Теорія Ймовір. та Матем. Статист.
Вип. 69, 2003

Theor. Probability and Math. Statist.
No. 69, 2004, Pages 39–53
S 0094-9000(05)00612-5
Article electronically published on February 7, 2005

CONTROLLED SEMI-MARKOV FIELDS WITH GRAPH-STRUCTURED COMPACT STATE SPACE

UDC 519.21

H. DADUNA, P. S. KNOPOV, AND R. K. CHORNEY

Abstract. We introduce locally acting distributed decision makers in the theory of semi-Markov decisions for systems for which both the domain and the action space are general and compact. Such decision makers are characterized by making decisions on the basis of the information gathered at their local neighborhood only. The state transient function of the system also is of a local structure. We consider general holding times of the systems and this results in semi-Markov properties in time. The neighborhood structure of the systems resembles in space the Markov property of spatial processes. Under some regularity assumptions, we reduce the optimal problems within the set of local strategies to the corresponding problems for deterministic Markov strategies.

1. Introduction

The optimal sequential control of stochastic systems can be described in most cases by models of semi-Markov processes. But there is extensive literature on the control of processes where the memoryless property of Markov systems is not used. Introducing nonexponential, respectively nongeometrical, sojourn times of the processes leads to considering semi-Markov processes and Markov renewal processes. Early papers on the stochastic dynamic optimization for semi-Markov processes (sometimes called renewal reward processes) are [8, 9, 7]. For further references to early works see [6]. More recent works on semi-Markov decision processes can be found in [10, 17, 18, 13] and the references therein. The general construction of semiregenerative control processes and semi-Markov control processes with an application to queueing systems can be found in the recent book by Kitaev and Rykov [11].

The basic definitions and results of the theory of semi-Markov processes and Markov renewal processes considered in this paper can be found in [1, 2] and in [11], where there is a special emphasis on the optimization theory.

In this paper we consider stochastic processes in continuous time with general compact state spaces which are structured by an underlying graph. At any node of the graph there is a local state space such that the global state space of the process is the product (indexed by the graph) of the local state spaces. The graph then defines a neighborhood structure for the states of the systems. These neighborhoods determine the local interactions of coordinates of the spatial process. Then for a fixed time moment the random state of

2000 *Mathematics Subject Classification.* Primary 60K15, 60K35, 90C40.

Key words and phrases. Semi-Markov processes, Markov renewal processes, optimal control, average asymptotic reward, renewal reward processes, random fields, local strategies.

The research of the second author was partially supported by a grant from Deutsche Forschungsgemeinschaft at Hamburg University.

the system, respectively of the corresponding stochastic process, is a random field with respect to the neighborhood graph.

We assume that the process has the semi-Markov property in time and that the transition kernels of the process have the spatial Markov property with respect to the underlying graph. Additionally we construct control structures to optimize the development of the system over time. Optimality here is meant with respect to the asymptotic reward criterion averaged in time.

Our goal is to find optimal strategies to control the system where the decision makers located at the nodes of the network can use in the decision making only information gathered in their neighborhoods. We find conditions under which there exist optimal stationary deterministic policies in the class of the strategies described above.

In this sense, our results extend those of [5] in the direction of interacting semi-Markov processes.

We recall necessary definitions of the theory of stochastic processes that depend on the spatial coordinates and time in Section 2.1. In Section 2.2 we introduce local structures for the policies and the transition mechanism of the processes. We assume that time is continuous but decisions can be made only at jump moments of the process. The random policies have, as usual, a conditional independence structure. We assume that the transition kernels of the jump chains (the so-called synchronized kernels) are of a similar structure. For the Markov case we then are in the setting of the paper [3]. In Section 3 we first provide abstract conditions for the existence of stationary deterministic strategies in the class of local structures and then show that under some weak smoothness assumptions these conditions are fulfilled.

2. Description of the system

2.1. **Preliminaries.** For systems with locally interacting coordinates the interaction structure is defined in terms of an undirected finite neighborhood graph $\Gamma = (V, B)$ where V is the set of vertices (nodes) and B is the set of edges. Denote by $\{k, j\}$ the edge of the graph connecting vertices k and j. The neighborhood of a vertex k is the set of nodes $N(k) = \{j\colon \{k, j\} \in B\}$. The complete neighborhood of a vertex k is the set $\widetilde{N}(k) = N(k) \cup \{k\}$, that is, the neighborhood of the vertex k including k itself. For any subset $K \subset V$ we define the neighborhood

$$N(K) = \bigcup_{k\in K} N(k) \setminus K$$

and complete neighborhood $\widetilde{N}(K) = N(K) \cup K$. For every node $i \in V$ let some compact metric space X_i with countable basis be given (it is called the local state space at the node i). Then $X := \prod_{i\in V} X_i$ is the global state space of the system. For every subset of nodes $K \subset V$ we denote by $x_K = (x_k, k \in K) \in X_K = \prod_{i\in K} X_i$ the marginal vector of a state x.

Let $\mathfrak{X}_i$ be the σ-algebra of Borel subsets of X_i, $\mathfrak{X}_K = \sigma\{\prod_{i\in K} \mathfrak{X}_i\}$ the σ-algebra generated by $\prod_{i\in K} \mathfrak{X}_i$, $K \subseteq V$, and $\mathfrak{X}_V = \mathfrak{X}$.

Throughout the paper we fix the underlying probability space $(\Omega, \mathcal{F}, \mathsf{P})$ on which all random variables are defined.

Definition 2.1 (Random fields)**.** (1) A random variable ξ defined on $(\Omega, \mathcal{F}, \mathsf{P})$ and assuming values in X is called a random field over (V, B) (or simply a random field over V). We denote by ξ_K the marginal random variables with values in the space X_K. For $K = \{k\}$ we write ξ_k.

(2) A random field ξ over (V, B) is called a Markov field if

$$\mathsf{P}\left(\xi_k \in C_k \mid \xi_{V-\{k\}} = x_{V-\{k\}}\right) = \mathsf{P}\left(\xi_k \in C_k \mid \xi_{N(k)} = x_{N(k)}\right) \quad \text{for all } x \in X, \qquad C_k \in \mathfrak{X}_k, \tag{1}$$

and $k \in V$ where $\mathsf{P}(\xi_k \in C_k \mid \xi_{N(k)} = x_{N(k)})$ and $\mathsf{P}(\xi_k \in C_k \mid \xi_{V-\{k\}} = x_{V-\{k\}})$ are the regular conditional probability on $(X_k, \mathfrak{X}_k)$ given $\xi_{N(k)} = x_{N(k)}$ and $\xi_{V-\{k\}} = x_{V-\{k\}}$, respectively.

