y) = \min \{C^+(EXP_V^+(a, i, s)) : V(a, i, s) \vdash_{ef} Y\},$$

where the symbol $\vdash_{ef}$ means that the deduction in question is provable by the means being at the automaton disposal.

Theorem 4. If Y is a valid formula such that $C_0^+(Y) < \infty$, then, for $\varepsilon \to 0$,

$$P(\{\omega : C^+(Y, \varepsilon, \omega) = C_0^+(Y)\}) \nearrow 1 .$$

Proof. Y is valid and $C_0^+(Y) < \infty$, hence, there is at least one formula, say $V(a_0, i_0, s_0)$, such that $V(a_0, i_0, s_0) \in \mathcal{T}$, $V(a_0, i_0, s_0) \vdash_{ef} Y$ and

$$C^+(EXP_V^+(a_0, i_0, s_0)) = C_0^+(Y) .$$

The statistical deducibility testing procedure $T(\alpha(Y), \langle M, N\rangle, \cdot)$ is completely defined by the pair $\langle M, N\rangle$, hence, for a fixed N there is just $N + 1$ possible procedures with at most $N + 1$ different values of $PE_1(\langle M, N\rangle)$. To minimize $PE_1(\langle M, N\rangle)$ under the minimal among these values requests, hence, to enlarge N, so $\varepsilon \to 0$ implies $N \to \infty$. However,

$$P(\{\omega\colon \alpha_1(\omega) = V(a_0, i_0, s_0)\}) > 0 ,$$

so, according to the supposed statistical independence and equal distribution of α_i we have

$$P(\{\omega : V(a_0, i_0, s_0) \in \{\alpha_1(\omega), \alpha_2(\omega), \ldots, \alpha_N(\omega)\}\}) =$$
$$= 1 - (1 - P(\{\omega : \alpha_1(\omega) = V(a_0, i_0, s_0)\}))^N$$

and this expression tends to 1 if N increases. Moreover, $V(a_0, i_0, s_0) \in \{\alpha_1(\omega), \ldots$ $\ldots, \alpha_N(\omega)\}$ implies $V(a_0, i_0, s_0) \in T^*(Y)$ as $V(a_0, i_0, s_0) \vdash_{ef} Y$, and this implies, again, that $C^+(Y, \varepsilon, \omega) = C_0^+(Y)$. Hence, if $\varepsilon \to 0$, then

$$P(\{\omega\colon C^+(Y, \varepsilon, \omega) = C_0^+(Y)\}) \to 1 ,$$

Q.E.D.

Theorem 5. Let α be a random variable defined on the probability space $\langle \Omega, \mathcal{S}, P\rangle$ and taking its values in the set $\mathcal{L}$ of formulas, let $K > 0$ be a given real number. Then, for $\varepsilon \to 0$,

$$E(\min\{K, C(\alpha(\cdot), \varepsilon, \cdot)\}/\{\omega\colon T(\alpha(\omega), \langle M, N\rangle, \omega) = \mathrm{t}\}) \to$$
$$\to E(\min\{K, C_0^+(\alpha(\cdot))\}/\{\omega\colon \alpha(\omega) \in \mathcal{T}\})$$

(here $E(\cdot/A)$ denotes the conditional expected value of the random variable in question conditioned by the random event A).

Proof. Because of the fact that $0 \leqq \min\{K, C(\alpha(\omega), \varepsilon, \omega)\} \leqq K$ for any $\omega \in \Omega$ we have

$$E(\min\{K, C(\alpha(\cdot), \varepsilon, \cdot)\}/\{\omega\colon T(\alpha(\omega), \langle M, N\rangle, \omega) = \mathrm{t}\}) =$$
$$= \int(\min\{K, C(\alpha(\cdot). \varepsilon, \cdot)\})\, dP(\cdot/\{\omega\colon T(\alpha(\omega), \langle M, N\rangle, \omega) = \mathrm{t}\}) =$$

