On finite-memory determinacy of games on graphs

Patricia Bouyer
LSV, CNRS, Univ. Paris-Saclay, ENS Paris-Saclay
France

Based on joint work with Stéphane Le Roux, Youssouf Oualhadj,
Mickael Randour, Pierre Vandenhove
(Published at CONCUR’20)
The talk in one slide

- Find good and simple controllers for systems interacting with an antagonistic environment

- Performance w.r.t. objectives / payoffs / preference relations

- Memoryless strategies
  - Finite-memory strategies

When are simple strategies sufficient to play optimally?
The setting - Example of a game

Reachability winning condition for $P_1$
The setting - Example of a game

Reachability winning condition for $P_1$

Use of colors to define winning condition/preference relation

$$* \ ( \ + \ )^\omega$$
The setting - Example of a game

Reachability winning condition for $P_1$

The game is played using strategies:

$$\sigma_i : S^*S_i \rightarrow E$$
Families of strategies

\[ \sigma_i : S^* S_i \rightarrow E \]

Subclasses of interest

- Memoryless strategy: \( \sigma_i : S_i \rightarrow E \)
- Finite-memory strategy: \( \sigma_i \) defined by a finite-state Mealy machine

« Reach the target »

Loop 5 times in the initial state

« Reach the target with energy 0 »

« Visit both \( s_1 \) and \( s_2 \) »

Every odd visit to \( s_0 \), go to \( s_1 \)
Every even visit to \( s_0 \), go to \( s_2 \)
The setting - Preference relation

A preference relation $\succeq$ is a total preorder on $C^\omega$.

$\pi \succeq \pi'$ and $\pi' \succeq \pi$ means that $\pi$ and $\pi'$ are equally appreciated

$\pi \succeq \pi'$ and $\pi' \not\succeq \pi$ means that $\pi'$ is preferred over $\pi$

Examples

- $W \subseteq C^\omega$ winning condition:
  $\pi \succeq \pi'$ if either $\pi' \in W$ or $\pi \notin W$
- Quantitative real payoff $f$
  $\pi \succeq \pi'$ if $f(\pi) \leq f(\pi')$
  Ex: MP, AE, TP

Zero-sum assumption:
- Preference of $P_1$ is $\succeq$
- Preference of $P_2$ is $\succeq^{-1}$
Payoffs based on energy

Focus on two memoryless strategies
Payoffs based on energy

Focus on two memoryless strategies

- Constraint on the energy level (EL)
Payoffs based on energy

Focus on two memoryless strategies

- Mean-payoff (MP): long-run average payoff per transition

$MP = 0$

$MP = \frac{1}{3}$

- Constraint on the energy level (EL)
- Mean-payoff (MP): long-run average payoff per transition
Payoffs based on energy

Focus on two memoryless strategies

- Constraint on the energy level (EL)
- Mean-payoff (MP): long-run average payoff per transition
Payoffs based on energy

Focus on two memoryless strategies

- Mean-payoff (MP): long-run average payoff per transition
- Constraint on the energy level (EL)
- Total-payoff (TP)

\[
TP = 0, \overline{TP} = 2
\]

\[
TP = 0, \overline{TP} = 1
\]
Payoffs based on energy

Focus on two memoryless strategies

- Mean-payoff (MP): long-run average payoff per transition
- Constraint on the energy level (EL)
- Total-payoff (TP)
Payoffs based on energy

Focus on two memoryless strategies

- Constraint on the energy level (EL)
- Mean-payoff (MP): long-run average payoff per transition
- Total-payoff (TP)
- Average-energy (AE)

\[ AE = \frac{4}{3} \]

\[ AE = 1 \]
Optimality of strategies

\[ \text{Out}(\sigma_n)^T \subseteq \text{Out}(\sigma'_n)^T \]

\[ \Rightarrow \sigma_n \text{ is better than } \sigma'_n \]

\[ \sigma_n \text{ optimal whenever it is better than any other } \sigma'_n \]

Remark

- To be distinguished from:
  - \( \epsilon \)-optimal
  - Subgame-perfect optimal (in our case: Nash equilibria)
A focus on memoryless strategies
When are memoryless strategies sufficient to play optimally?

Quite often!

Examples

- Reachability, safety, Büchi, parity, MP, EL $\geq 0$, TP, AE, etc…

Can we characterize when they are?

YES!

