Arena-independent Memory Bounds for Nash Equilibria in Reachability Games

James C. A. Main

F.R.S.-FNRS, Belgium and UMONS – Université de Mons

STACS 2024

Talk overview

- We consider turn-based multiplayer games on graphs with reachability and shortest-path objectives.
- We focus on constrained Nash equilibria in these games.
- Traditional constructions for finite-memory constrained Nash equilibria usually yield strategies with a size dependent on the arena.

In this talk

We provide constructions for finite-memory Nash equilibria in shortest-path and reachability games that depend only on the number of players.

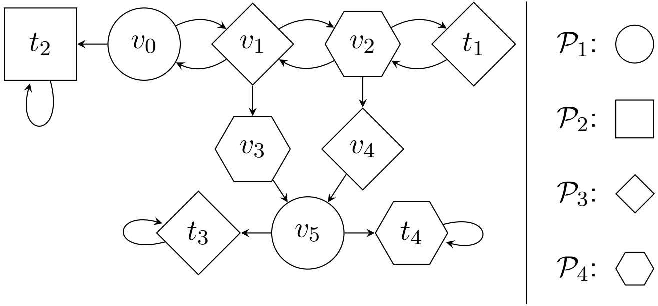
- The constructions presented here apply to infinite arenas.

Table of contents

Table of contents

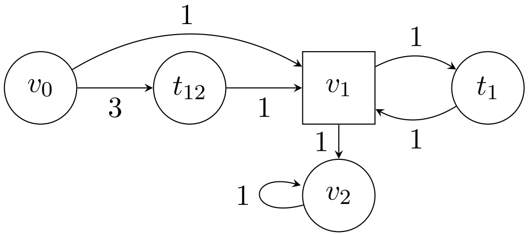
Multiplayer games on graphs

- An arena is a (possibly infinite) graph with vertices partitioned between n players.



- Plays are infinite sequences of vertices consistent with the edges, e.g., $v_0v_1v_2(v_1v_0)^\omega$. A history is a finite prefix of a play.
- In a game, each player has a cost function $\mathsf{cost}_i \colon \mathsf{Plays}(\mathcal{A}) \to \overline{\mathbb{R}}$.

Reachability and shortest-path games



- The reachability cost function is described by a target set $T \subseteq V$.
- A shortest-path cost function is described by a weight function $w\colon E \to \mathbb{N}$ and a target T. For any play $\pi = v_0v_1v_2\ldots$,

$$\mathsf{SPath}_w^T(\pi) = \begin{cases} +\infty & \text{if } T \text{ is not visited in } \pi \\ \sum_{\ell=0}^{r-1} w((v_\ell, v_{\ell+1})) & \text{else, where } r = \min\{r' \mid v_{r'} \in T\} \end{cases}$$

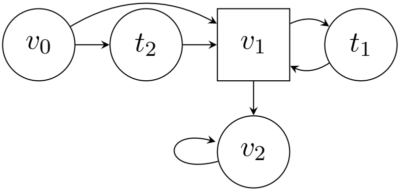
- We omit weight of 1 from illustrations in the following.

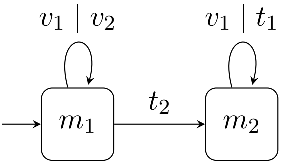
Strategies

- A strategy $\sigma_i \colon V^*V_i \to V$ of $\mathcal{P}_i$ maps a history to a vertex.
- A strategy profile $\sigma = (\sigma_i)_{i \leq n}$ is a tuple with one strategy per player.

Finite-memory strategies

A strategy is finite-memory if it can be encoded by a Mealy machine $(M, m_{\mathsf{init}}, \mathsf{up}, \mathsf{nxt}_i)$ where M is a finite set, $m_{\mathsf{init}} \in M$, $\mathsf{up} \colon M \times V \to M$ is an update function and $\mathsf{nxt} \colon M \times V_i \to V$ is a next-move function.





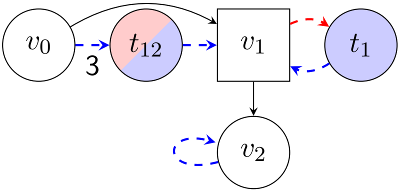
Nash equilibria

Nash equilibrium

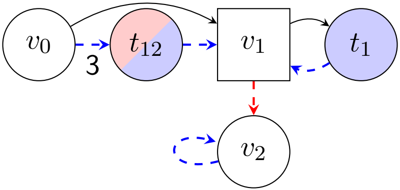
A strategy profile σ is a Nash equilibrium (NE) from $v_0 \in V$ if no player has an incentive to unilaterally deviate from σ, i.e., for all $i \leq n$ and all strategies σ'_i of $\mathcal{P}_i$:

$$\mathsf{cost}_i(\mathsf{Out}(\sigma, v_0)) \leq \mathsf{cost}_i(\mathsf{Out}((\sigma'_i, \sigma_{-i}), v_0)).$$

- Consider the shortest-path game with $T_1 = \{t_{12}, t_1\}$ and $T_2 = \{t_{12}\}$.