$$= \left[\int (\min \{K, C(\alpha(\cdot), \varepsilon, \cdot)\}) \, dP(\cdot / \{\omega: T(\alpha(\omega), \langle M, N \rangle, \omega) = t\} \cap \{\omega: \alpha(\omega) \in \mathscr{T}\})\right] .$$
$$. \, P(\{\omega: \alpha(\omega) \in \mathscr{T}\} / \{\omega: T(\alpha(\omega), \langle M, N \rangle, \omega) = t\}) +$$
$$+ \left[\int (\min \{K, C(\alpha(\cdot), \varepsilon, \cdot)\}) \, dP(\cdot / \{\omega: T(\alpha(\omega), \langle M, N \rangle, \omega) = t\} \cap$$
$$\cap \{\omega: \alpha(\omega) \in \mathscr{L} - \mathscr{T}\})\right] P(\{\omega: \alpha(\omega) \in \mathscr{L} - \mathscr{T}\} / \{\omega: T(\alpha(\omega), \langle M, N \rangle, \omega) = t\}) .$$

According to the definition of the first type probability of error and according to the assumption that we take into consideration only such statistical deducibility testing procedures for which this probability is majorized by ε, we obtain (using Bayes rule)

$$P(\{\omega: \alpha(\omega) \in \mathscr{L} - \mathscr{T}\} / \{\omega: T(\alpha(\omega), \langle M, N \rangle, \omega) = t\}) =$$
$$= PE_1(\langle M, N \rangle) \, P(\{\omega: \alpha(\omega) \in \mathscr{L} - \mathscr{T}\}) \, (P(\{\omega: T(\alpha(\omega), \langle M, N \rangle, \omega) = t)^{-1} \leqq \varepsilon c_1 ,$$

where

$$c_1 = \frac{P(\{\omega: \alpha(\omega) \in \mathscr{L} - \mathscr{T}\})}{P(\{\omega: T(\alpha(\omega), \langle M, N \rangle, \omega) = t\})}$$

is a constant value. This result and (5) give

(6)
$$\left[\int (\min \{K, C(\alpha(\cdot), \varepsilon, \cdot)\}) \, dP(\cdot / \{\omega: T(\alpha(\omega), \langle M, N \rangle, \omega) = t\} \cap \{\omega: \alpha(\omega) \in \mathscr{T}\})\right] .$$
$$. \, P(\{\omega: \alpha(\omega) \in \mathscr{T}\} / \{\omega: T(\alpha(\omega), \langle M, N \rangle, \omega) = t\}) \leqq$$
$$\leqq E(\min \{K, C(\alpha(\cdot), \varepsilon, \cdot)\}) \leqq$$
$$\leqq \left[\int (\min \{K, C(\alpha(\cdot), \varepsilon, \cdot)\}) \, dP(\cdot / \{\omega: T(\alpha(\omega), \langle M, N \rangle, \omega) = t\} \cap \{\omega: \alpha(\omega) \in \mathscr{T}\})\right] .$$
$$. \, P(\{\omega: \alpha(\omega) \in \mathscr{T}\} / \{\omega: T(\alpha(\omega), \langle M, N \rangle, \omega) = t\}) + K\varepsilon c_1 .$$

Moreover,

(7)
$$\left[\int (\min \{K, C(\alpha(\cdot), \varepsilon, \cdot)\}) \, dP(\cdot / \{\omega: T(\alpha(\omega), \langle M, N \rangle, \omega) = t\} \cap \{\omega: \alpha(\omega) \in \mathscr{T}\})\right] .$$
$$. \, P(\{\omega : \alpha(\omega) \in \mathscr{T}\} / \{\omega : T(\alpha(\omega), \langle M, N \rangle, \omega) = t\}) \geqq$$
$$\geqq \left[\int (\min \{K, C(\alpha(\cdot), \varepsilon, \cdot)\}) \, dP(\cdot / \{\omega: \alpha(\omega) \in \mathscr{T}\})\right] (1 - \varepsilon c_1) -$$

