The Multiple Dimensions of Mean-Payoff Games (Extended Abstract)

Laurent Doyen

LSV, ENS Paris-Saclay & CNRS

Outline We consider quantitative game models for the design of reactive systems working in resource-constrained environment. The game is played on a finite weighted graph where some resource (e.g., battery) can be consumed or recharged along the edges of the graph.

In mean-payoff games, the resource usage is computed as the long-run average resource consumption. In energy games, the resource usage is the initial amount of resource necessary to maintain the resource level always positive.

We review fundamental results about mean-payoff games that show the existence of memoryless optimal strategies, and the equivalence of mean-payoff games with finite-duration reachability games, as well as with energy games (which can also be viewed as safety games). These results provide conceptually simple backward-induction algorithms for solving mean-payoff games, and for constructing memoryless optimal strategies. It follows that mean-payoff games can be solved in NP $\cap$ coNP.

Then we consider games with multiple mean-payoff conditions for systems using multiple resources. In multi-dimension mean-payoff games, memory is necessary for optimal strategies, and the previous equivalence results with reachability and energy (safety) games no longer hold. First, infinite memory is necessary in general for optimal strategies. With infinite memory, the limit of the long-run average resource consumption may not exist, and it is necessary to distinguish between the limsup and the liminf of the long-run average resource consumption. Second, the equivalence with a multi-dimensional version of energy games holds only if the players are restricted to use finite-memory strategies, and in that case the limsup- and the liminf-value coincide.

The complexity of solving multi-dimension mean-payoff games is as follows, depending on which class of strategies is given to the player: NP-complete for memoryless strategies, coNP-complete for finite-memory strategies, NP $\cap$ coNP for infinite-memory strategies and a conjunction of limsup objectives, and coNP-complete for infinite-memory strategies and a conjunction of liminf objectives.

Games We consider two-player games of infinite duration, as a model of non-terminating reactive systems with controllable and uncontrollable actions. Such models have applications in the verification and synthesis of reactive systems [3, 36, 38], and fundamental connections with many areas of computer science, such as logic and automata [24, 27, 37].

The game is played for infinitely many rounds, on a finite graphs $\langle V, E\rangle$ with vertices $V = V_1 \uplus V_2$ partitioned into player-1 and player-2 vertices. Initially a token is placed on a designated vertex $v_0 \in V$. In each round the player owning the vertex v_i where the token lies moves the token to a successor vertex v_{i+1} along an edge $(v_i, v_{i+1}) \in E$ of the graph. The outcome of the game is an infinite path $v_0, v_1, \ldots$ called a *play*. In the traditional qualitative analysis, plays are classified as either winning or losing (for player 1). An *objective* is a set $\Omega \subseteq V^\omega$ of winning plays. The goal of player 1 is to achieve a winning play: the central qualitative question is to decide if there exists a strategy for player 1 such that for all strategies of player 2 the outcome is a winning play in Ω. Sets of winning plays defined by ω-regular conditions are central in verification and synthesis of reactive systems [36, 38] and have been extensively studied [10, 18, 26].

Mean-Payoff and Energy Games For the design of reactive systems working in resource-constrained environment, we consider weighted graphs $\langle V, E, w\rangle$ where $w : E \to \mathbb{Z}$ is a weight function that assigns a resource consumption to each edge of the graph. The limit-average value (or mean-payoff value) of a play $\rho = v_0, v_1, \ldots$ is defined in two variants, the limsup-average

$$\overline{\mathsf{MP}}(\rho) = \limsup_{n\to\infty} \frac{1}{n} \cdot \sum_{i=0}^{n-1} w(v_i, v_{i+1}),$$

and the liminf-average

$$\underline{\mathsf{MP}}(\rho) = \liminf_{n\to\infty} \frac{1}{n} \cdot \sum_{i=0}^{n-1} w(v_i, v_{i+1}).$$

The goal of player 1 is to maximize the mean-payoff value, and the associated decision problem is to decide, given a threshold value $\nu \in \mathbb{Q}$, whether there exists a winning strategy for player 1 for the mean-payoff objective $\overline{\mathsf{MP}}^{\geq\nu} = \{\rho \mid \overline{\mathsf{MP}}(\rho) \geq \nu\}$ (or, $\underline{\mathsf{MP}}^{\geq\nu} = \{\rho \mid \underline{\mathsf{MP}}(\rho) \geq \nu\}$) with a mean-payoff value at least ν. In the sequel we only consider the case $\nu = 0$, which can be obtained by shifting all weights in the graph by the given value ν (and scaling them to get integers).

