TRABAJOS DE ESTADISTICA Y
DE INVESTIGACION OPERATIVA
Vol. 32, Núm. 1, 1981, pp. 116 a 129

A MATHEMATICAL FRAMEWORK FOR LEARNING AND ADAPTION: (GENERALIZED) RANDOM SYSTEMS WITH COMPLETE CONNECTIONS*

Ulrich Herkenrath (Bonn)
and
Radu Theodorescu (Quebec)

ABSTRACT

The aim of this paper is to show that the theory of (generalized) random systems with complete connections may serve as a mathematical framework for learning and adaption. Chapter 1 is of an introductory nature and gives a general description of the problems with which one is faced. In Chapter 2 the mathematical model and some results about it are explained. Chapter 3 deals with special learning and adaption models.

1. Introduction

In various fields of science one is faced with systems which evolve on two different, but connected levels. Therefore such systems induce two processes so that the evolution of one process is adapted to the evolution of the other process. According to specific concrete situations this adaption may also be called learning. We want to discuss such phenomena within a mathematical framework.

* Work supported by the Deutsche Forschungsgemeinschaft and by the Natural Sciences and Engineering Research Council Canada grant A-7223. *AMS 1970 subject classification:* Primary 60G99; Secondary 90A15, 92A25, 94A35.

Key words and phrases. Random systems with complete connections, random automata, random environment, expediency, optimality, two-armed bandit problem.

The abstract mathematical model fitting the description of phenomena we are interested in is the so-called generalized random system with complete connections (abbreviated to GRSCC) and which consist of two sets, called *state space* and *event space*, and two *transition laws* which govern the evolution of the system on a step-by-step basis. It is always assumed that the system operates in discrete time. The first transition law describes the transition from a state to an event and the second the transition from a state and an event to a new state. These transition laws are not in general deterministic but depend on chance (or eventually on effects which are out of our knowledge so that they seem to happen in a random fashion).

There are basically two aspects of a theory dealing with learning and adaption. One aspect concerns the study of learning and adaption systems with fixed transition laws, e.g., with the aim to discuss its long-run behaviour, whereas the second aspect refers to situations when one can influence or control the transition laws in order to achieve a specific evolution. Therefore we distinguish between systems with fixed transition laws and systems with controllable transition laws.

A very general situation to which the second aspect is devoted is the following. A subject (this may also be an automaton) operates within an environment, i.e., it has to make choices and decisions dependent on its decisions and the evolution of the environment up to this time, which in turn influence the further development of the whole system. The system is constituted by the subject, the environment, and their laws of motion and interaction. Each special possible sequence of the subject's decisions and states of the environment resulting from the evolution of the system is evaluated by a real number (= value of this path). The subject wants to behave, in a certain sense, optimal or at least expedient with respect to the function evaluating the paths of the system. The problem is, that the subject does not know completely the rules or laws of motion which govern the system. So the subject is faced with the problem to increase its knowledge or information about the system's rules, because a gain of information will enable it to behave better. Since the only possibility to increase information about the system consists in observing responses of the environment to actions taken by the subject, the increase of information is a learning problem.

We think that the theory of GRSCC's provides an appropriate mathematical framework in which to study stochastic processes arising from learning and adaption mechanisms and to judge given learning procedures with respect to their efficiency or to find most efficient, i.e., optimal procedures. We have the advantage of having at our disposal a well developed theory which we can use in analyzing specific models. In Chapter 2 we present the concept of a GRSCC, a natural extension of the concept of a random system with complete connections (abbreviated to RSCC), and we indicate some of the results about it. We further illustrate the concepts of GRSCC and RSCC in Chapter 3 by means of special models from psychology, system theory, economics, and statistics.

2. The mathematical model

In Section 2.1 we present the concept of a GRSCC and explain its relation to interacting processes. We indicate some of the results known about (G) RSCC's in Section 2.2.

