Games Where You Can Play Optimally with Arena-Independent Finite Memory

Patricia Bouyer[1] Stéphane Le Roux[1] Youssouf Oualhadj[2]
Mickael Randour[3] **Pierre Vandenhove**[1,3]

[1]LSV – CNRS & ENS Paris-Saclay, Université Paris-Saclay, France

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

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

June 22, 2020 – MOVEP 2020

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.

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?

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] about memoryless determinacy.

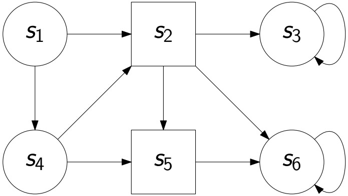
[1] Gimbert and Zielonka, "Games Where You Can Play Optimally Without Any Memory", 2005.

1 Memoryless determinacy

2 The need for memory

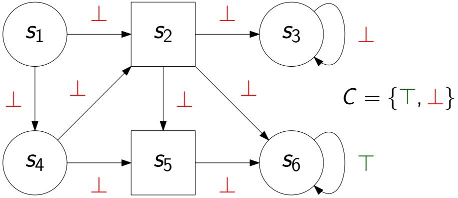
3 Arena-independent finite memory

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



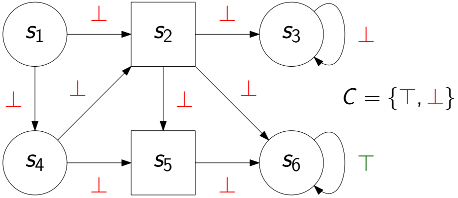
- Finite two-player arenas: S_1 (circles, for $\mathcal{P}_1$) and S_2 (squares, for $\mathcal{P}_2$), edges E.

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



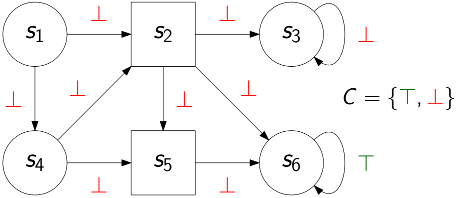
- Finite two-player arenas: S_1 (circles, for $\mathcal{P}_1$) and S_2 (squares, for $\mathcal{P}_2$), edges E.
- Set C of colors. Edges are colored.

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



- Finite two-player arenas: S_1 (circles, for $\mathcal{P}_1$) and S_2 (squares, 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, $\sqsubseteq^{-1}$.

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



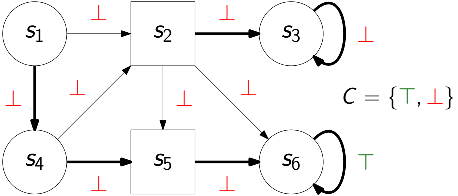
- Finite two-player arenas: S_1 (circles, for $\mathcal{P}_1$) and S_2 (squares, for $\mathcal{P}_2$), edges E.
- Set C of colors. Edges are colored.
- "Objectives" given by preference relations $\sqsubseteq \in C^\omega \times C^\omega$ (total preorder). Zero-sum, $\sqsubseteq^{-1}$.
- A strategy for $\mathcal{P}_i$ 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 ~~E^*~~ $S_i \to E$.

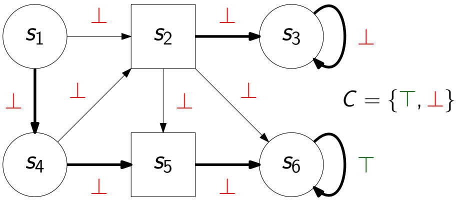


Memoryless determinacy

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 ~~E^*~~ $S_i \to E$.



E.g., for reachability, memoryless strategies suffice.
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.[2,3]

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

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

Memoryless determinacy

Good understanding of memoryless determinacy:

- **sufficient** conditions to guarantee memoryless optimal strategies for **both** players.[2,3]
- **sufficient** conditions to guarantee memoryless optimal strategies for **one** player.[4,5,6]

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

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

[4]Kopczynski, "Half-Positional Determinacy of Infinite Games", 2006.