$$- [P(\{\omega: T(\alpha(\omega), \langle M, N\rangle, \omega) = \mathfrak{f}\} \cap \{\omega: \alpha(\omega) \in \mathscr{T}\})] K(1 - \varepsilon c_1) =$$
$$= \left[\int (\min \{K, C(\alpha(\omega), \varepsilon, \cdot)\}) \, \mathrm{d}P(\cdot / \{\omega: \alpha(\omega) \in \mathscr{T}\})\right] (1 - \varepsilon c_1) -$$
$$- PE_1(\langle M, N\rangle) P(\{\omega: \alpha(\omega) \in \mathscr{T}\}) K(1 - \varepsilon c_1) .$$

If $\alpha(\omega) \in \mathscr{T}$, then $C(\alpha(\omega), \varepsilon, \omega) = C^+(\alpha(\omega), \varepsilon, \omega)$, hence (7) equals to

$$\left[\int (\min \{K, C^+(\alpha(\cdot), \varepsilon, \cdot)\}) \, \mathrm{d}P(\cdot / \{\omega: \alpha(\omega) \in \mathscr{T}\})\right] (1 - \varepsilon c_1) -$$
$$- PE_1(\langle M, N\rangle) P(\{\omega: \alpha(\omega) \in \mathscr{T}\}) K(1 - \varepsilon c_1) \geqq$$
$$\geqq \left[\int (\min \{K, C^+(\alpha(\cdot), \varepsilon, \cdot)\}) \, \mathrm{d}P(\cdot / \{\omega: \alpha(\omega) \in \mathscr{T}\})\right] (1 - \varepsilon c_1) -$$
$$- K \, P(\{\omega: C^+(\alpha(\omega), \varepsilon, \omega) \neq C_0^+(\alpha(\omega))\}) - K\varepsilon(1 - \varepsilon c_1) P(\{\omega: \alpha(\omega) \in \mathscr{T}\}) \geqq$$
$$\geqq E((\min \{K, C_0^+(\alpha(\cdot))\}) / \{\omega : \alpha(\omega) \in \mathscr{T}\}) - K \, P(\{\omega : C^+(\alpha(\omega), \varepsilon, \omega) \neq C_0^+(\alpha(\omega))\}) -$$
$$- K \, \varepsilon(1 - \varepsilon c_1) P(\{\omega: \alpha(\omega) \in \mathscr{T}\}) .$$

Theorem 4 gives that, for $\varepsilon \to 0$,

$$P(\{\omega: C^+(\alpha(\omega), \varepsilon, \omega) \neq C_0^+(\alpha(\omega))\}) \to 1 ,$$

so it follows immediately, that the right side of the last inequality in (8) tends to

$$E((\min \{K, C_0^+(\alpha(\cdot))\}) / \{\omega: \alpha(\omega) \in \mathscr{T}\}) ,$$

if $\varepsilon \to 0$. Combining this result with (6) we obtain that, for $\varepsilon \to 0$,

$$E(\min \{K, C(\alpha(\cdot), \varepsilon, \cdot)\} / \{\omega: T(\alpha(\omega), \langle M, N\rangle, \omega) = \mathfrak{t}\}) \to$$
$$\to E(\min \{K, C_0^+(\alpha(\cdot))\} / \{\omega: \alpha(\omega) \in \mathscr{T}\}) ,$$

Q.E.D.

In spite of the fact that from the formal point of view the foregoing theorem seems to be rather complicated, its intuitive sense is simple. In a practical case of automated experiment planning and executing the automaton is not allowed to continue this activity without any a priori given bound concerning the maximal admissible cost of such an experiment. Our investigating the random variable $\min \{K, C(\alpha(\cdot), \varepsilon, \cdot)\}$ instead of C itself reflects this condition and from the purely mathematical point of view this boundedness enables to prove the assertion of Theorem 5. Its sense is as follows: the expected value of expenses connected with verifying or disproving these formulas, which have been proclaimed to be theorems by the used statistical deducibility testing procedure tends to the expected value of the minimal expenses connected with verifying theorems, when the first type probability of error connected with the used testing procedure tends to 0. Hence, the smaller this probability of error is the better to orient oneself to verifying (and to neglect disproving) when the auxiliary hypotheses from $T^*(\alpha(\omega))$ are investigated in more details.