Remark 2.1. Equation (1) can be interpreted as follows:

$$\mathsf{P}\left(\xi_k \in C_k \mid \xi_{V-\{k\}} = x_{V-\{k\}}\right) = \mathsf{P}\left(\xi_k \in C_k \mid \xi_{V-\{k\}} = y_{V-\{k\}}\right), \quad \text{if } x_{N(k)} = y_{N(k)}.$$

Similar interpretations will apply throughout the paper.

In many problems related to biological, economical or engineering applications, random fields describe the state of a system at some fixed time moment. The evolution of that system over time is then described by a stochastic process η with random fields as one-dimensional marginals in time. In this case we write $\eta = (\eta^t)$ to emphasize the time-dependence. We assume that time is continuous, $t \in \mathbb{R}_+ = [0; \infty)$. The subscript k in η_k^t refers to the vertex k, thus η_k^t denotes the marginal distribution for time t and node k of some vector-valued process $\eta = (\eta^t : t \geq 0)$. Such processes, which vary in space and time and whose space variable is structured by some graph, are investigated in [3] using Markov chain techniques. It is the goal of this paper to drop the assumption that the holding times of the processes are memoryless, so that we investigate semi-Markov processes and Markov renewal processes.

Definition 2.2 (Markov renewal processes and semi-Markov processes)**.** Consider a stochastic jump process $\eta = (\eta^t : t \geq 0)$ whose state space is X. Assume that its paths are càdlàg, that is they are continuous from the right and have limits on the left. Corresponding to the process η is the sequence $(\xi, \tau) = \{(\xi^n, \tau^n),\ n = 0, 1, \ldots\}$. The $\mathbb{R}_+$-valued random variables τ^n are the interjump times and $\{\xi^n, n = 0, 1, \ldots\}$ is the sequence of states visited by the process just after the jumps. More precisely:

> Let $\sigma = \{\sigma^n : n = 0, 1, \ldots\}$ be the increasing sequence of jump times such that $\sigma^0 = 0$ and $\sigma^n = \sum_{i=0}^{n-1} \tau^i$, $n > 0$. Then $\eta^t = \xi^n$ if $t \in [\sigma^n, \sigma^{n+1})$, $n \in \mathbb{N}$.

If conditional distributions of the sequence $(\xi, \tau) = \{(\xi^n, \tau^n), n = 0, 1, \ldots\}$ do not depend on n and

$$\begin{aligned}\mathsf{P}&\left(\xi^n \in C, \tau^{n-1} \leq s \mid \xi^k = x^k, k = 0, 1, \ldots, n-1, \sigma^0 = 0, \tau^k \leq s^k, k = 0, 1, \ldots, n-2\right) \\ &= \mathsf{P}\left(\xi^n \in C, \tau^{n-1} \leq s \mid \xi^{n-1} = x^{n-1}\right), \qquad n > 0,\end{aligned}$$

then (ξ, τ) is called a homogeneous Markov renewal process. The process

$$\eta = \left(\eta^t : 0 \leq t < \sum_{i=0}^{\infty} \tau^i\right)$$

is called a homogeneous semi-Markov process. The transition function

$$\widetilde{Q}(C, s \mid x) = \mathsf{P}\left(\xi^{n+1} \in C, \tau^n \leq s \mid \xi^n = x\right)$$

is called a semi-Markov kernel both in the case of Markov renewal process and of semi-Markov process.

In what follows we assume that

$$\widetilde{Q}(C, 0 \mid x) = 0 \quad \text{for all } x \text{ and } C, \qquad \sum_{i=0}^{\infty} \tau^i = \infty,$$

and that $\sigma = \{\sigma^n \colon n = 0, 1, \dots\}$ has no finite accumulation point. The transition probability of the embedded Markov chain $\xi = (\xi^n = \eta(\sigma^n) \colon n \in \mathbb{N})$ is

$$Q(C \mid x) = \widetilde{Q}(C, \infty \mid x) = \mathsf{P}\left(\xi^{n+1} \in C, \tau^n \leq \infty \mid \xi^n = x\right) \qquad \text{for all } x \in X \text{ and } C \in \mathfrak{X}.$$

Let

$$\mathrm{T}(\cdot \mid x, y) = \mathsf{P}\left(\tau^n \leq \cdot \mid \xi^n = x, \xi^{n+1} = y\right), \qquad x, y \in X,$$

be the conditional distribution of the sojourn time τ^n given that the process is in a state x and will pass to a state y after τ^n is expired.

Before we present the details of the control structures for the semi-Markov processes we recall some definitions and theorems from the book [12].

Definition 2.3. Let X and A be topological spaces equipped with the Borel σ-algebras $\mathfrak{X}$ and $\mathfrak{A}$, respectively. Denote by $(2)^A_{\text{set}}$ the set of all subsets of A. Let

$$F \colon X \to (2)^A_{\text{set}}$$

be a set-valued function which associates with $x \in X$ a nonempty closed set

$$F(x) =: A(x) \subseteq A$$

such that

$$\Delta = \{(x, a) \colon x \in X,\, a \in A(x)\} \subset X \times A$$

is Borel measurable. In what follows the function F determines the set of admissible decisions (controls) $A(x)$ for any state x.

A map $F \colon X \to (2)^A_{\text{set}}$ is called open, closed, or Borel measurable if

$$\{x \colon F(x) \cap E \neq \varnothing\} \in \mathfrak{X}$$

for all open, closed, or Borel measurable subsets E of A, respectively.

A function $f \colon X \to A$ is called a selector for a map F if $f(x) \in F(x)$ for all $x \in X$.

Definition 2.4. Let X and A be topological spaces equipped with the Borel σ-algebras $\mathfrak{X}$ and $\mathfrak{A}$, respectively. A closed (open) measurable map F is called upper (lower) semicontinuous if the set $\{x \in X \colon F(x) \cap E \neq \varnothing\}$ is closed (open) for every closed (open) set $E \subseteq A$.

A map F is continuous if it is both upper and lower semicontinuous.

Theorem of Choice ([12, p. 74])**.** *Any Borel measurable map $F \colon X \to (2)^A_{\text{set}}$ has a Borel measurable selector*

$$f \colon (X, \mathfrak{X}) \to (A, \mathfrak{A}).$$

Theorem of Choice for semicontinuous maps ([12, p. 71])**.** *If A is a compact metric space, then any upper (lower) semicontinuous map $F \colon X \to (2)^A_{\text{set}}$ has a selector belonging to the Baire class* 1.

2.2. **Semi-Markov processes with local and synchronous transition mechanisms.** The goal of this section is to introduce a controlled semi-Markov process with locally interacting synchronous components, namely: a controlled semi-Markov random field similar to that constructed for the pure Markovian framework in [3] as controlled Markov random field.