2.1. *Definition and description*

2.1.1. We say that a tuple $\{(W, \mathbf{W}), (X, \mathfrak{X}), \Pi, P\}$ together with the set of natural numbers $\mathbb{N}$ serving as parameter set, is a *GRSCC* if and only if the following conditions are fulfilled: (i) $(W, \mathbf{W})$ and $(X, \mathfrak{X})$ are measurable spaces called the *state* and *event space*, respectively; (ii) Π is a stochatic kernel from $W \times X$ to $\mathbf{W}$ (i.e., Π is a function Π: $W \times X \times \mathbf{W} \to [0,1]$ such that for all $(w, x) \in W \times X$, $\Pi(w, x, .)$ is a probability measure on $\mathbf{W}$ and for all $B \in \mathbf{W}$, $\Pi(., ., B)$ is a measurable function); (iii) P is a stochastic kernel from W to $\mathfrak{X}$.

Furthermore we say that a GRSCC is an *RSCC* if and only if the following condition is fulfilled: there is a measurable application u from $W \times X$ into W such that $\Pi(w, x, B) = 1_B(u(w, x))$ for all $w \in W$, $x \in X$, and $B \in \mathbf{W}$, where 1_B stands for the indicator of B.

GRSCC's were introduced by LeCalvé and Theodorescu [19] and represent a natural generalization of the concept of an RSCC, which was defined by Iosifescu [15] (see also [17], p. 63). An RSCC in its turn generalizes the concepts of an OM-chain and that of a chain of infinite order, which were studied by Onicescu and Mihoc [24] and Doeblin and Fortet [4] respectively in the 1930ies (see [17]). Let us remark that the theory of these stochastic processes began with Onicescu and Mihoc [24] and was motivated by the study of certain contagion schemes and actuarial models.

The definition of a (G) RSCC can be extended to the nonhomogeneous case in the sense that all entities constituting it are allowed to depend on $t \in \mathbb{N}$. Up to a certain degree, nonhomogeneous GRSCC's can be reduced to homogeneous ones (see [17] and [18]).

2.1.2. A GRSCC describes the following interaction process (interaction between a state and event process) which reflects learning and adaption mechanisms.

The system starts at time instant $t = 1$ in a given state $w_1 \in W$. According to the stochastic kernel P an event $x_1 \in X$ takes place at $t = 1$ as a realization of the probability distribution $P(w_1, .)$. Given w_1 and x_1 the system moves to a new state $w_2 \in W$ at time instant $t = 2$ according to the stochastic kernel Π. Now in period $t = 2$ an event $x_2 \in X$ is observed as a realization of $P(w_2, .)$ etc. RSCC's are characterized by the fact that the transition from state and event to the next state is not stochastic but deterministic, i.e., given w_1 and x_1 then state w_2 equals $u(w_1, x_1)$.

2.2. *Some mathematical results*

2.2.1. Let $\{(W, \mathbf{W}), (X, \mathfrak{X}), \Pi, P\}$ be a GRSCC. For each starting element $w \in W$ there are (see [19]) a probability space $(\Omega, K, \mathbb{P}_w)$ and two sequences of random variables $\{\zeta_n : n \geq 1\}$ and $\{\zeta_n : n \geq 1\}$ on it with values in W and X respectively such that the state sequence $\{\zeta_n : n \geq 1\}$ and the event sequence $\{\zeta_n : n \geq 1\}$ are connected as explained above, i.e., the system moves from ζ_n to ξ_n by means of P and from (ζ_n, ξ_n) to ζ_{n+1} by means of Π. Moreover it turns out that the state sequence $\{\zeta_n : n \geq 1\}$ is a general Markov chain with initial state W and transition probability function $Q(w, B) = \int_X \Pi(w,x,B)\, P(w,dx)$, $w \in W$, $B \in \mathbf{W}$. $\{\zeta_n : n \geq 1\}$ and $\{\xi_n : n \geq 1\}$ are called the *associated processes* of a GRSCC.

An interesting fact is that the concept of a discrete parameter stochastic process with values in a Polish space and that of an RSCC can be regarded as equivalent. Indeed for each probability space $(\Omega, K, \mathbb{P})$ and each sequence of random variables $\{\xi_n : n \geqslant 1\}$ defined on it with values in a Polish space X there is an RSCC such that the sequence $\{\xi_n : n \geqslant 1\}$ is the associated (event process of this RSCC (see [23]).

