Arena-Independent Finite-Memory Strategies

Pierre Vandenhove[1,2]

Based on joint work with Patricia Bouyer[1], Stéphane Le Roux[1], Youssouf Oualhadj[3], Mickael Randour[2].

[1]LMF, Université Paris-Saclay, CNRS, ENS Paris-Saclay, France

[2]F.R.S.-FNRS & UMONS – Université de Mons, Belgium

[3]LACL – Université Paris-Est Créteil, France

April 23, 2021 – GT Model-Checking & Synthèse

Outline

Strategy synthesis for two-player turn-based games

Design **optimal** controllers for systems interacting with an **antagonistic** environment.

"Optimal" w.r.t. an objective or a specification.

Goal: interest in "simple" controllers

Finite-memory determinacy: when do **finite-memory** controllers suffice?

Inspiration

Results by Gimbert and Zielonka[1,2] about **memoryless** determinacy.

[1] Gimbert and Zielonka, "Games Where You Can Play Optimally Without Any Memory", 2005.

[2] Gimbert and Zielonka, "Pure and Stationary Optimal Strategies in Perfect-Information Stochastic Games with Global Preferences", 2009.

Content

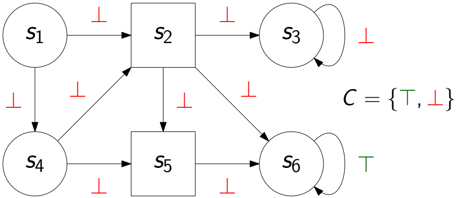
Overview of two papers:

- First results on deterministic games: *Games Where You Can Play Optimally with Arena-Independent Finite Memory*, CONCUR 2020.[3]
- Improvement and extension of our results to stochastic games: *Arena-Independent Finite-Memory Determinacy in Stochastic Games*, 2021.[4]

[3] https://drops.dagstuhl.de/opus/volltexte/2020/12836/

[4] https://arxiv.org/abs/2102.10104

Two-player turn-based zero-sum games on graphs



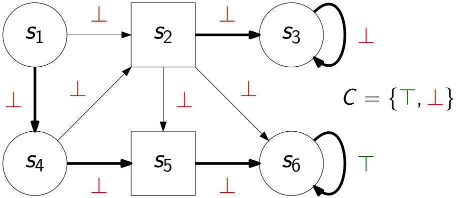
- **Finite** two-player arenas: S_1 ($\bigcirc$, for $\mathcal{P}_1$) and S_2 ($\square$, for $\mathcal{P}_2$), edges E.
- Set C of **colors**. Edges are colored.
- "Objectives" given by **preference relations** $\sqsubseteq \in C^\omega \times C^\omega$ (total preorder). **Zero-sum**.
- A strategy for $\mathcal{P}_i$ is a (partial) function $\sigma\colon E^* \to E$.

Memoryless determinacy

Question

Given a preference relation, do "simple" strategies suffice to play optimally in all arenas?

A strategy σ of $\mathcal{P}_i$ is *memoryless* if it is a function $\cancel{E^*}\ S_i \to E$.



E.g., for reachability, **memoryless** strategies suffice to play optimally.
Also suffice for safety, Büchi, co-Büchi, parity, mean payoff, energy, average-energy...

Memoryless determinacy

Good understanding of memoryless determinacy:

- **sufficient** conditions to guarantee memoryless optimal strategies for **both** players.[5,6]
- **sufficient** conditions to guarantee memoryless optimal strategies for **one** player.[7,8,9]
- **characterization of the preference relations admitting optimal memoryless strategies for both players.**[10]

[5] Gimbert and Zielonka, "When Can You Play Positionally?", 2004.

[6] Aminof and Rubin, "First-cycle games", 2017.

[7] Kopczyński, "Half-Positional Determinacy of Infinite Games", 2006.

[8] Gimbert, "Pure Stationary Optimal Strategies in Markov Decision Processes", 2007.