If a semi-Markov process η and associated Markov renewal process (ξ, τ) are related to a system of neighborhoods $\{N(k)\colon k \in V\}$ on (V, B), then it is natural to assume that with respect to the evolution over time the value $\xi_k^n = \eta_k(\sigma^n) = x_k$ of the kth vertex depends on the previous states of the whole system only through the values of the vertices in $\widetilde{N}(k)$ (including k) after transition moment σ^{n-1}. To describe this property we introduce the spatial Markov property for the kernel of the embedded jump chain for a semi-Markov process. This can be done similarly to Definition 2.1.

Since we concentrate on the local behavior of the embedded jump chain, we need the definition of synchronous and local kernels as described in [3, Section 2]. Note that the distributions of the sojourn time of the semi-Markov process remain globally determined.

Definition 2.5 (Synchronized and local transition kernels)**.** Let

$$\eta = (\eta^t \colon t \geq 0)$$

be a semi-Markov process with the state space X such that the associated Markov renewal process is $(\xi, \tau) = \{(\xi^n, \tau^n), n = 0, 1, \ldots\}$. The jump transition probabilities of η, respectively ξ, are called *local* [16, p. 100] if

$$\mathsf{P}\left\{\xi_k^{n+1} \in C_k \mid \xi^n = x^n, \ldots, \xi^0 = x^0\right\} = \mathsf{P}\left\{\xi_k^{n+1} \in C_k \mid \xi^n_{\widetilde{N}(k)} = x^n_{\widetilde{N}(k)}\right\} \tag{2}$$

for all $k \in V$, $x^0, \ldots, x^{n+1} \in X$, and $C_k \in \mathfrak{X}_k$, that is, the state of a vertex k just after the jump depends only on the state of its complete neighborhood at the preceding transition moment. The transition probabilities of X are called *synchronous* [16, p. 100] if

$$\mathsf{P}\left\{\xi_K^{n+1} \in C_K \mid \xi^n = x^n\right\} = \prod_{k \in K} \mathsf{P}\left\{\xi_k^{n+1} \in C_k \mid \xi^n = x^n\right\} \tag{3}$$
$$\text{for all } K \subset V \text{ and } x^n \in X$$

where $C_K = \prod_{k \in K} C_k \in \mathfrak{X}_K$.

If (2) and (3) hold for the process η, respectively for ξ or (ξ, τ), then η is called a *semi-Markov process with locally interacting synchronous components over* (Γ, X), shortly, a *semi-Markov random field*. We call (ξ, τ) a *Markov renewal process with locally interacting synchronous components over* (Γ, X), shortly a *Markov renewal field*.

Now we introduce the classes of admissible policies to control the interacting coordinates of the stochastic processes under consideration. Let (for simplicity) $\eta = (\eta^t \colon t \geq 0)$, or $(\xi, \tau) = \{(\xi^n, \tau^n)\colon n \in \mathbb{N}\}$ as defined in Definition 2.2, be a (uncontrolled) semi-Markov random field. In what follows we assume that the random field η is equipped with a control structure which governs the jump transition kernel and the distributions of the sojourn time of (ξ, τ). We mainly consider controls which are local in the sense to be specified now.

Definition 2.6 (Action spaces and local restrictions)**.** The sequence of decision moments (control moments) is $\tau = \{\tau^n, n = 0, 1, \ldots\}$.

(1) The set of actions (control values) that can be used at control moments is

$$A = \prod_{i \in V} A_i$$

where A_i is the set of possible actions (decisions) for the vertex i. Assume that A_i is a compact metric space with a countable basis and with the Borel σ-algebra $\mathfrak{A}_i$; $\mathcal{A}$ is the Borel σ-algebra in A.

(2) For a decision maker at a node i at time σ^n with a state $\xi^n = x$, let the set of control actions be restricted to $A_i^n(x) \subset A_i$. The set $A_i^n(x)$ is called the set of admissible actions (decisions) at time σ^n in state x.

(3) Assume that the set-valued maps

$$F_i^n \colon X \to (2)_{\text{set}}^{A_i}, \qquad x \to F_i^n(x) = A_i^n(x), \qquad i \in V,$$

depend on $x \in X$ only through $x_{\widetilde{N}(i)} \in X_{\widetilde{N}(i)}$ and are Borel measurable according to Definition 2.3. The maps F_i^n determine the admissibility as a local property. Thus we write $A_i^n(x) =: A_i^n(x_{\widetilde{N}(i)})$. We also assume that the sets

$$\Delta_i^n = \left\{ \left(x_{\widetilde{N}(i)}, a_i\right) \colon x_{\widetilde{N}(i)} \in X_{\widetilde{N}(i)}, a_i \in A_i^n\left(x_{\widetilde{N}(i)}\right) \right\}$$

are Borel measurable in the product space $X_{\widetilde{N}(i)} \times A_i$, and $\Delta^t = \prod_{i \in V} \Delta_i^t$ is Borel measurable in the product space $\widetilde{X} \times A$, where $\widetilde{X} = \prod_{i \in V} X_{\widetilde{N}(i)}$ and $\widetilde{\mathfrak{X}} = \sigma\{\prod_{i \in V} X_{\widetilde{N}(i)}\}$.

Definition 2.7 (Strategies)**.** (1) Let α_i^n denote the action chosen by the decision maker at the node i at the nth transition moment σ^n, and let $\alpha^n := (\alpha_i^n \colon i \in V)$ be the joint decision vector at time σ^n, $n \in \mathbb{N}$.

(2) A strategy (policy) δ to control the system with interacting components is defined as a vector of coordinate policies $\delta = (\delta_i, i \in V)$ where, for the node i,

$$\delta_i = \{\pi_i^0, \dots, \pi_i^n, \dots\}$$

is the sequence of transition probabilities

$$\pi_i^n = \pi_i^n \left(\cdot \mid x^0, a^0, t^0, \dots, x^{n-1}, a^{n-1}, t^{n-1}, x^n \right).$$

Thus π_i^n is a probability measure on $(A_i, \mathfrak{A}_i)$ for all

$$\left(x^0, a^0, t^0, \dots, x^{n-1}, a^{n-1}, t^{n-1}, x^n\right).$$

This measure depends measurably on the history

$$h^n = \left(x^0, a^0, t^0, \dots, x^{n-1}, a^{n-1}, t^{n-1}, x^n\right)$$

of the system up to the nth transition. Thus

$$\begin{aligned} &\mathsf{P}\big(\alpha_i^n \in B_i \mid \xi^0 = x^0, \alpha^0 = a^0, \tau^0 = t^0, \dots, \\ &\qquad\qquad \xi^{n-1} = x^{n-1}, \alpha^{n-1} = a^{n-1}, \tau^{n-1} = t^{n-1} \xi^n = x^n\big) \\ &\quad = \pi_i^n \left(B_i \mid x^0, a^0, t^0, \dots, x^{n-1}, a^{n-1}, x^n\right) \end{aligned}$$

for all $B_i \in \mathfrak{A}_i$.