In order to be able to prove some more assertions concerning the expenses connected with automatically proposed experiments let us introduce the notions if *support* and *effective support* of a formula. If Y is a formula from $\mathscr{L}$, then its support is denoted by $Sup(Y)$ and defined

$$Sup(Y) = \{V(a, i, s) : V(a, i, s) \vdash Y\} \subset \mathscr{V}\ .$$

The effective support $Sup_0(Y)$ is defined

$$Sup_0(Y) = \{V(a, i, s) : V(a, i, s) \vdash_{ef} Y\} \subset \mathscr{V}\ ,$$

where $V(a, i, s) \vdash_{ef} Y$ means that $V(a, i, s) \vdash Y$ and, at the same time, the automaton is able to prove Y from $V(a, i, s)$ using the theorem-prover being at its disposal.

Lemma 1. Let $Y_1, Y_2, \ldots, Y_n$ be formulas from $\mathscr{L}$. Then

a) $Sup(\bigvee_{j=1}^{n} Y_j) \supset \bigcup_{j=1}^{n} Sup(Y_j)$.

b) if $\vdash_{ef}$ is closed with respect to propositional calculus tautologies (i.e. for any such tautology Y $Sup_0(Y) = \mathscr{V} = \{V(a, i, s) : a \in I^2, i \in \mathscr{N}, s \in S\}$), then

$$Sup_0(\bigvee_{j=1}^{n} Y_j) \supset \bigcup_{j=1}^{n} Sup_0(Y_j).$$

Proof. (a) follows immediately from the definition of Sup.

(b) The formula

$$(V(a, i, s) \to Y_1) \to (V(a, i, s) \to \bigvee_{j=1}^{n} Y_j)$$

is a propositional calculus tautology, hence, it is deducible in the sense of $\vdash_{ef}$, which gives also

$$V(a, i, s) \to Y_1 \vdash_{ef} V(a, i, s) \to \bigvee_{j=1}^{n} Y_j$$

and

$$V(a, i, s) \vdash_{ef} \bigvee_{j=1}^{n} Y_j\ ,$$

hence, $Sup_0(\bigvee_{j=1}^{n} Y_j) \supset \bigcup_{j=1}^{n} Sup_0(Y_j)$. Q.E.D.

Clearly, for any Y, $Sup_0(Y) \subset Sup(Y)$. If Y is a theorem, then $Sup(Y) = \mathscr{V}$ (but $Sup_0(Y) \neq \mathscr{V}$, in general), if Y is a non-theorem, then $Sup(Y) \subset \mathscr{V} - \mathscr{T} = \mathscr{V} \cap \cap (\mathscr{L} - \mathscr{T})$.

We have defined the cost of an experiment $EXP_V(a, i, s)$ only for elementary formulas $V(a, i, s)$. Let us extend this definition to the case of an alternative of such elementary formulas setting

$$C(EXP_V(\bigvee_{j=1}^{n} V(a_j, i_j, s_j))) = \min \{C(EXP_V(a_j, i_j, s_j)) : j \leqq n\} .$$

In order to abbreviate our further reasonings let us introduce the following notion.