[9] Gimbert and Kelmendi, "Two-Player Perfect-Information Shift-Invariant Submixing Stochastic Games Are Half-Positional", 2014.

[10] Gimbert and Zielonka, "Games Where You Can Play Optimally Without Any Memory", 2005.

Gimbert and Zielonka's characterization

Let $\sqsubseteq$ be a preference relation. One of the two main results:

One-to-two-player memoryless lift[11]

If

- in all **one-player** arenas of $\mathcal{P}_1$, $\mathcal{P}_1$ has an optimal memoryless strategy,
- in all **one-player** arenas of $\mathcal{P}_2$, $\mathcal{P}_2$ has an optimal memoryless strategy,

then both players have an optimal memoryless strategy in all **two-player** arenas.

Extremely useful in practice. Very easy to recover memoryless determinacy of, e.g., parity and **mean-payoff** games.

[11] Gimbert and Zielonka, "Games Where You Can Play Optimally Without Any Memory", 2005.

Application: memoryless determinacy of mean payoff

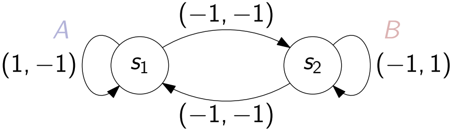
- Colors $C = \mathbb{Z}$. Objective: maximize (for $\mathcal{P}_1$) or minimize (for $\mathcal{P}_2$) the **mean payoff** (average weight by transition).

- In **one-player** arenas, simply **reach** and loop around the **simple cycle** with the **greatest** (for $\mathcal{P}_1$) or **smallest** (for $\mathcal{P}_2$) mean payoff $\rightsquigarrow$ memoryless strategy.

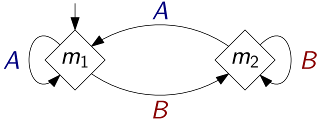
$\Longrightarrow$ Memoryless strategies also suffice to play optimally in **two-player** arenas!

The need for memory

Memoryless strategies do not always suffice.



- Büchi(A) $\wedge$ Büchi(B): requires **finite memory**.



- Mean payoff ≥ 0 in both dimensions: requires **infinite memory**.[12]

$\rightsquigarrow$ **Combinations of objectives** usually require memory.

[12] Chatterjee, Doyen, et al., "Generalized Mean-payoff and Energy Games", 2010.

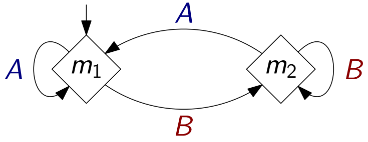
Finite memory

Finite memory ≈ memory structure + next-action function.

Memory structure

Memory structure $\mathcal{M} = (M, m_{\mathsf{init}}, \alpha_{\mathsf{upd}})$: finite set of states M, initial state m_{init}, update function $\alpha_{\mathsf{upd}} \colon M \times C \to M$.

Example for Büchi(A) ∧ Büchi(B) (**not yet a strategy!**):

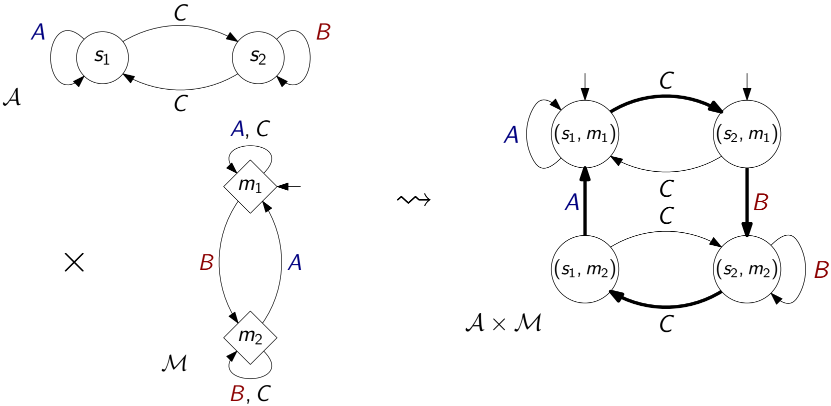


Given an arena $\mathcal{A} = (S_1, S_2, E)$: *next-action function* $\alpha_{\mathsf{nxt}} \colon S_i \times M \to E$.

