Mean-Payoff Games with Partial-Observation*
(Extended Abstract)

Paul Hunter, Guillermo A. Pérez, and Jean-François Raskin

Département d’Informatique, Université Libre de Bruxelles (U.L.B.), Belgium
{paul.hunter,guillermo.perez,jraskin}@ulb.ac.be

Abstract. Mean-payoff games are important quantitative models for open reactive systems. They have been widely studied as games of perfect information. In this paper we investigate the algorithmic properties of several subclasses of mean-payoff games where the players have asymmetric information about the state of the game. These games are in general undecidable and not determined according to the classical definition. We show that such games are determined under a more general notion of winning strategy. We also consider mean-payoff games where the winner can be determined by the winner of a finite cycle-forming game. This yields several decidable classes of mean-payoff games of asymmetric information that require only finite-memory strategies, including a generalization of perfect information games where positional strategies are sufficient. We give an exponential time algorithm for determining the winner of the latter.

1 Introduction

Mean-payoff games (MPGs) are two-player, infinite duration, turn-based games played on finite edge-weighted graphs. The two players alternately move a token around the graph; and one of the players (Eve) tries to maximize the (limit) average weight of the edges traversed, whilst the other player (Adam) attempts to minimize the average weight. Such games are particularly useful in the field of verification of models of reactive systems, and the perfect information versions of these games have been extensively studied \[4,7,8,10\]. One of the major open questions in the field of verification is whether the following decision problem, known to be in the intersection of the classes \(\text{NP} \cap \text{coNP}\):\(^1\) can be solved in polynomial time: Given a threshold \(\nu\), does Eve have a strategy to ensure a mean-payoff value of at least \(\nu\)?

In game theory the concepts of imperfect, partial and limited information indicate situations where players have asymmetric knowledge about the state of the game. In the context of verification games this partial knowledge is reflected in one or both players being unable to determine the precise location of the token amongst several equivalent vertices, and such games have also been extensively

\footnote{This work was supported by the ERC inVEST (279499) project.}

\footnote{From results in \[17\] and \[12\] it follows that the problem is also in \(\text{UP} \cap \text{coUP}\).}
studied \[2, 3, 9, 13, 16\]. Adding partial-observation to verification games results in an enormous increase in complexity, both algorithmically and in terms of strategy synthesis. For example, it was shown in [9] that for MPGs with partial-observation, when the mean payoff value is defined using \( \limsup \), the analogue of the above decision problem is undecidable; and whilst memoryless strategies suffice for MPGs with perfect information, infinite memory may be required. The first result of this paper is to show that this is also the case when the mean payoff value is defined using the stronger \( \liminf \) operator, closing two open questions posed in [9]. As a consequence, we generalize a result from [6] which uses the undecidability result from [9] to show several classical problems for mean-payoff automata are also undecidable.

To simplify our definitions and algorithmic results we initially consider a restriction on the set of observations which we term limited-observation. In games of limited-observation the current observation contains only those vertices consistent with the observable history, that is the observations are the belief set of Eve (see, e.g. [5]). This is not too restrictive as any MPG with partial-observation can be realized as a game of limited-observation via a subset construction. In Section 9 we consider the extension of our definitions to MPGs with partial-observation via this construction.

Our focus for the paper will be on games at the observation level, in particular we are interested in observation-based strategies for both players. Whilst observation-based strategies for Eve are usual in the literature, observation-based strategies for Adam have not, to the best of our knowledge, been considered. Such strategies are more advantageous for Adam as they encompass several simultaneous concrete strategies. Further, in games of limited-observation there is guaranteed to be at least one concrete strategy consistent with an observation-based strategy. Our second result is to show that although MPGs with partial-observation are not determined under the usual definition of strategy, they are determined when Adam can use an observation-based strategy.

In games of perfect information one aspect of MPGs that leads to simple (but not quite efficient) decision procedures is their equivalence to finite cycle-forming games. Such games are played as their infinite counterparts, however when the token revisits a vertex the game is stopped. The winner is determined by a finite analogue of the mean-payoff condition on the cycle now formed. Ehrenfeucht and Mycielski [10] and Björklund et al. [12] used this equivalence to show that positional strategies are sufficient to win MPGs with perfect information. Critically, a winning strategy in the finite game translates directly to a winning strategy in the MPG, so such games are especially useful for strategy synthesis.

We extend this idea to games of partial-observation by introducing a finite, perfect information, cycle-forming game played at the observation level. That is, the game finishes when an observation is revisited (though not necessarily the first time). In this reachability game winning strategies can be translated to finite-memory winning strategies in the MPG. This leads to a large, natural subclass of MPGs with partial-observation, forcibly terminating games, where