2.2.2. The main questions which arise in the study of GRSCC's concern convergence or limit properties of the associated processes. If one deals with the associated state process $\{\zeta_n : n \geqslant 1\}$ with starting point w_1 we can take advantage of its Markov property and make use of the theory of general Markov chains. So, if we assume that W and X are metric spaces both endowed with the σ-algebras of Borel sets and if W is compact, then we can show under certain continuity assumptions on the underlying stochastic kernels Π and P that $(1/n) \sum_{j=1}^{n} Q^j (w,.)$ tends to a limit $Q^\infty (w,.)$ as $n \to \infty$. Under additional assumptions which guarantee that "enough" probability mass distributed according to $\Pi (w, x, .)$ is spread out over the whole W, we get that $Q^n (w, .)$ converges to a limit $Q^\infty (\cdot)$ as $n \to \infty$ which does not depend on the starting point w. These results are explained and proved in [8].

The important special class of RSCC's often has to be handled separately, since there are some conditions appropiate for RSCC's which make no sense for GRSCC's. So the associated Markov process of a RSCC can be studied under contraction assumptions on u and P (see [17] and [22]).

Note that convergence properties of the n-step transition probability Q^n of $\{\zeta_n : n \geqslant 1\}$, i.e., $Q^n (w, A) = \mathbb{P}_w (\zeta_{n+1} \in A)$, imply corresponding properties for P^n, where P^n denotes the distribution of ξ_n, i.e., $P^n (w, A) = \mathbb{P}_w (\xi_n \in A)$. This is due to the following relation: $P^n (w, A) = \int_W Q^{n-1} (w, dw') P (w', A)$.

The associated event process $\{\xi_n : n \geqslant 1\}$ of a (G) RSCC can also be studied directly. This is done for GRSCC's in [19]. For a comprensive analysis of the associated processes of an RSCC we refer to the monographs of Iosifescu and Theodorescu [17] and Norman [22]. Recent results concern also functions of the event variables ξ_n, $n \geqslant 1$ (see [25]) and classification criteria (see [26] and [27]).

2.2.3. If we want to control, i.e., in a certain sense to optimize the functioning of a GRSCC, it is necessary to introduce a real-valued function $f\colon X \to \mathbb{R}$ which evaluates the events. If f plays the role of a penalty function, then we are interested in finding transition laws Π which lead to a reduction or eventually to a minimum of the induced penalty, having only partial or no knowledge about P. From a formal view point this problem is complementary to the problem of dynamic programming (see [14]) (we identify W with the state space of the program and X with the action space), where Π is the (known) law of nature and P under the control of the statistician.

If we want to characterize the learning performance of a transition law Π within a GRSCC by means of the penalty function f, we are led to consider the expected penalty in period n given the state ζ_n at period n:

$$\rho_n = \int_X f(x)\, P(\zeta_n, dx) = r(\zeta_n).$$

Given a history or path of the system up to ζ_n, i.e., given values for $\zeta_1 = w,\ \xi_1, \ldots, \zeta_n$, then we can compare the expected penalty in time n and in period $(n+1)$ conditioned by the path up to ζ_n. If for all $n \geqslant 1$ the latter one is less than or equal to the first one, then we call the transition law Π under consideration *absolutely expedient with respect to P*, or if P is not known exactly but only that $P \in \mathcal{P}_0$, where $\mathcal{P}_0$ is a subset of the set of all transition laws P from W to $\mathfrak{X}$, we call Π *absolutely expedient with respect to* $\mathcal{P}_0$. In other words, we call a transition law Π absolutely expedient with respect to $\mathcal{P}_0$, if and only if

$$\mathbb{E}_w\,(\rho_{n+1} \mid \zeta_1, \ldots, \zeta_n) = \mathbb{E}_w\,(\rho_{n+1} \mid \zeta_n) \leqslant \rho_n \quad \mathbb{P}_w - \text{a.s.}$$

for all $n \geqslant 1$, $w \in W$, and $P \in \mathcal{P}_0$. Here $\mathbb{E}_w$ denotes the expectation with respect to $\mathbb{P}_w$.

A sufficient condition (see [10, 11]) for absolute expediency with respect to $\mathcal{P}_0$ is, that the inequality

$$\int_X f(x)\, P^2(w, dx) \leqslant \int_X f(x)\, P(w, dx)$$

holds for all $w \in W$ and $P \in \mathcal{P}_0$. This criterion often permits with a reasonable amount of computation the verification of absolute expediency.