Memoryless strategies are sufficient to win mean-payoff games [23] and the associated decision problem lies in NP $\cap$ coNP [32, 43]. The solution of mean-payoff games is more intuitive by considering the class of energy games [6, 11] where the objective of player 1 is to maintain the accumulated resource level (i.e., the energy level) always nonnegative, given an initial credit value $c_0 \in \mathbb{N}$. The energy level of a finite path with initial credit c_0 is defined by

$$\mathsf{EL}_{c_0}(v_0, v_1, \ldots, v_n) = c_0 + \sum_{i=0}^{n-1} w(v_i, v_{i+1}),$$

and the energy objective is

$$\mathsf{EL}_{c_0}^{\geq 0} = \{v_0, v_1, \ldots \mid \forall n \in \mathbb{N} : \mathsf{EL}_{c_0}(v_0, v_1, \ldots, v_n) \geq 0\}.$$

The associated decision problem asks whether there exists c_0 such that player 1 has a winning strategy for the energy objective with initial credit c_0. It is important that the initial credit c_0 is not fixed to obtain the equivalence with mean-payoff games (see next paragraph). For fixed initial credit we get a variant of energy games (see Related Work). Note that the energy condition is a safety condition [1]: if a finite prefix of a play violates the energy condition, then every continuation violates the energy condition.

Memoryless Determinacy and Equivalence Results Intuitively, player 1 wins the energy game if he can ensure that whenever a cycle is formed in the play, the cycle has nonnegative sum of weights. The converse is also true [6, 39]. Consider a finite-duration game played analogously to the games we considered so far, but that stops whenever a cycle is formed, and declared won by player 1 if and only if the cycle is nonnegative. If player 1 wins the finite-duration game, then we can use his strategy to play the energy game and ensure that all cycles that are formed are nonnegative and therefore a finite initial credit c_0 is sufficient to survive along acyclic portions of the play (thus $c_0 \leq |V| \cdot W$ is sufficient, where W is the largest absolute weight in the graph). Conversely, if player 1 cannot avoid that a negative cycle is formed in the finite-duration game, then it is easy to show that player 2 can fix a strategy to ensure that only negative cycles are formed in the energy game, which would exhaust any arbitrary finite initial credit. Thus player 1 cannot win the energy game.

This argument reduces energy games to a reachability game in the finite tree obtained by unfolding the original graph and stopping a branch whenever a cycle is formed. Each leaf corresponds to the closing of a cycle, and the leaves associated with a positive cycle define the target nodes of a reachability objective for player 1. Using backward induction on the tree it is easy to establish that from all nodes where player 1 has a winning strategy for the reachability objective, finite initial credit is sufficient for the energy objective.

Finally, to establish the equivalence with mean-payoff games [6, 23], it is easy to see that if player 1 can ensure that only nonnegative cycles are formed along a play, than the mean-payoff value is nonnegative (both for the limsup and the liminf variants), and otherwise player 2 can ensure that only negative cycles are formed, and the mean-payoff value is negative. It follows that mean-payoff games are determined (i.e., if player 1 does not have a winning strategy for $\overline{\mathsf{MP}}^{\geq 0}$, then player 2 has a winning strategy for the complement $V^{\omega} \setminus \overline{\mathsf{MP}}^{\geq 0}$). Moreover, viewing energy games as a safety game it is easy to show that memoryless strategies are sufficient to win energy games, and that the same strategy can be used for the mean-payoff objective, showing the memoryless determinacy of mean-payoff games.