Definition 2. The cost C of experiments is called *continuous with respect to* $\vdash_{ef}$, if for any random variable α defined on the probability space $\langle \Omega, \mathcal{S}, P \rangle$, taking its values in $\mathcal{V}$ and such that $P(\{\omega : \alpha(\omega) = V(a, i, s)\}) > 0$ for any $V(a, i, s) \in \mathcal{V}$ and for any sequence $\{V(a_j, i_j, s_j)\}_{j=1}^{\infty}$,

$$\lim_{n \to \infty} P(\{\omega : \alpha(\omega) \in Sup_0\ V(\bigvee_{j=1}^{n} (a_j, i_j, s_j))\}) = 1$$

implies

$$\lim_{n \to \infty} C(EXP_V(\bigvee_{j=1}^{n} V(a_j, i_j, s_j))) = 0 .$$

Theorem 6. Let the cost C of experiments be continuous with respect to $\vdash_{ef}$. Suppose that there is, for any theorem Y and any $V(a, i, s) \in \mathcal{V}$, such a $V(a', i', s') \in \mathcal{V}$ that $V(a', i', s') \vdash_{ef} Y$ and $V(a, i, s) \in Sup_0(V(a', i', s'))$. Then there is a real continuous function $f : \langle 0, 1 \rangle \to \langle 0, 1 \rangle$ such that $\lim_{x \to 0} f(x) = 0$ and, moreover, for any theorem Y and any $\omega \in \Omega$, $T(\alpha(Y), \varepsilon, \omega) = \mathrm{t}$ implies $C(Y, \varepsilon, \omega) < f(\varepsilon)$.

Proof. Let Y be a theorem, let

$$T(\alpha(Y), \varepsilon, \omega) = T(\alpha(Y), \langle M(\varepsilon), N(\varepsilon) \rangle, \omega) = \mathrm{t} .$$

This means that $N(\varepsilon)$ formulas $V(a_1, i_1, s_1, \omega)$, $V(a_2, i_2, s_2, \omega)$, $\ldots$, $V(a_N, i_N, s_N, \omega)$ have been sampled, the relations $V(a_j, i_j, s_j) \vdash_{ef} Y$, $j \leqq N(\varepsilon)$, have been tested and at least $M(\varepsilon)$ among them have been proved to hold. Take any $V(a, i, s) \in \mathcal{V}$, denote by $V(a', i', s')$ one among the formulas satisfying $V(a, i, s) \in Sup_0(V(a', i', s'))$ and $V(a', i', s') \vdash_{ef} Y$, i.e. $V(a', i', s') \in Sup_0(Y)$. Denoting by α_0 the random variable sampling $V(a_j, i_j, s_j, \omega)$, $j \leqq N(\varepsilon)$, we have that

$$P(\{\omega : V(a, i, s) \in Sup_0(\bigvee_{k=1}^{m} V(a_{j_k}, i_{j_k}, s_{j_k}, \omega))\}) \geqq$$

$$\geqq P(\{\omega : V(a', i', s') \in \bigcup_{j=1}^{N(\varepsilon)} \{V(a_j, i_j, s_j, \omega)\}\}) =$$

$$= 1 - (1 - P(\{\omega : \alpha_0(\omega) = V(a', i', s')\}))^{N(\varepsilon)} ,$$

where $V(a_{j_k}, i_{j_k}, s_{j_k}, \omega)$, $k = 1, 2, \ldots, m$, $M(\varepsilon) \leqq m \leqq N(\varepsilon)$, denotes those formulas among $V(a_1, i_1, s_1, \omega), \ldots, V(a_N, i_N, s_N, \omega)$, for which $V(a_{j_k}, i_{j_k}, s_{j_k}, \omega) \vdash_{ef} Y$ holds. Hence,

$$P(\{\omega : \alpha(\omega) \in Sup_0(\bigvee_{k=1}^{m} V(a_{j_k}, i_{j_k}, s_{j_k}, \omega))\}) \geqq$$

$$\geqq \sum[(1 - (1 - P(\{\omega : \alpha_0(\omega) = V(a', i', s')\}))^{N(\varepsilon)}) P(\{\omega : \alpha_0(\omega) = V(a, i, s)\}) =$$

$$= E(1 - (1 - P(\{\omega : \alpha_0(\omega) = V(a', i', s')\}))^{N(\varepsilon)}),$$