(3) In parallel to the synchronous transitions and the locality of the transition kernels we always assume that the decision makers located at the nodes act conditionally independently given the history of the system. This leads to a control of the process

governed by a synchronous control kernel

$$\mathsf{P}\Big(\alpha^n \in \prod_{i\in V} B_i \mid \xi^0 = x^0, \alpha^0 = a^0, \tau^0 = t^0, \dots,$$
$$\xi^{n-1} = x^{n-1}, \alpha^{n-1} = a^{n-1}, \tau^{n-1} = t^{n-1}, \xi^n = x^n\Big)$$
$$= \prod_{j\in V} \mathsf{P}\big(\alpha_i^n \in B_i \mid \xi^0 = x^0, \alpha^0 = a^0, \tau^0 = t^0, \dots,$$
$$\xi^{n-1} = x^{n-1}, \alpha^{n-1} = a^{n-1}, \tau^{n-1} = t^{n-1}, \xi^n = x^n\big)$$
$$= \prod_{j\in V} \pi_i^n \left(B_i \mid x^0, a^0, t^0, \dots, x^{n-1}, a^{n-1}, t^{n-1}, x^n\right),$$
$$B_i \in \mathfrak{A}_i, \qquad a^s \in A, \qquad x^s \in X, \qquad t^s \in \mathbb{R}_+.$$

Definition 2.8 (Local strategies)**.** (1) If at transition times $\sigma^n, n = 0, 1, \dots$, the decision α_i^n at a node i is made according to the probability π_i^n only on the basis of the local history

$$h_i^n = \left(x^0_{\widetilde{N}(i)}, a_i^0, t^0, \dots, x^{n-1}_{\widetilde{N}(i)}, a_i^{n-1}, t^{n-1}, x^n_{\widetilde{N}(i)}\right)$$

of a neighborhood $\widetilde{N}(i)$ of the node i and if $\pi_i^n(F_i^n(x) \mid h_i^n) = \pi_i^n(A_i^n(x^n_{\widetilde{N}(i)}) \mid h_i^n) = 1$ and $x^s_{\widetilde{N}(i)} \in X_{\widetilde{N}(i)}$, $a_i^s \in A_i^s(x^s_{\widetilde{N}(i)})$, $t^s \in \mathbb{R}^+$, then π_i^n is called locally admissible. In this case, the sequence of transition probabilities (decisions) $\delta_i = \{\pi_i^n, n \in \mathbb{N}\}$ is called an admissible local strategy for the vertex i.

(2) An admissible local strategy $\delta = (\delta_i, i \in V)$ is called an admissible local Markov strategy if

$$\pi_i^n \left(\cdot \mid x^0_{\widetilde{N}(i)}, a_i^0, t^0, \dots, x^{n-1}_{\widetilde{N}(i)}, a_i^{n-1}, t^{n-1}, x^n_{\widetilde{N}(i)}\right) = \pi_i^n \left(\cdot \mid x^n_{\widetilde{N}(i)}\right), \qquad i \in V.$$

(3) An admissible local Markov strategy $\delta = (\delta_i, i \in V)$ is called an admissible local stationary (Markov) strategy if

$$\pi_i^{n'} \left(\cdot \mid x_{\widetilde{N}(i)}\right) = \pi_i^{n''} \left(\cdot \mid x_{\widetilde{N}(i)}\right), \qquad i \in V,$$

for all n', n'', and x.

(4) An admissible local stationary (Markov) strategy $\delta = (\delta_i, i \in V)$ is called an admissible local stationary deterministic (nonrandomized) strategy if $\pi_i(\cdot \mid x_{\widetilde{N}(i)})$, $i \in V$, are measures concentrated at a single point in $A_i^n(x_{\widetilde{N}(i)})$, $i \in V$, for all $x \in X$.

(5) We always assume that for local strategies the restriction sets are time invariant:

$$A^n(x) = A(x) = \prod_{i\in V} A_i \left(x_{\widetilde{N}(i)}\right) \qquad \text{for all } n \in \mathbb{N}.$$

The class of all admissible local strategies (with time invariant restriction sets) is denoted by LS; the subclass of admissible local Markov strategies is denoted by LS_M.

By LS_S we denote the class of admissible local stationary strategies with time invariant restriction sets: $\pi^n = \pi^{n'}$ for all n, $n' \in \mathbb{N}$.

By LS_D we denote the class of admissible local stationary deterministic strategies with time invariant restriction sets.

We now incorporate into our semi-Markov process framework the synchronization, locality of the transition kernel (similar to (2) and (3)), and the decision rules. Then we turn to the decisions depending on the transition mechanism and the decisions depending on the sojourn time behavior of the system.

The law of motion of the system is characterized by a set of time invariant transition probabilities. Whenever the state of the system is $\xi^n = x^n$ and a decision $\alpha^n = a^n$ is made, the transition probability is

$$\mathsf{P}\{\xi^{n+1} \in C \mid \xi^n = x^n, \alpha^n = a^n\} =: Q(C \mid x^n, a^n),$$

which is assumed to be independent of the past given the present generalized state that includes the present actions made. Then, given $\{\xi^n = x^n, \alpha^n = a^n\}$, the next state $\xi^{n+1} = x^{n+1}$ of the system is sampled according to $Q(\cdot \mid x^n, a^n)$ and thereafter the sojourn time in x^n, given x^n, a^n, x^{n+1}, is sampled according to some distribution function $\mathrm{T}(\cdot \mid x^n, a^n, x^{n+1})$ which is Borel measurable on $\Delta \times X$.

We further assume that this transition probability and the sojourn time distributions are independent of t, that is, the motion is homogeneous in time.

Applying a control policy δ to a semi-Markov process with interacting components η (semi-Markov random field), respectively to the associated Markov renewal process (ξ, τ) as defined in Definition 2.5, we call the pair (η, δ), respectively the triple (ξ, τ, δ), a controlled version of η, respectively (ξ, τ), using strategy δ.

It should be noted that in general even the embedded jump chain is not Markovian for such a controlled process because the sequence (α^n) of decisions according to π_i^n depends not only on states $x^n_{\widetilde{N}(i)}$, $i \in V$, but also on the previous (local) states $x^0_{\widetilde{N}(i)}, \dots, x^{n-1}_{\widetilde{N}(i)}$. An immediate consequence of the definition below is that if a Markov strategy δ is applied as control strategy for the semi-Markov random field, then we obtain an embedded Markov jump chain ξ from (η, δ), respectively from (ξ, τ, δ).

