Arena-Independent Finite-Memory Determinacy in Stochastic Games

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

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

[2]Univ Paris Est Creteil, LACL, France

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

August 2021 – CONCUR 2021

Outline

Strategy synthesis for zero-sum turn-based stochastic games

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

Interest in simple controllers

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

Deterministic $\rightsquigarrow$ stochastic games

Follow-up to CONCUR 2020 paper[1] about deterministic games.

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

1 Stochastic games

2 Arena-independent (or not) finite memory

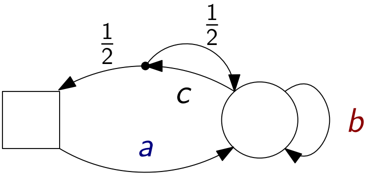
3 Three characterizations

1 Stochastic games

2 Arena-independent (or not) finite memory

3 Three characterizations

Two-player zero-sum stochastic games on graphs



- **Finite arena**: states S_1 ($\bigcirc$, for $\mathcal{P}_1$) and S_2 ($\square$, for $\mathcal{P}_2$), actions A, probabilistic transitions $S \times A \to \mathrm{Dist}(S)$.
- Set C of **colors**. Pairs in $S \times A$ are colored.
- A **strategy** for $\mathcal{P}_i$ is a function $\sigma\colon (SA)^* S_i \to \mathrm{Dist}(A)$.
- "**Objectives**" given by total preorders $\sqsubseteq\, \subseteq \mathrm{Dist}(C^\omega) \times \mathrm{Dist}(C^\omega)$. **Zero-sum**.

Strategy complexity

Main question

Given an objective, how *simple* can optimal strategies be?

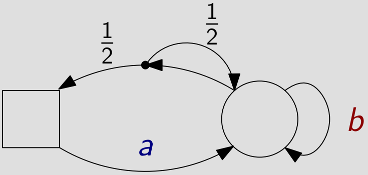
In stochastic games, two measures for strategy complexity:

- **randomization** of strategies;
- **memory** (how much information must be remembered).

Example

Büchi(a) $\wedge$ Büchi(b)

Objective: maximize the probability to see both a and b infinitely often.



Using randomization *or* memory is sufficient.

In this work: **pure** (i.e., **no randomization**) strategies with **finite memory**.

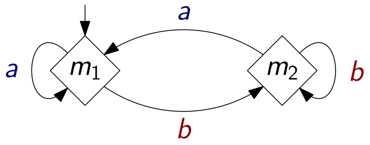
Finite memory

Finite-memory strategy ≈ 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 $\alpha_{\mathsf{upd}} \colon M \times C \to M$.

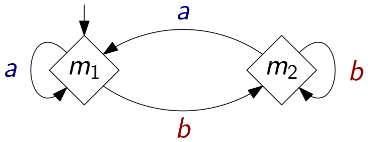
Example for Büchi(a) $\wedge$ Büchi(b) (**not a strategy yet!**):



Given an **arena**: pure **next-action function** $\alpha_{\mathsf{nxt}} \colon S_i \times M \to A$.

Two kinds of finite-memory determinacy

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



$\rightsquigarrow$ **There exists a memory structure that suffices for all arenas.**

- If the objective is to bring the sum back to 0 in this class of 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.

$\rightsquigarrow$ **for every arena in this class, there exists a memory structure. . .**

Pure arena-independent finite memory: examples

- **Memoryless** strategies are arena-independent (reachability,[2] Büchi and parity,[3] discounted sum,[4] one-counter games[5]...).
- All ω-**regular objectives** admit optimal arena-independent finite-memory strategies (Muller objectives,[6] generalized parity objectives[7]...).
- Lexicographic reachability-safety objectives (memory depends on the number of targets).[8]
- Finite but **not** arena-independent: variant of energy-parity games.[9]

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

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

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

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

[6] Chatterjee, "The complexity of stochastic Müller games", 2012.

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

[8] Chatterjee, Katoen, et al., "Stochastic Games with Lexicographic Reachability-Safety Objectives", 2020.

[9] Mayr et al., "Simple Stochastic Games with Almost-Sure Energy-Parity Objectives are in NP and coNP", 2021.

1 Stochastic games

2 Arena-independent (or not) finite memory

3 Three characterizations

One-to-two-player lift

One-to-two-player **stochastic** lift

If objective $\sqsubseteq$ is such that

- 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 $\mathcal{M}_1 \times \mathcal{M}_2$.