Finite memory

Playing with memory $\mathcal{M}$ in $\mathcal{A}$ $\approx$ playing memoryless in the arena $\mathcal{A} \times \mathcal{M}$.

Büchi(A) $\wedge$ Büchi(B):



An attempt at lifting [GZ05] to FM determinacy

- Lack of a good understanding of finite-memory determinacy.
- **Related work**: sufficient properties to preserve FM determinacy in Boolean combinations of objectives.[13]
- Our approach:

Hope: extend Gimbert and Zielonka's results

One-to-two-player lift for ~~memoryless~~ **finite-memory** determinacy?

[13] Le Roux, Pauly, and Randour, "Extending Finite-Memory Determinacy by Boolean Combination of Winning Conditions", 2018.

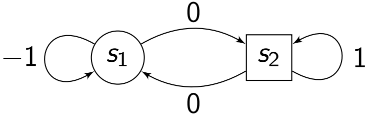
Counterexample to our hope

Let $C \subseteq \mathbb{Z}$. $\mathcal{P}_1$ wants to achieve a play $\pi = c_1 c_2 \ldots \in C^\omega$ s.t.

$$\limsup_n \sum_{i=1}^{n} c_i = +\infty \quad \text{or} \quad \exists^\infty n, \sum_{i=1}^{n} c_i = 0.$$

Optimal **FM** strategies in **one-player** arenas. . .

. . . not in **two-player** arenas: here, $\mathcal{P}_1$ wins but needs **infinite memory**.



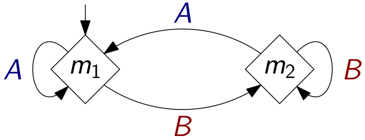
Intuition:

In one-player arenas, $\mathcal{P}_1$ can bound the needed memory in advance.

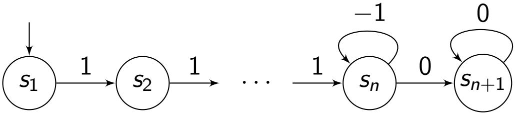
In two-player arenas, $\mathcal{P}_2$ can generate arbitrarily long sequences.

Distinction between the examples

- For Büchi(A) $\wedge$ Büchi(B), this structure suffices **for all** arenas for $\mathcal{P}_1$.



- The counterexample fails because in one-player arenas, the size of the memory is **dependent on the size of the arena**.



In this arena, $\mathcal{P}_1$ needs n memory states to win.

Arena-independent finite memory

Observation: for many objectives, one **fixed memory structure** suffices **for all arenas**.

"For all $\mathcal{A}$, does there exist $\mathcal{M}\ldots$?"
$\rightarrow$ **"Does there exist $\mathcal{M}$, for all $\mathcal{A}\ldots$?"**

Method: reproducing the approach of Gimbert and Zielonka **given an "arena-independent" memory structure $\mathcal{M}$**.

Characterization of arena-independent determinacy

Let $\sqsubseteq$ be preference relation and $\mathcal{M}_1$, $\mathcal{M}_2$ be memory structures.

One-to-two-player arena-independent lift[14]

If

- in all **one-player** arenas of $\mathcal{P}_1$, $\mathcal{P}_1$ has an optimal strategy with memory $\mathcal{M}_1$,
- in all **one-player** arenas of $\mathcal{P}_2$, $\mathcal{P}_2$ has an optimal strategy with memory $\mathcal{M}_2$,

then both players have an optimal strategy in all **two-player** arenas with memory $\mathcal{M}_1 \times \mathcal{M}_2$.

