Available online at www.sciencedirect.com
ScienceDirect
Applied Mathematics and Computation 198 (2008) 198–208
APPLIED MATHEMATICS AND COMPUTATION
www.elsevier.com/locate/amc

Stochastic shortest path problems with associative accumulative criteria

Yoshio Ohtsubo

Department of Mathematics, Faculty of Science, Kochi University, Kochi 780-8520, Japan

Abstract

We consider a stochastic shortest path problem with associative criteria in which for each node of a graph we choose a probability distribution over the set of successor nodes so as to reach a given target node optimally. We formulate such a problem as an associative Markov decision processes. We show that an optimal value function is a unique solution to an optimality equation and find an optimal stationary policy. Also we give a value iteration method and a policy improvement method.

Keywords: Shortest path problem; Markov decision process; Associative criterion; Invariant imbedding method; Optimality equation; Existence of optimal policy

1. Introduction

For a directed graph with nodes $1, 2, \ldots, K$ and with a cost (length or time) assigned to each arc, a stochastic shortest path problem is to select a probability distribution over all possible successor nodes at each node $i \neq K$ so as to reach a target node K with minimal associative accumulative cost.

Such a stochastic shortest path problem is analyzed by using the general theory of Markov decision processes in many references. Eaton and Zadeh [3] formulated such a problem as a pursuit problem and they showed that the optimal expected total cost is a unique solution to an optimality equation if at least one proper policy exists, and they gave an optimal value by a value iteration method. Derman in [4,5] considered the problem, where a target state (node) is absorbing, and proved that the problem has an optimal stationary policy and he gave several methods for obtaining optimal solutions. In [16], Sancho formulated Markov decision processes to analyze the problem and gave a policy iteration method. Bertsekas and Tsitsiklis [2] investigated the problem without the cost nonnegativity assumption and proved a natural generalization of the standard result for the deterministic shortest path problem within the framework of undiscounted finite state

E-mail address: ohtsubo@math.kochi-u.ac.jp

0096-3003/$ - see front matter
doi:10.1016/j.amc.2007.08.025

Markovian decision processes. In all of these, a criterion function is the expected total cost, which we call an additive case.

Also, Ohtsubo [12] considered a minimizing risk models in stochastic shortest path problems as undiscounted finite Markov processes and showed that an optimal value function is a unique solution to an optimality equation and found an optimal stationary policy by using an invariant imbedding method. General minimizing risk models in discounted Markov decision processes were investigated in White [17,18], Wu and Lin [19], Ohtsubo and Toyonaga [11,13] and Ohtsubo [14].

On the other hand, Maruyama in [9,10] investigated deterministic shortest path problems with associative criteria and show the existence and uniqueness of the optimal value. Especially in [10] he obtained a parameterized recursive equation for the class of the problem by using an invariant imbedding technique.

Furthermore the optimization problem for minimum criteria, which is associative, was first introduced by Bellman and Zadeh [1] as decision-making in fuzzy environment, and Iwamoto et al. [6–8] and Ohtsubo [15] formulated their optimization problem as finite horizon Markov decision processes and gave a optimal policy by using an invariant imbedding approach.

In this paper, we concern ourselves with a stochastic shortest path problem with an associative criterion, which is an expected accumulate cost $E_i^\pi[\bigcirc_{n=0}^{\tau} Y_n] = E_i^\pi[Y_0 \circ Y_1 \circ Y_2 \circ \cdots Y_\tau]$, where Y_n is a cost at nth step, $\circ$ is an operator with an associative property satisfying some conditions, τ is a hitting time to the target node K and E_i^π is an expectation operator when the starting node is i and a policy π is used. In Section 2, we give notations and formulate our model as undiscounted finite Markov decision processes with infinite horizon. In Section 3, we prove that the optimal value function is a unique solution to an optimality equation by using an invariant imbedding approach and that it is given by a value iteration method. We also show that there exists an optimal left continuous stationary policy. In Section 4, we give a policy improvement method for obtaining a optimal policy.

2. Notations and formulation

In this section, we formulate associative models in stochastic shortest path problems as Markov decision processes $\Gamma = ((X_n), (A_n), (Y_n), p)$ with a discrete time space $N = \{0, 1, 2, \ldots\}$. The state space S is a finite set $\{1, 2, \ldots, K\}$, where K is a target state, and we denote the state at time $n \in N$ by X_n. The action space A is finite and we denote the action at time $n \in N$ by A_n. The cost space E is a finite set $\{y_1, y_2, \ldots, y_\ell\}$, where $E \subset B$ for some subset B of R, and $Y_n \in E$ is a random cost function at time $n \in N$ with $Y_0 = e$, where e is a unit element defined below. We define conditional probability distributions by

$$q^a(j|i) = P(X_{n+1} = j | X_n = i, A_n = a),$$
$$\hat{q}^a_{ij}(y) = P(Y_{n+1} = y | X_n = i, X_{n+1} = j, A_n = a)$$

and set

$$p^a(j, y|i) = q^a(j|i)\hat{q}^a_{ij}(y) = P(X_{n+1} = j, Y_{n+1} = y | X_n = i, A_n = a)$$

for $i,j \in S$, $a \in A$ and $y \in E$. We use $S_B = S \times B$ as a new state space.

For a binary operator $\circ : R \times R \to R$ and a subset B of R, we assume that

(i) B is closed for the operator $\circ$, that is, $x \circ y \in B$ for any $x, y \in B$,
(ii) the operator $\circ$ is associative, that is, $(x \circ y) \circ z = x \circ (y \circ z)$ $(= x \circ y \circ z$, say) for any $x, y, z \in B$,
(iii) B has a unit element e, that is, $e \in B$ and $x \circ e = e \circ x = x$ for any $x \in B$,
(iv) $(B, \circ)$ is nondecreasing in the sense that $x \leqslant x \circ y$ and $x \leqslant y \circ x$ for any $x, y \in B$.

On the condition (iv), letting $x = e$, we notice that $y \geqslant e$ for any $y \in B$. Also we easily see under the conditions (i), (ii) and (iii) that if $x \geqslant e$ for any $x \in B$ and if $x \circ y \leqslant x \circ z$ and $y \circ x \leqslant z \circ x$ for any $x, y, z \in B$ satisfying $y \leqslant z$, then the condition (iv) holds. In algebra, $(B, \circ)$ satisfying the conditions (i), (ii) and (iii) is called a semigroup and it is also analogous to t-conorm (or s-norm) in fuzzy set theory (cf. [20]).

We give several examples in which $(B, \circ)$ satisfies the above conditions (cf. [9,20]).