We can solve mean-payoff games in NP by guessing a memoryless strategy for player 1 and checking that it induces only nonnegative reachable cycles in the graph game, which can be done in polynomial time using shortest-path algorithms. By memoryless determinacy, a coNP algorithm can guess a memoryless winning strategy for player 2. Hence mean-payoff games are in NP $\cap$ coNP [32, 43], which can be improved to UP $\cap$ coUP [30]. Mean-payoff games can be solved in $O(|V|\cdot|E|\cdot W)$, thus in P for weights encoded in unary [9]. It is a long-standing open question to know whether mean-payoff games can be solved in polynomial time for weights encoded in binary.

Multi-Dimension Mean-Payoff Games In multi-dimension mean-payoff games, the weight function $w : E \to \mathbb{Z}^k$ assigns a vector of resource consumption to each edge of the graph. The objective of player 1 is to ensure the mean-payoff objective in each dimension $1, \dots, k$, thus a conjunction of k one-dimension mean-payoff objectives with threshold 0. Hence for player 2 the objective is a disjunction of mean-payoff objectives. Simple examples show that infinite memory may be necessary for player 1, for both the limsup and the liminf variants [42]. Moreover, the conjunction of mean-payoff objectives is not equivalent to a conjunction of energy conditions (which requires to maintain the accumulated resource level always nonnegative in every dimension, for some vector of finite initial credit).

Since infinite memory may be necessary for player 1, we consider the problem of deciding the existence of a winning strategy with the following memory restrictions: memoryless strategies, finite-memory strategies, and the general case of infinite-memory strategies [42]. With memoryless strategies, the decision problem is NP-complete. With finite-memory strategies, multi-dimension mean-payoff games are equivalent to energy games, in both the limsup and the liminf variants; memoryless strategies are sufficient for player 2 and the problem is coNP-complete. In the general case with infinite-memory strategies, the problem is in NP $\cap$ coNP for conjunctions of limsup-average, and coNP-complete for conjunctions of liminf-average. In both cases memoryless strategies are sufficient for player 2. Thus in all cases player 2 does not need memory, but the proofs of the memoryless results rely on different techniques (see next paragraph).

Memoryless Strategies in Mean-Payoff Games In the results presented above, it is crucial to establish that memoryless strategies are sufficient for one player (and sometimes for both). We emphasize that different techniques can be used to prove those memorylessness results.

Edge induction and shuffling. A general technique is to use edge induction [25, 34]: to prove that a player is memoryless, consider for example a vertex v owned by that player with two outgoing edges, and consider the game G_1 obtained by removing the first outgoing edge, and the game G_2 obtained by removing the second outgoing edge. By induction we argue that memoryless strategies are sufficient for the player owning v in the games G_1 and G_2, and we need to show that no (arbitrary) strategy in the original game can achieve a better value than either the optimal (memoryless) strategy in G_1 or in G_2. Thus switching between the two outgoing transitions when visiting the vertex v does not give a better play (for the owner of v). Roughly, such a play can be decomposed into the appropriate shuffling of a play in G_1 and a play in G_2, and essentially it suffices to show (in case the owner of v is a minimizer, i.e., player 2) that the mean-payoff value of a shuffle of two plays is not smaller than the min of the values of the two plays, which holds for liminf-average (but not for limsup-average [33, 42]). Dually, in case the owner of v is a maximizer, i.e., player 1, we need to show that the mean-payoff value of a shuffle of two plays is not greater than the max of the values of the two plays.

This technique can be used to show that player 2 is memoryless in multi-dimension energy games, and in the liminf-average variant of mean-payoff games.

Backward induction. To show memorylessness in one-dimension energy games, a simpler proof uses a monotonicity property that if an outgoing edge is a good choice from a vertex v when the current accumulated resource level is x, then the same outgoing edge is a good choice in v for all accumulated resource levels $x' > x$. Therefore in every vertex there exists a choice of outgoing edge that is independent of the current resource level and defines a memoryless winning strategy. This argument is similar to the proof that safety games admit memoryless winning strategies: to win a safety game it is sufficient to always choose a successor vertex that lies in the winning set.