the sum is taken over all $V(a, i, s) \in \mathscr{V}$. Now, if $\varepsilon \to 0$, then $N(\varepsilon) \to \infty$ (see the proof of Theorem 4) and this implies that also $M(\varepsilon) \to \infty$ and $m \to \infty$ (if $N \to \infty$ and M is bounded, then $PE_1(\langle M, N \rangle) \to 1$ as can be easily seen and this contradicts the condition $PE_1(\langle M, N \rangle) < \varepsilon$). This gives that

$$P(\{\omega : \alpha(\omega) \in Sup_0(\bigvee_{k=1}^{m} V(a_{j_k}, i_{j_k}, s_{j_k}, \omega))\}) \to 1$$

and, applying the continuity condition we have that

$$\lim_{m \to \infty} C(EXP_V(\bigvee_{j=1}^{m} V(a_{j_k}, i_{j_k}, s_{j_k}, \omega))) = 0\,.$$

However, $C(EXP_V(\bigvee_{j=1}^{m} V(a_{j_k}, i_{j_k}, s_{j_k}, \omega))) = \min \{C(EXP_V(a_{j_k}, i_{j_k}, s_{j_k}, \omega)) : k \leqq m\} =$ $= \min (C(EXP_V(a, i, s))(\omega) : V(a, i, s, \omega) \in T^*(Y)(\omega)\} = C(Y, \varepsilon, \omega)$ according to Definition 1. Hence, for any $\varepsilon > 0$ there is $f(\varepsilon)$ such that f is a continuous function, $\lim_{x \to 0} f(x) = 0$ and $T(\alpha(Y), \varepsilon, \omega) = \mathrm{t}$ implies $C(Y, \varepsilon, \omega) < f(x)$. Q.E.D.

Remark. The condition that there is, for any $Y \in \mathscr{T}$ and $V(a, i, s) \in \mathscr{V}$ such a $V(a', i', s')$ that $V(a, i, s) \in Sup_0(V(a', i', s'))$ and $V(a', i', s') \in Sup_0(Y)$ can be called an *intermediate deducibility condition.* It sounds, intuitively, that any formula can be useful in our attempt to prove a theorem under the condition that an appropriate intermediate formula is chosen. Recall that, replacing $\vdash_{ef}$ by $\vdash$, this condition becomes trivial, as $V(a, i, s) \in Sup(V(a', i', s'))$ and $V(a', i', s') \in Sup(Y)$ implies immediately that $V(a, i, s) \in Sup(Y)$. More sophisticated results could be obtained using a generalized n-steps intermediate deducibility condition, but this problem will not be investigated here.

The conditions of Theorem 6 require not only that Y were proclaimed to be a theorem, but also that it were, indeed, a theorem. This assumption is not too natural, as the only we know (or the automaton knows) about the tested formula is the result of the statistical deducibility testing procedure. The next theorem shows that the assumption $Y \in \mathscr{T}$ can be replaced by another one, namely by that of a restriction of the cost of unsuccessful experiments. Put, for a real $K > 0$,

$$C_K(EXP_V(a, i, s)) = \min \{K, C(EXP_V(a, i, s))\}$$

and define $C_K(EXP_V(\bigvee_{j=1}^{n} V(a_j, i_j, s_j)))$ and $C_K(Y, \varepsilon, \omega)$ as above, just using C_K instead of C (a similar idea as in Theorem 5). Clearly, $C_K(Y, \varepsilon, \omega) \leqq K$ for any $\omega \in \Omega$ and $\varepsilon > 0$.