Definition 2.9. A pair (η, δ), respectively a triple (ξ, τ, δ), is called a controlled stochastic jump process with locally interacting synchronous components with respect to the finite interaction $\Gamma = (V, B)$ if

- $\xi = (\xi^n \colon n \in \mathbb{N})$ is a stochastic process with the state space $X = \prod_{i \in V} X_i$;
- $\tau = (\tau^n \colon n \in \mathbb{N})$ is a stochastic process with the state space $\mathbb{R}_+$;
- $\eta = (\eta^t \colon t \geq 0)$ is a stochastic process with the state space $X = \prod_{i \in V} X_i$;

and these processes are related pathwise as follows:

- the $\mathbb{R}_+$-valued random variables τ^n are the interjump times, and

  $$\{\xi^n, n = 0, 1, \dots\}$$

  is the sequence of states of the process entered just after the jump moments;
- $\sigma = \{\sigma^n \colon n = 0, 1, \dots\}$ with $\sigma^0 = 0$, and

  $$\sigma^n = \sum_{i=0}^{n-1} \tau^i, \qquad n > 0,$$

  is an increasing sequence of jump times. Then $\eta^t = \xi^n$, $n \in \mathbb{N}$, if $t \in [\sigma^n; \sigma^{n+1})$.

We assume that the conditional distribution function of τ^n given by

$$\left\{\xi^{n+1} = x^{n+1}, \xi^n = x^n, \alpha^n = a^n\right\}$$

is $\mathrm{T}(\cdot \mid x^n, a^n, x^{n+1})$, and it is defined as a Borel measurable function $\Delta \times X$. Let $\delta = (\delta_i \colon i \in V)$ be an admissible local strategy and the transitions of ξ be defined as follows:

$$
\begin{aligned}
&\mathsf{P}\left\{\xi_K^{n+1} \in C_K \mid \xi^0 = x^0, \alpha^0 = a^0, \ldots, \xi^{n-1} = x^{n-1}, \alpha^{n-1} = a^{n-1}, \xi^n = y, \alpha^n = a\right\} \\
&\quad \stackrel{(1)}{=} \mathsf{P}\left\{\xi_K^{n+1} \in C_K \mid \xi^0 = x^0, \ldots, \xi^{n-1} = x^{n-1}, \xi^n = y, \alpha^n = a\right\} \\
&\quad \stackrel{(2)}{=} \mathsf{P}\left\{\xi_K^{n+1} \in C_K \mid \xi^n = y, \alpha^n = a\right\} \\
&\quad \stackrel{(3)}{=} \prod_{j \in K} \mathsf{P}\left\{\xi_j^{n+1} \in C_j \mid \xi^n = y, \alpha^n = a\right\} \\
&\quad \stackrel{(4)}{=} \prod_{j \in K} \mathsf{P}\left\{\xi_j^{n+1} \in C_j \mid \xi^n_{\widetilde{N}(j)} = y_{\widetilde{N}(j)}, \alpha_j^n = a_j\right\} \\
&\quad \stackrel{(5)}{=} \prod_{j \in K} Q_j\{C_j \mid y_{\widetilde{N}(j)}, a_j\} \\
&\quad \stackrel{(6)}{=} Q_K(C_K \mid y, a), \qquad K \subseteq V,\ y \in X,\ a_j \in A_j\big(y_{\widetilde{N}(j)}\big).
\end{aligned}
\tag{4}
$$

If $K = V$, then we write $Q_V(C_V \mid y, a) = Q(C \mid y, a)$.

The Markov kernel $Q = \prod_{j \in V} Q_j$ is called local and synchronous.

This construction results in a semi-Markov kernel

$$\widetilde{Q}(C, s \mid x, a) = \mathsf{P}\left\{\xi^{n+1} \in C, \tau^n \leq s \mid \xi^n = x^n, \alpha^n = a^n\right\} =: Q(C \mid x^n, a^n).$$

The process η, respectively (ξ, τ), is called shortly a controlled time dependent semi-Markov random field, respectively a controlled time dependent Markov renewal field.

Some remarks on the modeling principle behind this definition are in order:

- if the first expression is well defined, then $\stackrel{(1)}{=}$ defines the memoryless (Markov) properties with respect to the decision process given the state process history;
- $\stackrel{(2)}{=}$ is an assumption similar to a Markovian transition law;
- $\stackrel{(3)}{=}$ expresses that the coordinates act synchronously;
- $\stackrel{(4)}{=}$ is due to the locality of the state transition law;
- the independence of time for the one-step transition probabilities then leads to the form $\stackrel{(5)}{=}$ of the Markov kernel, and $\stackrel{(6)}{=}$ again follows from the property of coordinates acting synchronously.

2.3. **A criteria for the optimality.** Consider a time dependent semi-Markov random field η. If the kth state of the system is x^k, the joint decision a^k is made, and the duration of the sojourn time in x^k is t^k, then a random reward $r(t^k, x^k, a^k)$ is earned. The function $r(s, x, a)$ is assumed to be Borel measurable on $[0; +\infty) \times \Delta$.

We evaluate the quality, respectively the optimality, of the applied strategies with respect to the following long time average reward measure:

$$\phi(x, \delta) = \liminf_{n \to \infty} \frac{\mathsf{E}_x^\delta \sum_{k=0}^n r\left(\tau^k, \xi^k, \alpha^k\right)}{\mathsf{E}_x^\delta \sum_{k=0}^n \tau^k} \tag{5}$$

where E_x^δ is the expectation associated with the controlled process (η, δ), respectively with (ξ, τ, δ), if $\xi^0 = x$.

The goal of our investigation is to find conditions for the existence of optimal Markovian policies.

Definition 2.10. A strategy $\delta^\star \in LS$ is called optimal in the class LS of local strategies if

$$\phi(x, \delta^\star) = \sup_{\delta \in LS} \phi(x, \delta) \quad \text{for all } x \in X.$$

Such a strategy $\delta^\star$ is shortly called locally optimal.

The quality criterion (5) is introduced by Ross [15] and used in [4] as well. In [15] only stationary policies are studied such that the decisions are made on the basis of the present state only. In this case it can be argued that the value of the measure (5) depends only on the transition kernels $Q(\cdot \mid x, a)$, on $\pi(\cdot \mid x)$, and on conditional expectations

$$\tau(x,a) = \int_X \int_0^\infty t\, d\mathrm{T}(t \mid x,a,y)\, Q(dy \mid x,a), r(x,a),$$
$$r(x,a) = \int_X \int_0^\infty r(t \mid x,a)\, d\mathrm{T}(t \mid x,a,y)\, Q(dy \mid x,a).$$

Then one can restrict the proofs to a special form of sojourn time distributions, say to one-point distributions:

$$\mathrm{T}(t \mid x,a,y) = \begin{cases} 1, & t \geq \tau(x,a); \\ 0, & t < \tau(x,a), \end{cases}$$
$$r(t,x,a) = \begin{cases} 0, & t < \tau(x,a); \\ r(x,a), & t \geq \tau(x,a). \end{cases}$$

Since we consider general policies, such a property does not hold for the asymptotic averaged reward in general.