Nested memoryless objectives. As the edge-induction technique does not work to show that player 2 is memoryless in limsup-average mean-payoff games, we use a specific result, based on nested (iterated) elimination of losing vertices: if from a given vertex v player 1 cannot win for one of the limsup-average mean-payoff objective, then v is a losing vertex for player 1 (and player 2 can use a memoryless strategy to win). Given a set $L \subseteq V$ of such losing vertices, all vertices from which player 2 can ensure to reach L are also losing for player 1. Note that player 2 can use a memoryless strategy to reach L. As long as such losing vertices (for player 1) exist, the above argument shows that player 2 has a memoryless strategy to win. The conclusion of the argument is to show that if no losing vertex remains (for any of the one-dimension objectives), then player 1 wins for the multi-dimension objective from every remaining vertex [42].

Related Work We give pointers to the literature related to multiple mean-payoff conditions in graphs and games. Deriving a deterministic algorithm from the coNP result for multi-dimension mean-payoff games gives an algorithm that is exponential in the number of vertices. The hyperplane separation technique gives an algorithm that is polynomial for fixed dimension k and bounded weights [21]. The technique has been extended to obtain a pseudo-polynomial algorithm for solving multi-dimension energy games of fixed dimension with a fixed initial credit [31]. Strategy synthesis is studied in [20] showing that exponential memory is sufficient for winning strategies in multi-dimension energy games (and thus in multi-dimension mean-payoff games under finite-memory strategies). Finitary variants of mean-payoff objectives have been considered in [17].

The vector of Pareto-optimal thresholds for multi-dimension mean-payoff games is studied in [8], showing that deciding if there exists a vector of threshold in a given polyhedron that can be ensured by player 1 is complete for $\mathrm{NP}^{\mathrm{NP}}$. Games with a Boolean combination of mean-payoff objectives are undecidable [41], and their restriction to finite-memory strategies is inter-reducible with Hilbert's tenth problem over the rationals [40]. Mean-payoff conditions have been used to define quantitative specification frameworks with appealing expressiveness and closure properties [2, 15]. The special case of a Boolean combination of mean-payoff objectives defined by a one-dimension weight function gives rise to interval objectives, solvable in NP $\cap$ coNP [29].

We discuss three directions to extend the model of two-player mean-payoff games. First, the mean-payoff objective has been combined with Boolean objectives such as the parity condition, a canonical form to express ω-regular conditions, in one dimension [13, 19] and in multiple dimensions [5, 20]. Note that parity games can be reduced to (one-dimension) mean-payoff games [30], which can be reduced to discounted-sum games, themselves reducible to simple stochastic games [43]. Second, the mean-payoff objective has been considered in Markov decision processes [7, 12] and stochastic games [4, 14], in combination with parity condition [16] and under finite-memory strategies [14]. Third, mean-payoff and energy games (in single dimension) with partial observation have been shown undecidable [22, 28]. Special assumptions on the structure of the game lead to decidable subclasses [28]. Partial-observation energy games with fixed initial credit are Ackermann-complete [35].