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

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

Memoryless determinacy

Good understanding of memoryless determinacy:

- **sufficient** conditions to guarantee memoryless optimal strategies for **both** players.[2,3]
- **sufficient** conditions to guarantee memoryless optimal strategies for **one** player.[4,5,6]
- **characterization of the preference relations admitting optimal memoryless strategies for both players.**[7]

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

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

[4] Kopczynski, "Half-Positional Determinacy of Infinite Games", 2006.

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

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

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

Gimbert and Zielonka's characterization[8]

Let $\sqsubseteq$ be a preference relation. Two results:

1. Characterization of memoryless determinacy w.r.t. properties of $\sqsubseteq$.

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

Gimbert and Zielonka's characterization[8]

Let $\sqsubseteq$ be a preference relation. Two results:

1. Characterization of memoryless determinacy w.r.t. properties of $\sqsubseteq$.
2. Corollary:

One-to-two-player memoryless lifting

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.

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

Gimbert and Zielonka's characterization[8]

Let $\sqsubseteq$ be a preference relation. Two results:

1. Characterization of memoryless determinacy w.r.t. properties of $\sqsubseteq$.
2. Corollary:

One-to-two-player memoryless lifting

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., mean-payoff and parity games.

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

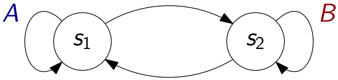
1 Memoryless determinacy

2 The need for memory

3 Arena-independent finite memory

The need for memory

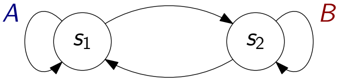
Memoryless strategies do not always suffice.



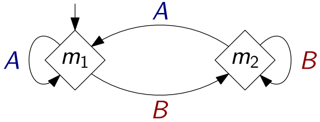
- Büchi(A) $\wedge$ Büchi(B): requires **finite memory**.

The need for memory

Memoryless strategies do not always suffice.

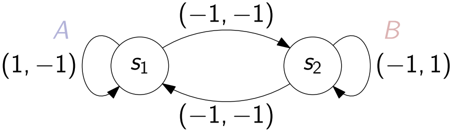


- Büchi(A) ∧ Büchi(B): requires **finite memory**.

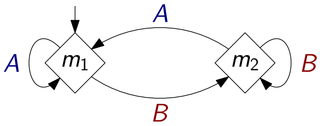


The need for memory

Memoryless strategies do not always suffice.



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

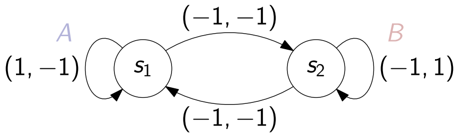


- Mean payoff ≥ 0 in both dimensions: requires **infinite memory**.[9]

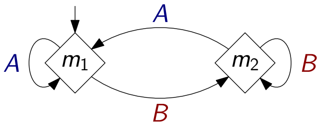
[9] Chatterjee, Doyen, et al., "Generalized Mean-payoff and Energy Games", 2010.

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**.[9]

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

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

An attempt at lifting [GZ05] to FM determinacy

- Lack of a good understanding of finite-memory determinacy.

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.[10]

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

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.[10]
- Our approach:

Hope: extend Gimbert and Zielonka's results

One-to-two-player lifting for ~~memoryless~~ **finite-memory** determinacy.

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

Counterexample

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

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

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

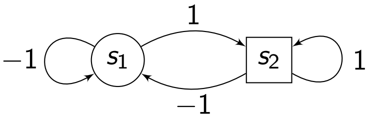
Counterexample

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

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

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

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



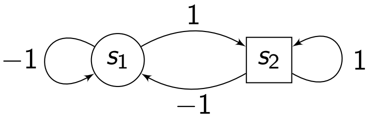
Counterexample

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

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

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

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



Intuition:

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

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

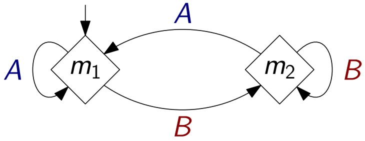
1 Memoryless determinacy

2 The need for memory

3 Arena-independent finite memory

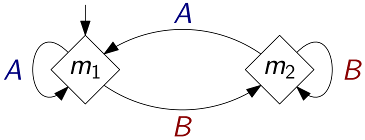
Arena-independent memory

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



Arena-independent memory

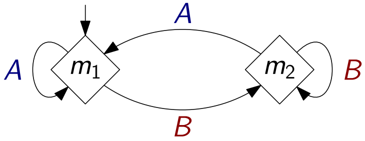
- For Büchi(A) $\wedge$ Büchi(B), this structure suffices to play optimally on **all** arenas for $\mathcal{P}_1$.