3. The existence of optimal strategies in LS

In what follows we need some auxiliary assumptions which we collect here:

Assumption I. $\tau(x,a) \geq m > 0$ for $(x,a) \in \Delta$.

Assumption II. $\tau(x,a) \leq M < \infty$ for $(x,a) \in \Delta$.

Assumption III. There exists a nonnegative measure μ on $(X, \mathfrak{X})$ such that

a) $\mu(C) \leq Q(C \mid x,a)$ for $(x,a) \in \Delta$ and $C \in \mathfrak{X}$;
b) $\mu(X) > 0$.

Denote by $\varkappa^n$ the random history of the system up to the time σ^n. Then $\varkappa^n$ assumes values of the form $h^n = (x^0, a^0, t^0, \ldots, x^{n-1}, a^{n-1}, t^{n-1}, x^n)$. Recall that the restriction sets for the local strategies are time independent.

Now we define the conditional expectations

$$\begin{aligned} \tau\left(x^k, a^k\right) &= \mathsf{E}_x^\delta\left\{\tau^k \mid \varkappa^k = h^k, \alpha^k = a^k\right\} \\ &= \int_X \int_0^\infty t\, d\mathrm{T}\left(t \mid x^k, a^k, y\right)\, Q\left(dy \mid x^k, a^k\right) \end{aligned}$$

and

$$\begin{aligned} r\left(x^k, a^k\right) &= \mathsf{E}_x^\delta\left\{r\left(\tau^k, \xi^k, \alpha^k\right) \mid \varkappa^k = h^k, \alpha^k = a^k\right\} \\ &= \int_X \int_0^\infty r\left(t, x^k, a^k\right)\, d\mathrm{T}\left(t \mid x^k, a^k, y\right)\, Q\left(dy \mid x^k, a^k\right). \end{aligned}$$

Using this notation we get

$$\mathsf{E}_x^\delta \sum_{k=0}^n r\left(\tau^k, \xi^k, \alpha^k\right) = \mathsf{E}_x^\delta \sum_{k=0}^n r\left(\xi^k, \alpha^k\right) \tag{6}$$

and

$$\mathsf{E}_x^\delta \sum_{k=0}^n \tau^k = \mathsf{E}_x^\delta \sum_{k=0}^n \tau^k \left(\xi^k, \alpha^k\right).$$

Theorem 3.1. *Let Assumption* I *hold. If there are a constant q and a bounded function $v(x)$ on X such that*

$$v(x) = \sup_{a \in A(x)} \left\{ r(x,a) + \int v(y) Q(dy \mid x,a) - q\tau(x,a) \right\}, \qquad x \in X, \tag{7}$$

then

$$\sup_{\delta \in LS} \phi(x,\delta) \le q, \qquad x \in X. \tag{8}$$

Further if

$$v(x) = \max_{a \in A(x)} \left\{ r(x,a) + \int v(y) Q(dy \mid x,a) - q\tau(x,a) \right\}, \qquad x \in X,$$

and

$$v(x) = r(x, \delta^\star(x)) + \int v(y)\, Q(dy \mid x, \delta^\star(x)) - q\tau(x, \delta^\star(x)), \qquad x \in X, \tag{9}$$

for some strategy $\delta^\star \in LS_D$, then the strategy $\delta^\star$ is locally optimal and moreover

$$\phi(x, \delta^\star) \equiv q.$$

Proof. We have

$$\mathsf{E}_x^\delta \left\{ \sum_{k=1}^n \left[v\left(\xi^k\right) - v\left(\xi^k\right) - \mathsf{E}^\delta \left\{ v\left(\xi^k\right) \mid \varkappa^{k-1}, \alpha^{k-1} \right\} \right] \right\} = 0 \tag{10}$$

for every strategy δ. It follows from (4) and (7) that

$$\begin{aligned} \mathsf{E}^\delta \left\{ v\left(\xi^k\right) \mid \varkappa^{k-1} = h^{k-1}, \alpha^{k-1} = a^{k-1} \right\} &= \int v(y)\, Q\left(dy \mid x^{k-1}, a^{k-1}\right) \\ &= r\left(x^{k-1}, a^{k-1}\right) + \int v(y)\, Q\left(dy \mid x^{k-1}, a^{k-1}\right) - q\tau\left(x^{k-1}, a^{k-1}\right) \\ &\quad - r\left(x^{k-1}, a^{k-1}\right) + q\tau\left(x^{k-1}, a^{k-1}\right) \\ &\le v\left(x^{k-1}\right) - r\left(x^{k-1}, a^{k-1}\right) + q\tau\left(x^{k-1}, a^{k-1}\right). \end{aligned} \tag{11}$$

Substituting (11) into (10), we obtain

$$\mathsf{E}_x^\delta \left\{ \sum_{k=1}^n r\left(\xi^{k-1}, \alpha^{k-1}\right) + v(\xi^n) - v\left(x^0\right) \right\} \le q \mathsf{E}_x^\delta \sum_{k=1}^n \tau\left(\xi^{k-1}, \alpha^{k-1}\right). \tag{12}$$

It follows from Assumption I and (6) that

$$\phi(x,\delta) \le q, \qquad x \in X, \tag{13}$$

since $v(x)$ is bounded. This proves inequality (8).

To prove the second part of the theorem we note that relations (11)–(13) become equalities for the strategy $\delta^\star$. □

Remark 3.1. The proof of Theorem 3.1 is similar to that of Theorem 2 in [15] where the case of stationary strategies is considered. Moreover our Assumption I is weaker than the condition used in [15].

Let $M(X)$ be a Banach space of bounded and Borel measurable functions on X equipped with the sup-norm $\|u\| = \sup_{x\in X} |u(x)|$. Denote by ρ the metric on $M(X)$ induced by this norm. Define the operator $\mathfrak{U}$ on $M(X)$ as follows:

$$\mathfrak{U}u(x) = \sup_{a\in A(x)} \left\{ r(x,a) + \int u(y)\, Q'(dy \mid x,a) \right\} \tag{14}$$

where

$$Q'(C \mid x,a) = Q(C \mid x,a) - \frac{1}{M}\mu(C)\tau(x,a). \tag{15}$$

Theorem 3.2. *Let Assumptions* I–III *hold. Moreover assume that*

(1) *the operator $\mathfrak{U}$ maps some metric subspace $S(X) \subseteq M(X)$ (with the metric ρ induced by the norm in $M(X)$) into itself;*
(2) *whatever function $u \in S(X)$, the mapping $F_u : X \to (2)^A_{\mathrm{set}}$ that defines, for every $x \in X$, the set*

$$F_u(x) := \left\{ a\colon a \in A(x),\ \mathfrak{U}u(x) = r(x,a) + \int u(y)\, Q'(dy \mid x,a) \right\}$$

is Borel measurable.