Theorem 7. Let a real $K > 0$ be given, let the cost C of experiments be continuous with respect to $\vdash_{ef}$, let the intermediate deducibility condition hold, let α be a random variable defined on the probability space $\langle \Omega, \mathscr{S}, P \rangle$ and taking its values in the set $\mathscr{L}$ of all formulas. Then, for $\varepsilon \to 0$,

$$E(C_K(Y, \varepsilon, \omega) / \{\omega : T(\alpha(\omega), \varepsilon, \omega) = t\}) \to 0 .$$

Proof. Clearly, as the expected value is a linear functional, we have

$$E(C_K(Y, \varepsilon, \omega) / \{\omega : T(\alpha(\omega), \varepsilon, \omega) = t\}) =$$

$$= E(C_K(Y, \varepsilon, \omega) / \{\omega : T(\alpha(\omega), \varepsilon, \omega) = t\} \cap \{\omega : \alpha(\omega) \in \mathscr{T}\}) P(\{\omega : \alpha(\omega) \in \mathscr{T}\}) +$$

$$+ E(C_K(Y, \varepsilon, \omega) / \{\omega : T(\alpha(\omega), \varepsilon, \omega) = t\} \cap \{\omega : \alpha(\omega) \in \mathscr{L} - \mathscr{T}\}) P(\{\omega : \alpha(\omega) \in \mathscr{L} - \mathscr{T}\}) .$$

The first component tends to 0 when $\varepsilon \to 0$ because of Theorem 6 and the fact that C_K is uniformly bounded by K. The other component can be majorized by the expression

$$K\, P(\{\omega : T(\alpha(\omega), \varepsilon, \omega) = t\} \cap \{\omega : \alpha(\omega) \in \mathscr{L} - \mathscr{T}\}) \leqq K\varepsilon$$

according to the fact that the first order probability of error is supposed to be majorized uniformly by ε. Hence, both the components tend to 0 when $\varepsilon \to 0$ and the theorem is proved. Q.E.D.

A more detailed investigation of our statistically based mechanized experiment planning procedure seems to be justifiable specially in the case of some concrete automaton-environment system, as its features could justify some more conditions imposed to the cost C, testing procedure T etc. and these conditions would enable to prove some more detailed properties concerning the proposed experiment planning procedure. As the kind of such an investigation would differ significantly from the pure mathematical considerations presented in this paper we have decided to limit ourselves, in this paper, to the theoretical reasonings in the extent given above.

(Received November 6, 1976.)

[1] А. Н. Богомолов, В. А. Твердохлебов: Целенаправленное поведение автоматов. „Наукова думка", Киев 1975.

[2] I. M. Havel, I. Kramosil: Randomized Incidental Phenomena: A New Tool for Robot Decision Making. In preparation.

[3] I. M. Havel, I. Kramosil: Probabilistic Methods in Robot Decision Making. In: Celostátní konference o kybernetice, Praha, listopad 1976, sborník prací, 66–80.

[4] I. Kramosil: Statistical Estimation of Deducibility in Formalized Theories. Proceedings of the Fourth Conference on Probability Theory, 1971, Brasov, 281–298.

[5] I. Kramosil: A Method for Statistical Testing of an at Random Sampled Formula. Kybernetika *9* (1973), 3, 162–173.

[6] I. Kramosil, L. Zwinogrodzki: Statistical Methods for Comparing Theorem Proving Algorithms. Kybernetika *10* (1974), 3, 221–240.

[7] I. Kramosil: A Probabilistic Approach to Automaton-Environment Systems. Kybernetika *11* (1975), 3, 173–206.

[8] I. Kramosil: A Selection-Based Formal Description of an Environment. Kybernetika *12* (1976), 3, 127–150.

[9] L. Siklóssy, J. Dreussi: Simulation of Executing Robots in Uncertain Environments. Research Report TR-16, Computer Science Department, University of Texas, 1973.

[10] A. Špaček: Statistical Estimation of Provability in Boolean Logic. Transactions of the Second Prague Conference on Information Theory, Statistical Decision Functions, Random Processes, Prague 1960, 609–626.

[11] O. Štěpánková, I. M. Havel: A Logical Theory of Robot Problem Solving. Artificial Intelligence *7* (1976), 129–161.

[12] O. Štěpánková, I. M. Havel: Incidental and State-Dependent Phenomena in Robot Problem Solving. In: Proceedings of the AISB Summer Conference, Edinburgh 1976, 266–278.

Dr. Ivan Kramosil, CSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 180 76 Praha 8. Czechoslovakia.