References

1. B. Alpern, A. J. Demers, and F. B. Schneider. Safety without stuttering. *Information Processing Letters*, 23(4):177–180, 1986.
2. R. Alur, A. Degorre, O. Maler, and G. Weiss. On omega-languages defined by mean-payoff conditions. In *Proc. of FoSSaCS: Foundations of Software Science and Computational Structures*, LNCS 5504, pages 333–347. Springer, 2009.
3. R. Alur, T. A. Henzinger, and O. Kupferman. Alternating-time temporal logic. *Journal of the ACM*, 49:672–713, 2002.
4. N. Basset, M. Z. Kwiatkowska, U. Topcu, and C. Wiltsche. Strategy synthesis for stochastic games with multiple long-run objectives. In *TACAS*, LNCS 9035, pages 256–271. Springer, 2015.
5. A. Bohy, V. Bruyère, E. Filiot, and J.-F. Raskin. Synthesis from LTL specifications with mean-payoff objectives. In *Proc. of TACAS: Tools and Algorithms for the Construction and Analysis of Systems*, LNCS 7795, pages 169–184. Springer, 2013.
6. P. Bouyer, U. Fahrenberg, K. G. Larsen, N. Markey, and J. Srba. Infinite runs in weighted timed automata with energy constraints. In *Proc. of FORMATS: Formal Modeling and Analysis of Timed Systems*, LNCS 5215, pages 33–47. Springer, 2008.
7. T. Brázdil, V. Brožek, K. Chatterjee, V. Forejt, and A. Kučera. Markov decision processes with multiple long-run average objectives. *Logical Methods in Computer Science*, 10(1:13), 2014.
8. R. Brenguier and J.-F. Raskin. Pareto curves of multidimensional mean-payoff games. In *Proc. of CAV: Computer-Aided Verification*, LNCS 9207, pages 251–267. Springer, 2015.
9. L. Brim, J. Chaloupka, L. Doyen, R. Gentilini, and J.-F. Raskin. Faster algorithms for mean-payoff games. *Formal Methods in System Design*, 38(2):97–118, 2011.

10. J. R. Büchi and L. H. Landweber. Solving sequential conditions by finite-state strategies. *Transactions of the American Mathematical Society*, 138:295–311, 1969.
11. A. Chakrabarti, L. de Alfaro, T. A. Henzinger, and M. Stoelinga. Resource interfaces. In *Proc. of EMSOFT: Embedded Software*, LNCS 2855, pages 117–133. Springer, 2003.
12. K. Chatterjee. Markov decision processes with multiple long-run average objectives. In *Proc. of FSTTCS: Foundation of Software Technology and Theoretical Computer Science*, LNCS 4855, pages 473–484. Springer, 2007.
13. K. Chatterjee and L. Doyen. Energy parity games. In *Proc. of ICALP: Automata, Languages and Programming (Part II)*, LNCS 6199, pages 599–610. Springer, 2010.
14. K. Chatterjee and L. Doyen. Perfect-information stochastic games with generalized mean-payoff objectives. In *Proc. of LICS: Logic in Computer Science*, pages 247–256. ACM, 2016.
15. K. Chatterjee, L. Doyen, H. Edelsbrunner, T. A. Henzinger, and P. Rannou. Mean-payoff automaton expressions. In *Proc. of CONCUR: Concurrency Theory*, LNCS 6269, pages 269–283. Springer, 2010.
16. K. Chatterjee, L. Doyen, H. Gimbert, and Y. Oualhadj. Perfect-information stochastic mean-payoff parity games. In *Proc. of FoSSaCS: Foundations of Software Science and Computation Structures*, LNCS 8412, pages 210–225. Springer-Verlag, 2014.
17. K. Chatterjee, L. Doyen, M. Randour, and J.-F. Raskin. Looking at mean-payoff and total-payoff through windows. In *Proc. of ATVA: Automated Technology for Verification and Analysis*, LNCS 8172, pages 118–132. Springer, 2013.
18. K. Chatterjee and T. A. Henzinger. A survey of stochastic ω-regular games. *J. Comput. Syst. Sci.*, 78(2):394–413, 2012.
19. K. Chatterjee, T. A. Henzinger, and M. Jurdziński. Mean-payoff parity games. In *Proc. of LICS: Logic in Computer Science*, pages 178–187. IEEE Computer Society, 2005.
20. K. Chatterjee, M. Randour, and J.-F. Raskin. Strategy synthesis for multi-dimensional quantitative objectives. *Acta Inf.*, 51:129–163, 2014.
21. K. Chatterjee and Y. Velner. Hyperplane separation technique for multidimensional mean-payoff games. In *Proc. of CONCUR: Concurrency Theory*, pages 500–515, 2013.
22. A. Degorre, L. Doyen, R. Gentilini, J.-F. Raskin, and S. Toruńczyk. Energy and mean-payoff games with imperfect information. In *Proc. of CSL: Computer Science Logic*, LNCS 6247, pages 260–274. Springer, 2010.
23. A. Ehrenfeucht and J. Mycielski. Positional strategies for mean payoff games. *Int. Journal of Game Theory*, 8(2):109–113, 1979.
24. E. A. Emerson and C. Jutla. Tree automata, mu-calculus and determinacy. In *Proc. of FOCS: Foundations of Computer Science*, pages 368–377. IEEE, 1991.
25. H. Gimbert and W. Zielonka. When can you play positionally? In *Proc. of MFCS: Mathematical Foundations of Computer Science*, LNCS 3153, pages 686–697. Springer, 2004.
26. E. Grädel, W. Thomas, and T. Wilke, editors. *Automata, Logics, and Infinite Games: A Guide to Current Research*, LNCS 2500. Springer, 2002.
27. Y. Gurevich and L. Harrington. Trees, automata, and games. In *Proc. of STOC: Symposium on Theory of Computing*, pages 60–65. ACM Press, 1982.
28. P. Hunter, G. A. Pérez, and J.-F. Raskin. Mean-payoff games with partial-observation - (extended abstract). In *Proc. of RP: Reachability Problems*, LNCS 8762, pages 163–175. Springer, 2014.