A very useful consequence of absolute expediency is that the sequence $\{\rho_n : n \geq 1\}$ constitutes a supermartingale with respect to the increasing sequence of σ-algebras generated by the sequence $\{\zeta_n : n \geq 1\}$. Therefore the sequence $\{\rho_n : n \geq 1\}$ converges a.s., if, e.g., the ρ_n s are all nonnegative. This in turn allows to draw conclusions about the asymptotic expected penalty and often about convergence of the sequence $\{\zeta_n : n \geq 1\}$ itself.

Based on the expected penalty, we can define an asymptotic optimality criterion for transition laws Π by asking that this asymptotic expected penalty takes its least possible value. Whereas absolute expediency is a structural, i.e., a qualitative property of a transition law, optimality is a quantitative concept. So for a given scheme only very special transition laws will be optimal: perhaps we can only find nearly optimal ones. Therefore optimality and ϵ-optimality (i.e., nearly optimal laws) are defined for a whole family $\mathcal{F}$ of transition laws with respect to a set of admissible transition laws $\mathcal{P}_0$, in the sense that we call a family $\mathcal{F}$ of transition laws from $W \times X$ to $\mathbf{W}$ *optimal with respect to* $\mathcal{P}_0$, if and only if for each $w \in W$ there is a transition law $\Pi \in \mathcal{F}$ such that

$$\lim_{n \to \infty} \mathbb{E}_w^{\Pi} \rho_n = \inf_{w \in W} \int_X f(x) P(w, dx)$$

for all $P \in \mathcal{P}_0$. Here $\mathbb{E}_w^{\Pi}$ denotes the expectation with respect to the probability measure $\mathbb{P}_w$ when $\Pi \in \mathcal{F}$ is used. Therefore $\lim_{n \to \infty} \mathbb{E}_w^{\Pi} \rho_n$ is the asymptotic expected penalty, if the system starts in w and the transition law $\Pi \in \mathcal{F}$ is applied.

Furthermore we call a family $\mathcal{F}$ of transition laws from $W \times X$ to $\mathbf{W}$ *ϵ-optimal with respect to* $\mathcal{P}_0$, if and only if for each $w \in W$ and $\epsilon > 0$ there is a transition law $\Pi \in \mathcal{F}$ such that

$$\lim_{n \to \infty} \mathbb{E}_w^{\Pi} \rho_n - \inf_{w \in W} \int_X f(x) P(w, dx) < \epsilon$$

for all $P \in \mathcal{P}_0$. This means that under the transition law Π the asymptotic expected penalty approaches for all $P \in \mathcal{P}_0$ its least possible value up to ϵ.

3. Special models

In what follows we intend to discuss some models with fixed transition laws in Section 3.1 and some models with controllable transition laws in Section 3.2. The link between the first kind and the second kind of models will be the well known Bush-Mosteller learning model. For some other mathematical models dealing with learning and adaption we refer to the survey papers of Fu [5] and Narendra and Thathachar [21].

3.1. *Models with fixed transition laws*

3.1.1. In [2] and [20] Brock and Mirman study the growth of an economy with respect to a quantity "capital per head". In their model capital is an amount of money which indicates the value of all goods and resources available in the economy. In each time period of the development of the economy a certain amount of this capital is invested into a production process and the rest is free for consumption. Production and consumption are allowed to depend on chance. Under certain homogenity assumptions on the production and consumption functions we may consider the quantity capital per head, i.e., capital divided by the number of economic agents, instead of the total amount of capital. Then, if k_n is the capital per head at the beginning of period n and the environmental situation, i.e., the chance variable, takes on the value x_n, the output of the production process is given by $f(k_n, x_n)$, where f denotes the production function and the actual consumption is given by $c(k_n, x_n)$, where c denotes the consumption function which is chosen so as to maximize a certain utility criterion. Thus given k_n and the state of environment x_n capital per head at the beginning of period $(n+1)$ will be

$$k_{n+1} = f(k_n, x_n) - c(k_n, x_n).$$

Given an initial value k_1, we want to know, whether the adjustment from k_n to k_{n+1} by means of f, c, and the distribution of the chance variable leads to the convergence of the sequence of variables $\{k_n: n \geqslant 1\}$. A limit k_∞ of this sequence is usualy called *steady state* of the underlying economy.