Reduces the sufficiency of pure arena-independent FM strategies in **two-player** zero-sum games to the same problem in **one-player** games.

History of one-to-two-player lifts

Classes of arenas/memory for which **two-player zero-sum** memory determinacy reduces to **one-player** determinacy.

Arenas \ Mem. req.	Memoryless	Arena-ind. FM	Mildly growing
Finite deterministic	[GZ05][10]	[BLORV20][11]	[Koz21][12]
Finite stochastic	[GZ09][13]	**New article**	
Infinite deterministic	[CN06,Kop08][14]		

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

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

[12] Kozachinskiy, "One-to-Two-Player Lifting for Mildly Growing Memory", 2021.

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

[14] Colcombet and Niwiński, "On the positional determinacy of edge-labeled games", 2006; Kopczyński, "Half-positional Determinacy of Infinite Games", 2008.

Further characterization in one-player games

Two-player $\rightsquigarrow$ one-player. What about one-player games?

One-player characterization

Let $\sqsubseteq$ be an objective expressible as a payoff function $C^\omega \to \mathbb{R}$.

$\mathcal{M}$ suffices to play optimally with **pure** strategies in all **one-player** stochastic arenas of $\mathcal{P}_1$
if and only if $\sqsubseteq$ is $\mathcal{M}$-**monotone** and $\mathcal{M}$-**selective**.

Further characterization in one-player games

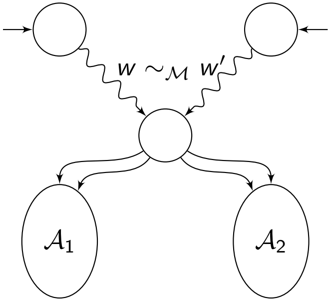
One-player characterization

Let $\sqsubseteq$ be an objective expressible as a payoff function $C^\omega \to \mathbb{R}$.

$\mathcal{M}$ suffices to play optimally with **pure** strategies in all **one-player** stochastic arenas of $\mathcal{P}_1$
if and only if $\sqsubseteq$ is $\mathcal{M}$-**monotone** and $\mathcal{M}$-**selective**.

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

$\mathcal{M}$-**monotony**: for all one-player arenas $\mathcal{A}_1, \mathcal{A}_2$, one arena among $\mathcal{A}_1$ and $\mathcal{A}_2$ suffices to play optimally after all prefixes in the same equivalence class of $\sim_{\mathcal{M}}$.



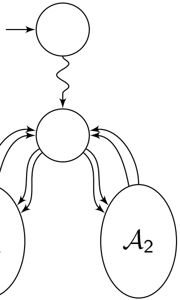
Further characterization in one-player games

One-player characterization

Let $\sqsubseteq$ be an objective expressible as a payoff function $C^\omega \to \mathbb{R}$.

$\mathcal{M}$ suffices to play optimally with **pure** strategies in all **one-player** stochastic arenas of $\mathcal{P}_1$
if and only if $\sqsubseteq$ is $\mathcal{M}$-**monotone** and $\mathcal{M}$-**selective**.

$\mathcal{M}$-**selective**: close idea with cycles on the same state of $\mathcal{M}$.



Further characterization in one-player games

One-player characterization

Let $\sqsubseteq$ be an objective expressible as a payoff function $C^\omega \to \mathbb{R}$.

$\mathcal{M}$ suffices to play optimally with **pure** strategies in all **one-player** stochastic arenas of $\mathcal{P}_1$
if and only if $\sqsubseteq$ is $\mathcal{M}$-**monotone** and $\mathcal{M}$-**selective**.

A less general characterization was known in deterministic games[15,16] but not in stochastic games.

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

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

AIFM and subgame-perfect strategies

The proof was helped by the new following characterization:

Subgame perfection and optimality under pure AIFM strategies

Given an objective, if a memory structure $\mathcal{M}$ suffices to play **optimally** in all arenas with pure strategies,
then $\mathcal{M}$ also suffices to play **subgame-perfect** pure strategies.

Summary

Contributions

- Further the understanding of arena-independent finite-memory determinacy.
- Characterizations of **arena-independent** finite-memory determinacy.

Ongoing work

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

Thanks! Questions?