29. P. Hunter and J.-F. Raskin. Quantitative games with interval objectives. In *Proc. of FSTTCS: Foundation of Software Technology and Theoretical Computer Science*, volume 29 of *LIPIcs*, pages 365–377. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2014.
30. M. Jurdziński. Deciding the winner in parity games is in UP $\cap$ co-UP. *Inf. Process. Lett.*, 68(3):119–124, 1998.
31. M. Jurdziński, R. Lazić, and S. Schmitz. Fixed-dimensional energy games are in pseudo-polynomial time. In *Proc. of ICALP (II): Automata, Languages, and Programming*, LNCS 9135, pages 260–272. Springer, 2015.
32. A. V. Karzanov and V. N.Lebedev. Cyclical games with prohibitions. *Mathematical Programming*, 60:277–293, 1993.
33. E. Kelmendi. *Two-Player Stochastic Games with Perfect and Zero Information*. PhD thesis, Université de Bordeaux, 2016.
34. E. Kopczyński. Half-positional determinacy of infinite games. In *Proc. of ICALP: Automata, Languages and Programming (Part II)*, LNCS 4052, pages 336–347. Springer, 2006.
35. G. A. Pérez. The fixed initial credit problem for partial-observation energy games is Ack-complete. *Information Processing Letters*, 118:91–99, 2017.
36. A. Pnueli and R. Rosner. On the synthesis of a reactive module. In *Proc. of POPL*, pages 179–190. ACM Press, 1989.
37. M. O. Rabin. *Automata on Infinite Objects and Church's Problem*. American Mathematical Society, Boston, MA, USA, 1972.
38. P. J. Ramadge and W. M. Wonham. The control of discrete event systems. *Proceedings of the IEEE*, 77:81–98, 1989.
39. J.-F. Raskin. A tutorial on mean-payoff and energy games. In *Dependable Software Systems Engineering*, volume 45, pages 179–201, 2016.
40. Y. Velner. Finite-memory strategy synthesis for robust multidimensional mean-payoff objectives. In *Proc. of CSL-LICS: Joint Meeting of Computer Science Logic (CSL) and Logic in Computer Science (LICS)*, pages 79:1–79:10. ACM, 2014.
41. Y. Velner. Robust multidimensional mean-payoff games are undecidable. In *Proc. of FoSSaCS: Foundations of Software Science and Computation Structures*, LNCS 9034, pages 312–327. Springer, 2015.
42. Y. Velner, K. Chatterjee, L. Doyen, T. A. Henzinger, A. M. Rabinovich, and J.-F. Raskin. The complexity of multi-mean-payoff and multi-energy games. *Inf. Comput.*, 241:177–196, 2015.
43. U. Zwick and M. Paterson. The complexity of mean payoff games on graphs. *Theor. Comput. Sci.*, 158(1&2):343–359, 1996.