The model described above can be brought into the form of an RSCC and analyzed by means of the methods known from there (see [7]).

3.1.2. The most well known learning model which is formalized as an RSCC is that introduced by Bush and Mosteller [3] (see also [16]) and abbreviated here to BM-model. In such a model a learning subject has to choose in each one of a sequence of time periods an alternative s out of a finite set S. As a consequence of its choice it gets a response t out of a finite set T according to a transition probability L from S to T. If the learning subject has an initial probability distribution m_1 over the set of alternatives reflecting its preference, then it chooses in period 1 an alternative s_1 according to m_1 and as consequence of the choice s_1 and the response t_1 following s_1, the subject's attitude to the different alternatives is changed by means of a function h; h assigns to each $m \in \mathcal{M}(S)$, (here $\mathcal{M}(S)$ is the set of all probability measures on S) and $(s, t) \in S \times T$ a new probability measure $h(m, s, t)$ on S. Obviously this model can be formalized as an RSCC by setting: $W = \mathcal{M}(S)$, $X = S \times T$, $u(w, x) = h(m, s, t)$, and

$$P(m, S' \times T') = \sum_{\substack{s' \in S' \\ t' \in T'}} L(s', t') \, m(s').$$

3.1.3. The extension of an RSCC to a GRSCC means for BM-models to allow that the change of subject's attitude to the alternatives from period to period does not only depend on the chosen alternative and the response, but also on chance. This amounts to the fact, that the function h is subsistuted by a transition probability Π. This model may be called a *generalized BM-model.*

In the context of a (generalized) BM-model the associated Markov process $\{\zeta_n: n \geqslant 1\}$ describes the subject's tendency for the different alternatives. So convergence of the sequence $\{\zeta_n: n \geqslant 1\}$ means that, as a consequence of the experience, the subject makes by choosing alternatives and being confronted with responses, the changes of its attitude to the alternatives become smaller and smaller as time goes on, so that the subject finally reaches an attitude which remains unchanged.

Another learning model from mathematical psychology which may be formalized as a GRSCC is discussed by Rhenius [28].

3.2. *Models with controllable transition laws*

3.2.1. A model which contains the generalized BM-model as a special case is the so-called *general control system* (abbreviated to GCS). It was introduced [10, 11] as a model for the interaction of a random automaton with a random environment. There is an extensive literature dealing with models of this kind. We refer to [5], [21], and [30].

The GCS consists of a random automaton A, a random environment E, and an interaction rule χ between E and A. A in turn consists of measurable spaces $(A, \mathcal{A})$, $(B, \mathcal{B})$, and $(C, \mathcal{C})$ called input space, output space, and state space of A respectively, and a stochastic kernel Γ from $C \times A$ to $\mathcal{C} \times \mathcal{B}$ (law of motion of A). E consists of two measurable spaces $(E, \mathcal{E})$ and $(B, \mathcal{B})$, called space of environmental situations and imput space of the environment respectively, and of a stochastic kernel Ψ from $E \times B$ to $\mathcal{E}$ (law of motion of E). Note that the space B respresents the outputs of the automaton as well as the inputs for the environment. The interaction rule χ is a stochastic kernel from $E \times B$ to $\mathcal{A}$ which produces the inputs for the automaton in dependence of the environmental situation. Then given starting points $(e, b, c) \in E \times B \times C$ there are stochastic processes representing the sequence of environmental situations, of outputs of the automaton, of states of the automaton, and of outputs of the environment = inputs for the automaton, which evolve according to the transition laws of the GCS.

With each GCS we can associate a GRSCC. Take: $W = E \times B \times C$; $X = A$; $\Pi\ (e, b, c, a, E' \times B' \times C') = \Psi\ (e, b, E')\ \Gamma\ (c, a, B' \times C')$ for $(e, b, c) \in E \times B \times C, a \in A, E' \in \mathcal{E}, B' \in \mathcal{B}$, and $C' \in \mathcal{C}$; $P\,(e, b, c, A') = \chi\ (e, b, A')$ for $(e, b, c) \in E \times B \times C$ and $A' \in \mathcal{A}$.