And this is a beautiful result by Gimbert and Zielonka, CONCUR’05
The memoryless story

**Sufficient conditions**

- Sufficient conditions to guarantee memoryless optimal strategies for both player [GZ04,AR17]
- Sufficient conditions to guarantee memoryless optimal strategies for one player (« half-positional ») [Kop06,Gim07,GK14]

- Characterization of the preference relations admitting optimal memoryless strategies for both players in all finite games [GZ05]
The Gimbert-Zielonka characterization for memory less determinacy (1)

Let $\leq$ be a preference relation.

It is said:

• monotone whenever

\[ \implies \]

• selective whenever

\[ \implies \]
The Gimbert-Zielonka characterization for memory less determinacy (2) [GZ05]

Characterization - Two-player games

The two following assertions are equivalent:
1. All finite games have memoryless optimal strategies for both players
2. Both $\Xi$ and $\Xi^{-1}$ are monotone and selective

Characterization - One-player games

The two following assertions are equivalent:
1. All finite $P_1$-games have (uniform) memoryless optimal strategies
2. $\Xi$ is monotone and selective
Assume all $P_1$-games have optimal memoryless strategies.
Assume $\Xi$ is monotone and selective.

The case of one-player games

One best choice between and (monotony).

No reason to swap at $t$ (selectivity).

No memory required at $t$!
Applications

Lifting theorem

- If in all finite one-player game for player $P_i$, $P_i$ has uniform memoryless optimal strategies, then both players have memoryless optimal strategies in all finite two-player games.

Very powerful and extremely useful in practice!

Discussion

- Easy to analyse the one-player case (graph analysis)
  - Mean-payoff, average-energy [BMRL15]
- Allows to deduce properties in the two-player case
Discussion of examples

Examples

- Reachability, safety:
  - Monotone (though not prefix-independent)
  - Selective
- Parity, mean-payoff:
  - Prefix-independent hence monotone
  - Selective
- Priority mean payoff [GZ05]
- Average-energy games [BMRL15]
  - Lifting theorem!!
Discussion

Winning condition for $P_1$:

$((MP \in \mathbb{Q}) \land \text{Büchi}(A)) \lor \text{coBüchi}(B)$

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{i=1}^{n} c_i \in \mathbb{Q}$$

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{i=1}^{n} c_i \in \mathbb{Q}$$
Discussion

Winning condition for $P_1$:

\[((M \in Q) \land \text{Büchi}(A)) \lor \text{coBüchi}(B)\]

- In all one-player games, $P_1$ has a memoryless uniform optimal strategy
- Hence: the winning condition is monotone and selective

\begin{align*}
(0, B) &\quad (1, B) \\
(1, A) &\quad (0, A)
\end{align*}

- $P_1$ wins this game:
  - Infinitely often, give the hand back to $P_2$
  - Play for a long time the edge labelled $(0, B)$ to approach 0
  - Play for a long time the edge labelled $(1, B)$ to approach 1
- It requires infinite memory!
Discussion

Winning condition for $P_1$:

$$((\text{MP} \in \mathbb{Q}) \land \text{Büchi}(A)) \lor \text{coBüchi}(B)$$

If only $\subseteq$ is monotone and selective, $P_1$ might not have a memoryless optimal strategy.
Finite-memory strategies
We need memory!

Objectives/preference relations become more and more complex

- Büchi($A$) $\land$ Büchi($B$) requires finite memory
We need memory!

Objectives/preference relations become more and more complex

- Büchi(A) \( \land \) Büchi(B) requires finite memory

- \( MP_1 \geq 0 \land MP_2 \geq 0 \) requires infinite memory
Can we lift [GZ05] to finite memory?

A priori no...

Consider the following winning condition for $P_1$:

$$\liminf_{n} \sum_{i=1}^{n} c_i = +\infty \text{ or } \exists \infty n \text{ s.t. } \sum_{i=1}^{n} c_i = 0$$

- Optimal finite-memory strategies in one-player games
- But not in two-player games!!

$P_1$ wins but uses infinite memory!
How do we formalize finite memory?

**Standardly**

- A strategy $\sigma_i$ of player $P_i$ has finite memory if it can be encoded as a Mealy machine $(M, m_{\text{init}}, \alpha_{\text{upd}}, \alpha_{\text{next}})$ where $M$ is finite, $m_{\text{init}} \in M$,
  $\alpha_{\text{upd}} : M \times S \to M$ and $\alpha_{\text{next}} : M \times S_i \to E$
  - $(M, m_{\text{init}}, \alpha_{\text{upd}})$ is a memory mechanism
  - $\alpha_{\text{next}}$ gives the next move

**To have an abstract theorem...**

- The memory mechanism should not speak about information specific to particular games, hence:
  - $\alpha_{\text{upd}}$ should not speak of states
  - $\alpha_{\text{upd}}$ can speak of colors
    (notion of « chromatic strategy » by Kopczynski)
Arena-independent memory management