Then there exists a strategy in LS_D that is optimal in the class LS (a locally optimal strategy).

Proof. We show that the operator $\mathfrak{U}$ possesses the following properties:

a) $\mathfrak{U}$ is isotone, that is, $\mathfrak{U}u_1(x) \geq \mathfrak{U}u_2(x)$ for all $x \in X$ if $u_1(x) \geq u_2(x)$ for all $x \in X$;
b) for any nonnegative constant c,

$$\mathfrak{U}(u(x)+c) \leq \mathfrak{U}u(x) + \alpha c, \qquad x \in X,$$

where $\alpha = 1 - (m/M)\mu(X)$.

Indeed, property a) follows explicitly from Assumption III a), while property b) follows from the following argument:

$$\begin{aligned}
\mathfrak{U}(u(x)+c) &= \sup_{a\in A_x} \left\{ r(x,a) + \int u(y) Q'(dy \mid x,a) + c\int Q'(dy \mid x,a) \right\} \\
&= \sup_{a\in A_x} \left\{ r(x,a) + \int u(y)\, Q'(dy \mid x,a) + c\left(1 - \frac{1}{M}\tau(x,a)\mu(X)\right) \right\} \\
&\leq \sup_{a\in A_x} \left\{ r(x,a) + \int u(y)\, Q'(dy \mid x,a) + c\left(1 - \frac{m}{M}\mu(X)\right) \right\} \\
&= \mathfrak{U}u(x) + \alpha c.
\end{aligned}$$

Properties a) and b) guarantee that $\mathfrak{U}$ is a contraction operator with contraction coefficient $\alpha = 1-(m/M)\mu(X) \in [0;1)$, since the measure μ is substochastic and $m \leq M$. Indeed, apply the operator $\mathfrak{U}$ to both sides of the inequality $u_1(x) \leq u_2(x) + \rho(u_1,u_2)$. Then we obtain from a) and b) that

$$\mathfrak{U}u_1(x) \leq \mathfrak{U}\left[u_2(x) + \rho(u_1,u_2)\right] \leq \mathfrak{U}u_2(x) + \alpha\rho(u_1,u_2).$$

Thus

$$\mathfrak{U}u_1(x) - \mathfrak{U}u_2(x) \leq \alpha\rho(u_1,u_2).$$

Interchanging u_1 and u_2 we obtain

$$|\mathfrak{U}u_1(x) - \mathfrak{U}u_2(x)| \leq \alpha\rho(u_1,u_2), \qquad x \in X,$$

that is

$$\rho(\mathfrak{U}u_1, \mathfrak{U}u_2) \leq \alpha\rho(u_1, u_2).$$

Then the Banach fixed point theorem implies that there exists a function $u^\star(x)$ in $S(X)$ such that

$$u^\star(x) = \sup_{a \in A(x)} \left\{ r(x,a) + \int u^\star(y) Q'(dy \mid x, a) \right\}, \qquad x \in X. \tag{16}$$

Now we derive from (16) and (15) that the optimality condition (7) holds with

$$v(x) = u^\star(x) \quad \text{and} \quad q = M^{-1} \int u^\star(y)\, \mu(dy).$$

Assumption (2) of Theorem 3.2 yields that the function $F_{u^\star}$ is Borel measurable. According to Definition 2.3 we obtain that $F_{u^\star}(x) \neq \varnothing$ for all $x \in X$. Thus $F_{u^\star}(x) \subseteq A(x)$ by definition. Therefore the set $F_{u^\star}(x)$ of maximizers $\mathfrak{U}u^\star(x)$ contains only local policies.

Now, by the Theorem of Choice, there exists a Borel measurable $A(x)$-valued selector $\delta^\star$, that is,

$$\delta^\star(x) \in \left\{ a \in A(x) \colon \mathfrak{U}u^\star(x) = r(x,a) + \int u^\star(y)\, Q'(dy \mid x, a) \right\},$$

which means that

$$\mathfrak{U}u^\star(x) = r(x, \delta^\star(x)) + \int u^\star(y)\, Q'(dy \mid x, \delta^\star(x)),$$

whence (9) follows for the strategy $\delta^\star \in LS_D$, and the strategy $\delta^\star$ is locally optimal by Theorem 3.1. □

Remark 3.2. Under the assumptions of Theorem 3.2 we conclude that in the set of optimal strategies there is at least one deterministic strategy which makes decisions on the basis of the actual state only. The optimal value $\phi(x, \delta^\star)$ under this strategy depends, according to the observation of Ross [15], only on the transition kernels $Q(\cdot\,|x,a)$ and $\pi^n(\cdot, x)$ and mean conditional sojourn times $\tau(x,a)$ and mean conditional rewards $r(x,a)$. Thus $\phi(x, \delta^\star)$ is insensitive to variations of the shape of the conditional sojourn times distribution and the conditional rewards distribution as long as their conditional means remain invariant. Therefore, without loss of generality, we can restrict the computation of $\phi(x, \delta^\star)$ to processes with deterministic sojourn times and special reward functions as given at the end of Section 2.3.

Our next theorem provides an easy tool to check smoothness conditions for semi-Markov models such that conditions 1) and 2) of Theorem 3.2 are fulfilled.

Theorem 3.3. *Consider a semi-Markov random field in the sense of Definition* 2.9. *Recall that the local state spaces X_i, local decision states A_i, and therefore the global state space X and global decision space A are compact metric spaces with a countable basis. Let the mapping F, which associates with any state x the time invariant restriction set $A(x)$, that is,*

$$F \colon X \longrightarrow (2)^A_{\text{set}}, \qquad x \to F(x) = \prod_{i \in V} F_i(x) = A(x),$$

be continuous and let Assumptions I, II, *and* III *be satisfied. Further we assume that*

(1) *the mean value functions $r(x,a)$ and $\tau(x,a)$ are continuous in $(x,a) \in \Delta$;*
(2) *the transition probability $Q(\cdot \mid x, a)$ is weakly continuous in $(x,a) \in \Delta$.*

Then in the class LS_D there exists an optimal strategy $\delta^\star$, and the function $\delta^\star \colon X \to A$ can be chosen in the Baire class 1.