The method of deriving convergence properties of the associated

Markov process $\{\zeta_n : n \geqslant 1\}$ by means of continuity assumptions on the underlying stochastic kernels, which is known from the theory of GRSCC's (see [8]), is applied in [9] to study the associated processes of a GCS.

If the outputs of the environment = inputs for the automaton are evaluted by a (real-valued) penalty function $f : A \to \mathbb{R}$, we are interested to know whether the transition law of the automaton has the property that it influences the evolution of the system to the effect that the resulting penalty values are reduced or minimized in a certain sense, whereas one has only partial or no knowledge about the transition law of the environment and the interaction rule.

In certain situations we can even control the transition law of the automaton, i.e., we can choose a transition law out of a certain set of transition laws. Our device should be such that we choose a law which behaves most efficiently in the above sense. This amounts to the problem to "learn" which automaton outputs lead by means of X to automaton inputs with small penalty values. This problem was investigated in [10, 11] by making use of the methods described in Section 2.2.3.

Let us note also that absolute expediency and (ϵ-) optimality of special BM-models are examined for example in [21].

3.2.2. Let us consider the so-called *two-armed bandit problem*, a problem well-known in statistics. Here a controller has to choose in each time period one of two experiments (= arms), the outcomes of which are success (= 0) or failure (= 1). The probabilities of success are unknown. The controller wants to minimize his expected proportion of failures. This will be achieved, if he plays the arm with the greater probability for success. So the player has to learn by experimenting what arm has the better success probability and simultaneously he has to minimize the proportion of failures. In [12] two learning or decision procedures for this problem are formalized by means of an RSCC with an additional penalty function. It is shown that the optimality concept for GCS's coincides with the criterion of minimizing the expected proportion of failures and that the two procedures are ϵ-optimal.

3.2.3. Rothschild [29] formulates a problem arising in economics, which may be studied in a simplified version as a two-armed bandit problem in which the sucess probabilities are not completely unknown, but an a priori distribution for these quantities is given. The problem consists of finding a most profitable price for an economic good. In each time period a store owner can charge a price for the good and as a result of his choice the good will be sold (= success) or not (= failure) in this time period. The probability for sale depends on the price which is claimed and this dependence is unspecified up to an a priori distribution over the success probabilities. Of course the store owner wants to behave in such a way as to maximize his expected total profit in case of a finite planning horizon or his expected discounted total profit in case of an infinite horizon. If only two prices are considered this leads to a two-armed bandit problem. The above problem of finding out a most profitable price may be studied also by means of a GRSCC with an additional penalty function, as well as by stochastic approximation techniques [13].

3.2.4. In mathematical economics several authors (e.g., [1] and [6]) have dealt with the problem of a planner, who has to quote a price for an economic good such that production and consumption of this good are equal, i.e., excess supply of the good is zero. If it is assumed that the excess supply is not uniquely determined by the price but depends on the price only by means of a stochastic kernel, which in turn is nearly unknown, then the planner is confronted with the problem to learn something about this atochastic kernel in order to quote the right price. This learning process, which should lead to an equilibrium price, is called *tatonnement process* in the economic literature and can be handled by means of a GRSCC.

REFERENCES

[1] Arrow K., Hahn F.: *General competitive analysis.* Holden Day, San Francisco 1972.

[2] Brock W. A., Mirman L. J.: *Optimal economic growth and uncertainty: The discounted case.* J. Econom. Theory 4 (1972), 479-513.

[3] Bush R. R., Mosteller F.: *Stochastic models for learning.* Wiley, New York 1955.

[4] Doeblin W., Fortet R.: *Sur les chaînes à liaisons complètes.* Bull. Soc. Math. France 65 (1937), 132-148.

[5] Fu K.S.: *Learning control systems. Review and outlook.* IEEE Trans. Automatic Control 15 (1970), 210-221.

[6] Green J. R., Majumdar M.: *The nature of stochastic equilibria.* Econometrica 43 (1975), 647-660.

[7] Herkenranth U.: *Der assoziierte Markov Prozess verallgemeinerter zufälliger Systeme mit vollständigen Bindungen.* Ph. D. Thesis, University of Bonn, Dept. Applied Math., Bonn, 1976.

