ARTICLE IN PRESS YINCO:4232

Information and Computation ••• (••••) •••–•••

Contents lists available at ScienceDirect

Information and Computation

www.elsevier.com/locate/yinco

Meet your expectations with guarantees: Beyond worst-case synthesis in quantitative games ☆

Véronique Bruyère [a], Emmanuel Filiot [c], Mickael Randour [a,b,c,*], Jean-François Raskin [c]

[a] Computer Science Department, Université de Mons (UMONS), Belgium
[b] LSV – CNRS & ENS Cachan, France
[c] Département d'Informatique, Université Libre de Bruxelles (ULB), Belgium

ARTICLE INFO

Article history:
Received 29 October 2014
Available online xxxx

Keywords:
Two-player quantitative game on graph
Markov decision process
Beyond worst-case synthesis
Worst-case and expected value
Mean-payoff
Shortest path

ABSTRACT

Classical analysis of two-player quantitative games involves an adversary (modeling the environment of the system) which is purely antagonistic and asks for strict guarantees while Markov decision processes model systems facing a purely randomized environment: the aim is then to optimize the expected payoff, with no guarantee on individual outcomes. We introduce the beyond worst-case synthesis problem, which is to construct strategies that guarantee some quantitative requirement in the worst-case while providing a higher expected value against a particular stochastic model of the environment given as input.
We study the beyond worst-case synthesis problem for two important quantitative settings: the mean-payoff and the shortest path. In both cases, we show how to decide the existence of finite-memory strategies satisfying the problem and how to synthesize one if one exists. We establish algorithms and we study complexity bounds and memory requirements.

1. Introduction

Classical models. Two-player zero-sum quantitative games [1–3] and Markov decision processes (MDPs) [4,5] are two popular formalisms for modeling decision making in adversarial and uncertain environments respectively. In the former, two players compete with opposite goals (zero-sum), and we want strategies for player 1 (the system) that ensure a given *minimal performance against all possible strategies* of player 2 (its environment). In the latter, the system plays against a stochastic model of its environment, and we want strategies that ensure a *good expected overall performance*. Those two models are well studied and simple optimal memoryless strategies exist for classical objectives such as mean-payoff [6,1,7] or shortest path [8,9]. But both models have clear weaknesses: strategies that are good for the worst-case may exhibit suboptimal behaviors in probable situations while strategies that are good for the expectation may be terrible in some unlikely but possible situations.

What if we want both? In practice, we would like to have strategies that are both ensuring (*a*) some worst-case threshold no matter how the adversary behaves (i.e., against any arbitrary strategy) and (*b*) a good expectation against the expected

☆ E. Filiot is a F.R.S.-FNRS Research Associate, M. Randour is a F.R.S.-FNRS Postdoctoral Researcher, J.-F. Raskin is supported by ERC Starting Grant (279499: inVEST). Work supported by European project CASSTING (FP7-ICT-601148).

* Corresponding author.
E-mail address: mickael.randour@gmail.com (M. Randour).

http://dx.doi.org/10.1016/j.ic.2016.10.011
0890-5401/

Table 1
Overview of decision problem complexities and (tight) memory requirements for winning strategies of player 1 in games (worst-case), MDPs (expected value) and the BWC setting (combination).

		Worst-case	Expected value	**BWC**
mean-payoff	complexity	NP ∩ coNP	P	**NP ∩ coNP**
	memory	memoryless		**pseudo-poly.**
shortest path	complexity	P		**pseudo-poly. / NP-hard**
	memory	memoryless		**pseudo-poly.**

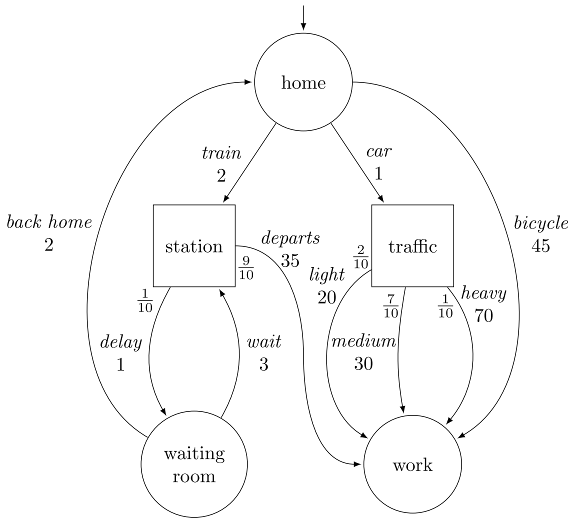

Fig. 1. Player 1 wants to minimize its expected time to reach "work", but while ensuring it is less than one hour in all cases.

behavior of the adversary (given as a stochastic model). This is the subject of this paper: we show how to construct finite-memory strategies that ensure both (*a*) and (*b*). We consider finite-memory strategies for player 1 as they can be implemented in practice (as opposed to infinite-memory ones). Player 2 is not restricted in his choice of strategies, but we will see that simple strategies suffice. Our problem, the **beyond worst-case synthesis problem**, is interesting for any quantitative measure, but we give here a thorough study of two classical ones: the *mean-payoff*, and the *shortest path*. Our results are summarized in Table 1.

Example. Let us consider the weighted game in Fig. 1 to illustrate the *shortest path* context. Circle states belong to player 1, square states to player 2, integer labels are durations in minutes, and fractions are probabilities that model the expected behavior of player 2. Player 1 wants a strategy to go from "home" to "work" such that "work" is *guaranteed* to be reached within 60 minutes (to avoid missing an important meeting), and player 1 would also like to minimize the expected time to reach "work". First, note that the strategy that minimizes the expectation is to take the car (expectation is 33 minutes) but this strategy is excluded as there is a possibility to arrive after 60 minutes (in case of heavy traffic). Bicycle is safe but the expectation of this solution is 45 minutes. We can do better with the following strategy: try to take the train, if the train is delayed three times consecutively, then go back home and take the bicycle. This strategy is safe as it always reaches "work" within 58 minutes and its expectation is $\approx 37,45$ minutes (so better than taking directly the bicycle). Observe that this simple example already shows that, unlike the situation for classical games and MDPs, strategies using memory are strictly more powerful than memoryless ones. Our algorithms are able to decide the existence of (and synthesize) such finite-memory strategies.

Contributions. Our main results are the following. First, for the mean-payoff value, we provide an algorithm (Theorem 3) that implies NP ∩ coNP-membership of the problem, which would reduce to P if mean-payoff games were proved to be in P, a long-standing open problem [3,10]. Pseudo-polynomial memory may be necessary and always suffices (Theorem 37). Finally, we observe that infinite-memory strategies are strictly more powerful that finite-memory strategies (Sect. 4.11). Second, for the shortest path, we provide a pseudo-polynomial time algorithm (Theorem 38), and show that the associated decision problem is NP-hard (Theorem 41). According to a very recent result by Haase and Kiefer [11], our reduction even proves PP-hardness, which suggests that the problem does not belong to NP at all otherwise the polynomial hierarchy would collapse. Pseudo-polynomial memory may be necessary and always suffices (Theorem 39). In the case of the shortest path problem, infinite-memory strategies give no additional power in comparison with finite-memory strategies (Remark 40).

Please cite this article in press as: V. Bruyère et al., Meet your expectations with guarantees: Beyond worst-case synthesis in quantitative games, Inf. Comput. (2016), http://dx.doi.org/10.1016/j.ic.2016.10.011