

ON CONTROLLED MARKOV CHAINS WITH OPTIMALITY REQUIREMENT AND SAFETY CONSTRAINT

SHUN-PIN HSU¹, ARISTOTLE ARAPOSTATHIS² AND RATNESH KUMAR³

¹Department of Electrical Engineering
National Chung-Hsing University
Taichung, 402, Taiwan
shsu@nchu.edu.tw

²Department of Electrical and Computer Engineering
The University of Texas at Austin
Austin, TX 78712-0240, USA
ari@ece.utexas.edu

³Department of Electrical and Computer Engineering
Iowa State University
Ames, IA 50011-3060, USA
rkumar@ece.iowa.edu

Received December 2008; revised June 2009

ABSTRACT. We study the control of completely observed Markov chains subject to generalized safety bounds and optimality requirement. Originally, the safety bounds were specified as unit-interval valued vector pairs (lower and upper bounds for each component of the state probability distribution). In this paper, we generalize the constraint to be any linear convex set for the distribution to stay in, and present a way to compute a stationary control policy which is safe and at the same time long-run average optimal. This policy guarantees the safety of the system as it is on its 'limiting status', and is derived through a linear programming formulation with its feasibility problem explored. To assure the safety of the system's transient behavior under the policy assumed to induce a unique limiting distribution in the interior of the constraint set, we present a finitely-terminating iterative algorithm to compute the maximal invariant safe set (MISS) such that starting from which any initial distribution incurs a sequence of future distributions that are safe also. A theoretic upper bound for the number of iterations is provided. Furthermore, a simplified algorithm that might require less calculation is also introduced and illustrated in numerical examples. In particular, we obtain the closed-form representation for the MISS of two-state system based on at most one iteration of the algorithm.

Keywords: Controlled Markov chains, Safety control, Stochastic discrete event system, Reliability

1. **Introduction.** Controlled Markov Chains (CMC) are among the most widely used models for the study of stochastic systems (see e.g. [9]). One of the main objective of this model is to optimize the system performance through implementation of the calculated policy that minimizes the induced long-run average cost. It is well known that, under certain conditions, this *optimal* policy can be characterized by the stochastic dynamic programming equation called *Bellman's ergodic optimality equation*, which can be solved by the *policy iteration* or *value iteration* algorithm (see e.g. [7]). When the CMC model is subject to other constraints (e.g., in cost), the model is called *constrained CMC* or *CMC with constraints*. To solve the associated dynamic programming equation with constraints, a linear programming formulation involving the concept of state-action frequency was