[8] Herkenrath U.: *The associated Markov process of generalized random systems with complete connections.* Rev. Roumaine Math. Pures Appl. (1979), 24, 243-254.

[9] Herkenrath U.: *Modelle für die Interaktion eines stochastischen Automaten mit einer zufälligen Umwelt.* Z. Angew. Math. Mech. (1978) (to appear).

[10] Herkenrath U., Theodorescu R.: *General control systems.* Information Sci. 14 (1978) 57-73.

[11] Herkenrath U., Theodorescu R.: *Expediency and optimality for general control systems.* Coll. Internat. C.N.R.S., n° 276, Theorie de l'information, Cachan, July 4-8, 1977. Editions du C.N.R.S., Paris, 1978, 449-455.

[12] Herkenrath U., Theodorescu R.: *On certain aspects of the two-armed bandit problem.* Elektron. Informationsverarbeit. Kybermetik, 1978, 14, 527-535.

[13] Herkenrath U., Theodorescu R.: *Ona stochastic approximation procedure applied to the bandit problem.* Elektron, Informationsverarbeit. Kybernetik, 1979, 15, 301-307.

[14] Hinderer K.: *Foundations of nonstationary dynamic programming with discrete time parameter.* Springer, Berlin 1970.

[15] Iosifescu M.: *Random systems with complete connections with an arbitrary set of states.* Rev. Roumaine Math. Pures Appl. 8 (1963), 611-645; 9 (1964), 91-92.

[16] Iosifescu M., Theodorescu R.: *On the Bush-Mosteller stochastic models for learning.* Math. Psychology 2 (1965), 196-203.

[17] Iosifescu M., Theodorescu R.: *Random processes and learning.* Springer, New York 1969.

[18] LeCalvé G., Theodorescu R.: *Systèmes aléatoires à liaisons complètes totalelement non-homogènes.* C.R. Acad. Sci, Paris Ser. A 265 (1967), 347-349.

[19] LeCalvé G., Theodorescu R.: *Systèmes aléatoires généralisés à liaisons complètes.* Z. Wahrscheinlichkeitstheorie Verw. Gebiete 19 (1971), 19-28.

[20] Mirman L. J.: *The steady state behavior of a class of one sector growth models with incertain technology.* J. Econom. Theory 6 (1973), 219-242.

[21] Narendra K. S., Thathachar M.A.L.: *Learning automata - a survey.* IEEE Trans Systems, Man, Cybernet, 4 (1974), 323 - 334.

[22] Norman M. F.: *Markov processes and learning models.* Academic Press, New York 1972.

[23] Pruscha H.: *Die Statistische Analyse von ergodischen Ketten mit vollständigen Bindungen.* Ph. D. Thesis, University of Munich, Dept. Math., Munich 1974.

[24] Onicescu O., Mihoc G.: *Sur les chaînes de variables statistiques.* Bull. Sci. Math. 59 (1935), 174-192.

[25] Pruscha H., Theodorescu R.: *Functions of event variables of a random system with complete connections.* J. Multivariate Anal. 7 (1977), 336-362.

[26] Pruscha H., Theodorescu R.: *Recurrence and transience for discrete parameter stochastic processes.* Trans. Eighth Prague Conference on Information Theory, Statistical Decision Functions, and Random processes, Prague, Aug. 28-Sept. 1. 1978, Academia, Prague, 1978, 97-105.

[27] Pruscha H., Theodorescu R.: *On a non-Markovian model with linear transition rule.* Coll. Math., (in print).

[28] Rhenius D.: *Markoffsche Entscheidungsmodelle mit unvollständiger Information und Anwendungen in der Lerntheorie.* Ph. D. Thesis, University of Hamburg, Dept. of Stochastics, Hamburg, 1971.

[29] Rothschild M.: *A two-armed bandit theory of market princing.* J. Econom. Theory 9 (1974), 185-202.

[30] Simon J. C.: *Computer oriented learning processes.* Noordhoff, Leyden 1976.

U. Herkenrath
University of Bonn
Institute of Applied Mathematics
Wegelerstr. 6
5300 Bonn
Federal Republic of Germany

R. Theodorescu
Laval University
Department of Mathematics
Quebec, Que.
Canada
G1K 7P4