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

Arena-independent memory

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



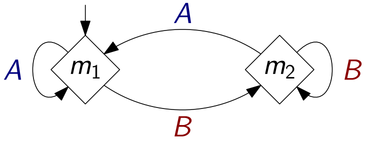
- The counterexample fails because in one-player arenas, the size of the memory is **dependent on the size of the arena**.
- Observation: for many objectives, one **fixed memory structure** suffices **for all arenas**.

"For all $\mathcal{A}$, does there exist $\mathcal{M}$. . . ?"

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

Arena-independent memory

- For Büchi(A) $\wedge$ Büchi(B), this structure suffices to play optimally on **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**.
- Observation: for many objectives, one **fixed memory structure** suffices **for all arenas**.

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

Method: reproducing the approach of Gimbert and Zielonka **given a memory structure** $\mathcal{M}$.

Characterization of arena-independent determinacy

Let $\sqsubseteq$ be preference relation, $\mathcal{M}$ be a memory structure.

1. Characterization of "playing with $\mathcal{M}$ is sufficient" in terms of properties of $\sqsubseteq$.

Characterization of arena-independent determinacy

Let $\sqsubseteq$ be preference relation, $\mathcal{M}$ be a memory structure.

1. Characterization of "playing with $\mathcal{M}$ is sufficient" in terms of properties of $\sqsubseteq$.
2. Corollary:

One-to-two-player lifting

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 $\mathcal{M}_1 \otimes \mathcal{M}_2$.

Characterization of arena-independent determinacy

Let $\sqsubseteq$ be preference relation, $\mathcal{M}$ be a memory structure.

1. Characterization of "playing with $\mathcal{M}$ is sufficient" in terms of properties of $\sqsubseteq$.
2. Corollary:

One-to-two-player lifting

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 $\mathcal{M}_1 \otimes \mathcal{M}_2$.

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

Applicability and limits

- **Applies to** objectives with optimal **arena-independent** strategies:
  - generalized reachability,[11]
  - generalized parity,[12]
  - window parity,[13]
  - lower- and upper-bounded (multi-dimensional) energy games.[14,15]

---

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

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

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

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

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

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

Applicability and limits

- **Applies to** objectives with optimal **arena-independent** strategies:
  - generalized reachability,[11]
  - generalized parity,[12]
  - window parity,[13]
  - lower- and upper-bounded (multi-dimensional) energy games.[14,15]

- **Does not apply to**, e.g., multi-dimension lower-bounded energy objectives:[16] the size of the finite memory depends on the arena.

---

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

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

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

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

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

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

Conclusion

Key observation: many objectives require arena-independent memory.

Conclusion

Key observation: many objectives require arena-independent memory.

Contributions

- Characterization of arena-independent finite-memory determinacy.
- One-to-two-player lifting.
- Generalization of Gimbert and Zielonka's work.

Conclusion

Key observation: many objectives require arena-independent memory.

Contributions

- Characterization of arena-independent finite-memory determinacy.
- One-to-two-player lifting.
- Generalization of Gimbert and Zielonka's work.

Future work

Understand (arena-**dependent**) finite-memory determinacy through the study of one-player arenas.

Conclusion

Key observation: many objectives require arena-independent memory.

Contributions

- Characterization of arena-independent finite-memory determinacy.
- One-to-two-player lifting.
- Generalization of Gimbert and Zielonka's work.

Future work

Understand (arena-**dependent**) finite-memory determinacy through the study of one-player arenas.

Thanks!