Proof. We use some ideas of the proof of Theorem 3 in [5] in order to apply our Theorem 3.2. To do this we need the following property which combines Lemmas 1 and 2 in [5]:

> Let A be a compact set and $F\colon X \to (2)^A_{\text{set}}$ a continuous mapping. Moreover let the function $u\colon \Delta \subseteq X \times A \to \mathbb{R}$ be bounded and continuous. Then the function $u^\star(x) = \sup_{a\in F(x)=A(x)} u(x,a)$ is continuous.

Applying (1), (2), and the latter property to the maximizing operator $\mathfrak{U}$ we obtain from (14) that $\mathfrak{U}$ maps the space $B(X) \subseteq M(X)$ of real-valued functions on X into itself. Therefore condition (1) of Theorem 3.2 holds. To prove condition (2) of Theorem 3.2 we show that the function $F_u\colon X \to (2)^A_{\text{set}}$, $x \to F_u(x) = A(x)$ is upper semicontinuous for all $u \in C(X)$. According to [12, p. 61], this result follows if the following holds for all fixed $u \in C(X)$:

> If $(x^n, a^n) \to (x,a)$ as $n \to \infty$, then $a \in F_u(x)$, since $x^n \in X$ and $a^n \in F_u(x^n)$, $n \in \mathbb{N}$.

It follows from the definition of $a^n \in F_u(x^n)$ that

$$r\left(x^n, a^n\right) + \int u(y)\, Q'\left(dy \mid x^n, a^n\right) = \sup_{\widetilde{a}\in A(x^n)} \left\{ r(x^n, \widetilde{a}) + \int u(y)\, Q'(dy \mid x^n, \widetilde{a}) \right\}.$$

Then the left-hand side of this equation is a continuous function of (x,a). At the same time the right-hand side is a continuous function of x. Letting $n \to \infty$ we obtain

$$r(x,a) + \int u(y)\, Q'(dy \mid x, a) = \sup_{\widetilde{a}\in A(x)} \left\{ r(x, \widetilde{a}) + \int u(y)\, Q'(dy \mid x, \widetilde{a}) \right\},$$

whence $a \in F_u(x)$.

In view of Theorem 3.2 we conclude that there exists an optimal local stationary deterministic strategy $\delta^\star \in LS_D$. □

Bibliography

1. E. Çinlar, *Introduction to Stochastic Processes*, Prentice-Hall, Englewood Cliffs, New Jersey, 1975. MR0380912 (52:1809)
2. R. R. Disney and P. C. Kiessler, *Traffic Processes in Queueing Networks. A Markov Renewal Approach*, The Johns Hopkins University Press, London, 1987. MR0896033 (89b:90081)
3. H. Daduna, P. S. Knopov, and R. K. Chorney, *Local control of Markovian processes of interaction on a graph with a compact set of states*, Kibernetika Sistem. Analiz **2001**, no. 3, 62–77; English transl. in Cybernet. Systems. Anal. **37** (2001), no. 3, 348–361. MR1872136 (2002h:90111)
4. L. G. Gubenko and E. S. Statland, *Controlled semi-Markov processes*, Kibernetika **1972**, no. 2 26–29; English transl. in Cybernetics **8** (1972), 200–205. MR0323423 (48:1779)
5. L. G. Gubenko and E. S. Statland, *Controlled semi-Markov processes*, Teor. Veroyatnost. Matem. Statist. **7** (1972), 51–64; English transl. in Theory Probab. Mathem. Statist. **7** (1972), 47–61. MR0334957 (48:13275)
6. K. Hinderer, *Foundations of Non-Stationary Dynamic Programming with Discrete Time Parameter*, Lecture Notes in Operations Research and Mathematical Systems, vol. 33, Springer-Verlag, Berlin, 1970. MR0267890 (42:2791)
7. R. A. Howard, *Research in semi-Markov decision structures*, Journal of the Operational Research Society of Japan **6** (1964), 163–199. MR0172699 (30:2918)
8. W. S. Jewell, *Markov renewal programming. I. Formulation, finite return models*, Operations Research **11** (1963), 938–948. MR0163374 (29:677)
9. W. S. Jewell, *Markov renewal programming. II. Infinite return models, example*, Operations Research **11** (1963), 948–971. MR0163375 (29:678)
10. M. Kitaev, *Elimination of randomization in semi-Markov decision models with average cost criterion*, Optimization **18** (1987), no. 3, 439–446. MR0882524 (88f:90177)
11. M. Y. Kitaev and V. V. Rykov, *Controlled Queueing Systems*, CRC Press, Boca Raton, 1995. MR1413045 (97h:90001)
12. K. Kuratowski, *Topology*, vol. 2, Academic Press, New York, 1968. MR0259835 (41:4467)

13. Vasquez F. Luque and Alcaraz M. T. Robles, *Controlled semi-Markov models with discounted unbounded costs*, Bol. Soc. Math. Mex. **39** (1994), no. 1–2, 51–68. MR1338681 (96d:93089)
14. R. T. Rockafellar, *Measurable dependence of convex sets and functions on parameters*, J. Math. Anal. Appl. **28** (1969), 4–25. MR0247019 (40:288)
15. S. M. Ross, *Average cost semi-Markov processes*, Journal of Applied Probability **7** (1970), no. 3, 649–656. MR0303635 (46:2771)
16. N. B. Vasilyev, *Bernoulli and Markov stationary measures in discrete local interactions*, Locally Interacting Systems and Their Application in Biology, (R. L. Dobrushin, V. I. Kryukov, and A. L. Toom, eds.), Lecture Notes in Mathematics, vol. 653, Springer-Verlag, Berlin, 1978, pp. 99–112. MR0505437 (80a:60133)
17. O. Vega-Amaya, *Average optimality in semi-Markov control models on Borel spaces: unbounded cost and controls*, Bol. Soc. Math. Mex. **38** (1993), no. 1–2, 47–60. MR1313106 (95j:90090)
18. K. Wakuta, *Arbitrary state semi-Markov decision processes with unbounded rewards*, Optimization **18** (1987), no. 3, 447–454. MR0882525 (88f:90179)

Universität Hamburg, Fakultät Mathematik, Bundesstrasse 55, D–20146 Hamburg, Germany

E-mail address: amos801@math.uni-Hamburg.de

Glushkov Institute for Cybernetics, National Academy of Sciences of Ukraine, Academician Glushkov Avenue 40, Kiev–187, 03680, Ukraine

E-mail address: knopov1@yahoo.com

Inter-Regional Academy of Personnel Management, Department of Mathematics, Frometivs'ka Street 2, Kiev–39, 03039, Ukraine

Received 24/JAN/2003

Translated by THE AUTHORS