In short: the study of **one-player arenas** is sufficient to determine whether playing with **arena-independent finite memory** suffices.

We recover [GZ05] with $\mathcal{M}_1 = \mathcal{M}_2 = (\{m_{\mathsf{init}}\}, m_{\mathsf{init}}, (m_{\mathsf{init}}, c) \mapsto m_{\mathsf{init}})$.

[14] Bouyer, Le Roux, et al., "Games Where You Can Play Optimally with Arena-Independent Finite Memory", 2020.

Proof technique

One-to-two-player **memoryless** lift[15]

If both players have optimal memoryless strategies in **one-player** arenas, then both players have optimal memoryless strategies in **two-player** arenas.

[15] Gimbert and Zielonka, "Games Where You Can Play Optimally Without Any Memory", 2005.

Issue for arena-independent lift

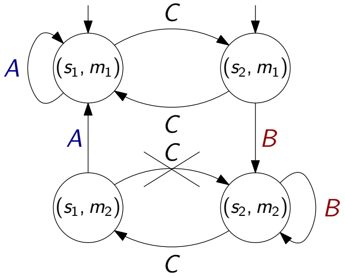
Let $\mathcal{M}_1$, $\mathcal{M}_2$ be memory structures.

One-to-two-player arena-independent lift

If in all **one-player** arenas of $\mathcal{P}_1/\mathcal{P}_2$, $\mathcal{P}_1/\mathcal{P}_2$ has an optimal strategy with memory $\mathcal{M}_1/\mathcal{M}_2$,
then both players have an optimal strategy in all **two-player** arenas with memory $\mathcal{M}_1 \times \mathcal{M}_2$.

Same inductive argument on all **product** arenas with $\mathcal{M}_1 \times \mathcal{M}_2$?

Issue: product arenas are not closed by edge removals.



Solution: covered arenas

We consider the broader class of *covered* arenas.

Covered arenas

An arena $\mathcal{A} = (S, S_1, S_2, E)$ is *covered* by $\mathcal{M} = (M, m_{\mathsf{init}}, \alpha_{\mathsf{upd}})$ from $S_{\mathsf{init}} \subseteq S$ if there exists a function $\phi\colon S \to M$ such that $\phi(S_{\mathsf{init}}) = \{m_{\mathsf{init}}\}$ and for all $(s, c, s') \in E$, $\alpha_{\mathsf{upd}}(\phi(s), c) = \phi(s')$.

Products are covered. Arenas covered by $\mathcal{M}$ are closed by edge removals.

One-to-two-player arena-independent lift

Proof sketch:

- Hypothesis: strategies with memory $\mathcal{M}_1$, $\mathcal{M}_2$ in all one-player arenas
- $\rightsquigarrow$ Memoryless strategies in one-player product arenas with $\mathcal{M}_1$, $\mathcal{M}_2$
- $\rightsquigarrow$ Memoryless strategies in one-player arenas *covered* by $\mathcal{M}_1 \times \mathcal{M}_2$
- $\rightsquigarrow$ Memoryless strategies in two-player arenas *covered* by $\mathcal{M}_1 \times \mathcal{M}_2$ (induction on edges)
- $\rightsquigarrow$ Strategies with memory $\mathcal{M}_1 \times \mathcal{M}_2$ in all two-player arenas.

Applicability and limits

- **Applies to** objectives with optimal **arena-independent** strategies:
  - generalized reachability,[16]
  - generalized parity,[17]
  - window parity,[18]
  - lower- and upper-bounded (multi-dimensional) energy games.[19,20]

- **Does not apply to**, e.g., multi-dimension lower-bounded energy objectives:[21] the size of the finite memory depends on the arena.

---

[16] Fijalkow and Horn, "The surprizing complexity of reachability games", 2010.

[17] Chatterjee, Henzinger, and Piterman, "Generalized Parity Games", 2007.