Conclusion

Key observation: many objectives require arena-independent memory.

Contributions

- Characterization of arena-independent finite-memory determinacy.
- One-to-two-player lifting.
- Generalization of Gimbert and Zielonka's work.

Future work

Understand (arena-**dependent**) finite-memory determinacy through the study of one-player arenas.

Thanks!

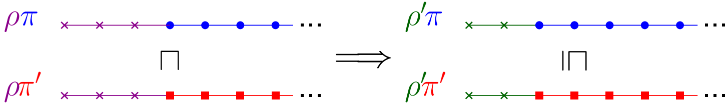
Appendix

Gimbert and Zielonka's characterization[17]

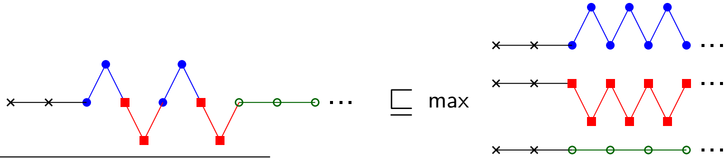
Let $\sqsubseteq$ be a preference relation.

Both players admit optimal memoryless strategies in all arenas
if and only if

1. $\sqsubseteq$ and $\sqsubseteq^{-1}$ are monotone: not sensitive to changing prefixes.



2. $\sqsubseteq$ and $\sqsubseteq^{-1}$ are selective: mixing cycles is useless.



[17] 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 **highest/lowest** mean-payoff (for $\mathcal{P}_1/\mathcal{P}_2$) $\rightsquigarrow$ memoryless strategy.

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

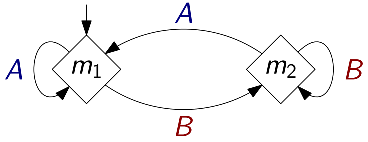
Finite memory

Finite memory $\approx$ memory structure + next-action function.

Memory structure

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

Example for Büchi(A) $\wedge$ Büchi(B) (**not yet a strategy!**):

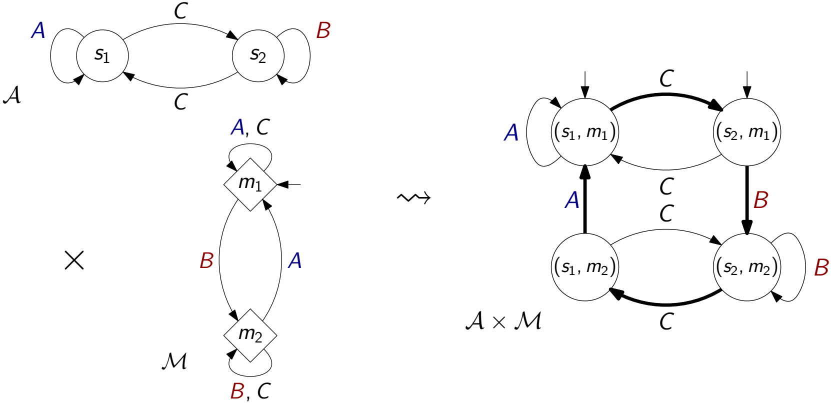


Given an arena $\mathcal{A} = (S_1, S_2, E)$: *next-action function* $\alpha_{\text{nxt}} \colon S_i \times M \to E$.

FM example

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

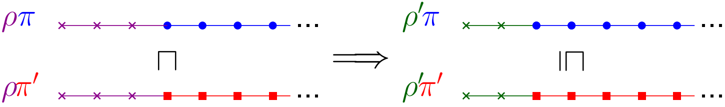
Büchi(A) $\wedge$ Büchi(B):



Characterization of arena-independent finite memory

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

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

Let $\sqsubseteq$, $\mathcal{M}$. $\mathcal{P}_1$ and $\mathcal{P}_2$ have optimal strategies with memory $\mathcal{M}$ in all arenas **if and only if** $\sqsubseteq$ and $\sqsubseteq^{-1}$ are $\mathcal{M}$-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 skeleton. 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 skeleton. 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 $K_3 \in \mathcal{R}(C)$,

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