**Memory skeleton**

- $\mathcal{M} = (M, m_{\text{init}}, \alpha_{\text{upd}})$ with $m_{\text{init}} \in M$ and $\alpha_{\text{upd}} : M \times C \rightarrow M$

**Strategy with memory $\mathcal{M}$**

- Additional next-move function: $\alpha_{\text{next}} : M \times S_i \rightarrow E$

The above skeleton is sufficient for the winning condition

[$\text{Büchi}(A) \land \text{Büchi}(B)$]
Example

Game arena $\mathcal{A}$:

Product game $\mathcal{A} \times \mathcal{M}$:

- $(s_1, m_1) \mapsto (s_1, s_2)$
- $(s_1, m_2) \mapsto (s_1, s_1)$
- $(s_2, m_1) \mapsto (s_2, s_2)$
- $(s_2, m_2) \mapsto (s_2, s_1)$

• One can however not apply the [GZ05] result to product games!
Memory-dependent monotony and selectivity

Let $\preceq$ be a preference relation and $\mathcal{M}$ a memory skeleton.

It is said:

- $\mathcal{M}$-monotone whenever

- $\mathcal{M}$-selective whenever

We look at how $\mathcal{M}$ classifies prefixes and cycles.
### Formal definitions of \( \mathcal{M} \)-monotony and \( \mathcal{M} \)-selectivity

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

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

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

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

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

\[
[w(K_1 \cup K_2)^* K_3] \sqsubseteq [wK_1^*] \cup [wK_2^*] \cup [wK_3].
\]
Our characterization for $\mathcal{M}$-determinacy

Characterization - Two-player games

The two following assertions are equivalent:
1. All finite games have optimal $\mathcal{M}$-strategies for both players
2. Both $\Xi$ and $\Xi^{-1}$ are $\mathcal{M}$-monotone and $\mathcal{M}$-selective

Characterization - One-player games

The two following assertions are equivalent:
1. All finite $P_1$-games have (uniform) optimal $\mathcal{M}$-strategies
2. $\Xi$ is $\mathcal{M}$-monotone and $\mathcal{M}$-selective

$\rightarrow$ We recover [GZ05] with $\mathcal{M} = \mathcal{M}_{\text{triv}}$
Applications

Transfer/Lifting theorem

- If in all finite one-player game for player $P_i$, $P_i$ has optimal $\mathcal{M}_i$-strategies, then both players have optimal $\mathcal{M}_1 \times \mathcal{M}_2$-strategies in all finite two-player games.

Very powerful and extremely useful in practice!

Subclasses of games

- If both $\Xi$ and $\Xi^{-1}$ are $\mathcal{M}$-monotone and $\mathcal{M}$-selective, then both players have optimal memoryless strategies in all $\mathcal{M}$-covered games.
Memory-covered arenas

If the game has enough information from $M$, then memoryless strategies will be sufficient.

Covered arenas = same properties as product arenas
Example of application

\[ \exists \text{ defined by a conjunction of reachability } \text{Reach}(\bullet) \land \text{Reach}(\circ) \]

\[ M_1 \]
\[ C \setminus \{\bullet\} \quad m_1 \rightarrow m_2 \quad C \]

\[ \exists \text{ is } M_1\text{-monotone, but not } M_1\text{-selective} \]

\[ M_2 \]
\[ C \setminus \{\bullet, \circ\} \quad m_3 \quad m_1 \rightarrow m_2 \rightarrow m_3 \rightarrow m_2 \rightarrow m_1 \quad C \setminus \{\circ\} \]

\[ \exists \text{ is } M_2\text{-selective} \]

\[ \exists \text{ is } M_1\text{-monotone and } M_2\text{-selective} \]

\[ \exists^{-1} \text{ is } M_1\text{-monotone and } M_{\text{triv}}\text{-selective} \]

\[ \Rightarrow \text{Memory } M_2 \text{ is sufficient for both players!!} \]
Conclusion

A generalization of [GZ05]

- To arena-independent finite memory
- Applies to generalized reachability or parity, lower- and upper-bounded (multi-dimension) energy games

Limitations

- Does only capture arena-independent finite memory
- Hard to generalize (remember counter-example)
- Does not apply to multi-dim. MP, MP+parity, energy+MP (infinite memory)
Conclusion

Other approaches

• Sufficient conditions giving half-memory management results
• Compositionality w.r.t. objectives [LPR18]

Further work

• Understand the arena-dependent framework
• Infinite arenas
• Probabilistic setting
• Other concepts (Nash equilibria)