[18] Bruyère, Hautem, and Randour, "Window parity games: an alternative approach toward parity games with time bounds", 2016.

[19] Bouyer, Markey, et al., "Average-energy games", 2018.

[20] Bouyer, Hofman, et al., "Bounding Average-Energy Games", 2017.

[21] Chatterjee, Randour, and Raskin, "Strategy synthesis for multi-dimensional quantitative objectives", 2014.

Other characterization (not shown here)

Second generalized result: characterization of arena-independent finite-memory determinacy in **one-player** arenas with two properties of $\sqsubseteq$.

Memory requirements of stochastic games

- **Pure** and memoryless strategies also suffice for many objectives: (maximize the probability of) reachability,[22] parity,[23] energy,[24] (maximize the expected value of) discounted sum.[25]

- For some objectives, there is a "constant" blow-up (e.g., *weak parity*; memoryless $\rightsquigarrow$ arena-independent).

[22] Condon, "The Complexity of Stochastic Games", 1992.

[23] Chatterjee, Jurdzinski, and Henzinger, "Quantitative stochastic parity games", 2004.

[24] Brázdil, Brozek, and Etessami, "One-Counter Stochastic Games", 2010.

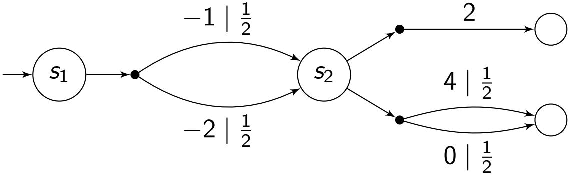
[25] Shapley, "Stochastic Games", 1953.

Greater memory requirements in stochastic games

Objective: maximize the probability of

$$\mathsf{Disc}_{\geq 0} = \{w = w_1 w_2 \ldots \in \mathbb{Q}^\omega \mid \sum_{i=1}^{\infty} \frac{1}{2^{i-1}} w_i \geq 0\}.$$

- Memoryless strategies suffice in **deterministic** games.
- Arena-independent FM strategies do **not** suffice in **stochastic** games.



$\rightsquigarrow$ Possible to win with probability $\frac{3}{4}$ **with memory**.

Results about stochastic games

One-to-two-player **stochastic** lift[26]

If

- in all **one-player** **stochastic** arenas (i.e., MDPs) of $\mathcal{P}_1$, $\mathcal{P}_1$ has a **pure** optimal strategy with memory $\mathcal{M}_1$,
- in all **one-player** **stochastic** arenas (i.e., MDPs) of $\mathcal{P}_2$, $\mathcal{P}_2$ has a **pure** optimal strategy with memory $\mathcal{M}_2$,

then both players have a **pure** optimal strategy in all **two-player** **stochastic** arenas with memory $\mathcal{M}_1 \times \mathcal{M}_2$.

Also:

- characterization in terms of two properties of $\sqsubseteq$.
- equivalence between the existence of arena-independent *subgame perfect* strategies and of arena-independent *optimal* strategies.

[26] Bouyer, Oualhadj, et al., "Arena-Independent Finite-Memory Determinacy in Stochastic Games", 2021.

Summary

Key observation: **arena-independent** memory often suffices.

Contributions

- **One-to-two-player lift** in deterministic and stochastic games.
- Characterization of **arena-independent** finite-memory determinacy.

Ongoing work

- Understand the arena-**dependent** case.
- Similar one-to-two-player lift for **infinite** arenas.

Thanks! Questions?

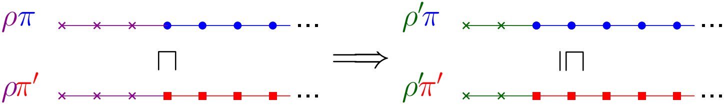
Appendix

Gimbert and Zielonka's characterization[27]

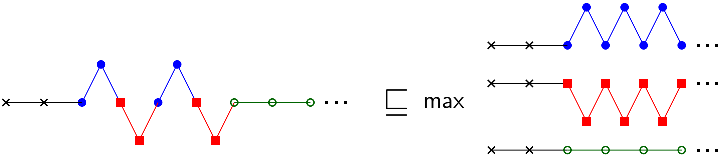
Let $\sqsubseteq$ be a preference relation.

$\mathcal{P}_1$ admits optimal memoryless strategies in **one-player** arenas
if and only if

1. $\sqsubseteq$ is monotone: not sensitive to changing prefixes.



2. $\sqsubseteq$ is selective: mixing cycles is useless.

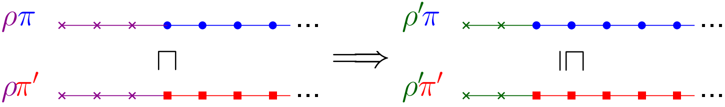


[27] Gimbert and Zielonka, "Games Where You Can Play Optimally Without Any Memory", 2005.

Characterization of arena-independent finite memory

Let $\sqsubseteq$. Let $\mathcal{M} = (M, m_{\text{init}}, \alpha_{\text{upd}})$.

- We classify prefixes according to $\mathcal{M}$:
  for $\rho, \rho' \in C^*$, $\rho \sim_{\mathcal{M}} \rho'$ iff $\alpha_{\text{upd}}(m_{\text{init}}, \rho) = \alpha_{\text{upd}}(m_{\text{init}}, \rho')$.
- From monotone to $\mathcal{M}$-monotone: same with $\rho \sim_{\mathcal{M}} \rho'$.



- Similar extension of selective to $\mathcal{M}$-selective by classifying cycles in the memory structure.

Proposition

$\mathcal{P}_1$ has optimal strategies with memory $\mathcal{M}$ in all **one-player** arenas **if and only if** $\sqsubseteq$ is $\mathcal{M}$-monotone and $\mathcal{M}$-selective.

Formal definitions of $\mathcal{M}$-monotony and $\mathcal{M}$-selectivity

Definition ($\mathcal{M}$-monotony)

Let $\mathcal{M} = (M, m_{\mathsf{init}}, \alpha_{\mathsf{upd}})$ be a memory structure. A preference relation $\sqsubseteq$ is $\mathcal{M}$-*monotone* if for all $m \in M$, for all $K_1, K_2 \in \mathcal{R}(C)$,

$$\exists\, w \in L_{m_{\mathsf{init}},m},\ [wK_1] \sqsubset [wK_2] \implies \forall\, w' \in L_{m_{\mathsf{init}},m},\ [w'K_1] \sqsubseteq [w'K_2].$$

Definition ($\mathcal{M}$-selectivity)

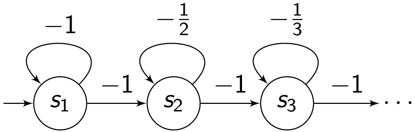
Let $\mathcal{M} = (M, m_{\mathsf{init}}, \alpha_{\mathsf{upd}})$ be a memory structure. A preference relation $\sqsubseteq$ is $\mathcal{M}$-*selective* if for all $w \in C^*$, $m = \widehat{\alpha_{\mathsf{upd}}}(m_{\mathsf{init}}, w)$, for all $K_1, K_2 \in \mathcal{R}(C)$ such that $K_1, K_2 \subseteq L_{m,m}$, for all $K_3 \in \mathcal{R}(C)$,

$$[w(K_1 \cup K_2)^* K_3] \sqsubseteq [wK_1^*] \cup [wK_2^*] \cup [wK_3].$$

Greater memory requirements in **infinite** arenas

Objective: get the largest mean payoff.

- Memoryless strategies suffice in finite (even stochastic) arenas.
- Infinite memory is required in one-player deterministic infinite arenas.[28]



⇝ Possible to get 0 at the limit **with infinite memory**.

[28] Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, 1994.