Not an NE



Memoryless NE

- Incomparable NE cost profiles may co-exist and some require memory.

Table of contents

Zero-sum games

- We rely on properties of two-player zero-sum games to construct NEs.

Zero-sum game

A two-player game $\mathcal{G} = (\mathcal{A}, (\mathsf{cost}_1, \mathsf{cost}_2))$ is zero-sum if $\mathsf{cost}_1 = -\mathsf{cost}_2$.

- A strategy σ_1 of $\mathcal{P}_1$ ensures $c \in \overline{\mathbb{R}}$ from v if for all strategies σ_2 of $\mathcal{P}_2$, $\mathsf{cost}_1(\mathsf{Out}((\sigma_1, \sigma_2), v)) \leq c$. For $\mathcal{P}_2$, we reverse the inequality.
- The value of v, $\mathsf{val}(v)$, is the infimum cost $\mathcal{P}_1$ can ensure from it.
- A strategy σ_i of $\mathcal{P}_i$ is optimal from v if it ensures $\mathsf{val}(v)$.
- From multi-player games, we derive zero-sum games that oppose one player to the others.

Coalition game

Given a game $\mathcal{G}$ and a player $\mathcal{P}_i$, we let $\mathcal{G}_i$ be the zero-sum game on the same graph where all other players ally against $\mathcal{P}_i$.

Zero-sum reachability and shortest-path games

In a zero-sum reachability game with target T, vertices are either:

- in $W_1(\mathsf{Reach}(T))$, from which $\mathcal{P}_1$ can force a visit to T;
- in $W_2(\mathsf{Safe}(T))$, from which $\mathcal{P}_2$ can avoid T infinitely.

Theorem ([Maz01][1])

In a zero-sum reachability game, both players have uniform optimal (i.e., winning) memoryless strategies.

- In a shortest-path game, $\mathcal{P}_2$ may not have an optimal strategy.

Theorem

In a zero-sum shortest-path game, for all $\alpha \in \mathbb{N}$, there exists a memoryless strategy σ_2^α of $\mathcal{P}_2$ such that, for all $v \in V$:

1. *if $v \in W_2(\mathsf{Safe}(T))$, T cannot be visited from v under σ_2^α;*
2. *σ_2^α ensures a cost of at least $\min\{\mathsf{val}(v), \alpha\}$.*

[1] Mazala, "Infinite Games".

Table of contents

Simplifying NE outcomes

- Not all outcomes can be induced by finite-memory strategy profiles.
- We simplify NE outcomes via a characterisation based on:
  - values of vertices in coalition shortest-path games $\mathcal{G}_i$;
  - winning regions in coalition reachability games.

Lemma (NE outcomes with a simple decomposition)

Let $\rho = v_0v_1v_2\dots$ be an NE outcome in an n-player shortest-path game. There exists an NE outcome π from v_0 that can be decomposed as $h_1 \cdot \ldots \cdot h_k \cdot \pi'$ such that

1. *h_j is a simple history ending in the jth visited target;*
2. *π' is a simple play or of the form hc^ω with hc a simple history;*
3. *for all $j \leq k$, $w(h_j)$ is minimum among histories sharing their first and last vertices with h_j that traverse a subset of the vertices of h_j;*
4. *for all $i \leq n$, $\mathsf{SPath}_w^{T_i}(\pi) \leq \mathsf{SPath}_w^{T_i}(\rho)$.*

Obtaining arena-independent memory bounds

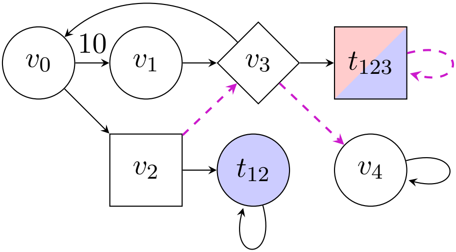
- An outcome with k segments can be achieved by a Mealy machine with k states.
- We build on these Mealy machines and include information to track deviations.
- We punish players with memoryless strategies when they deviate from the intended outcome to obtain NEs.

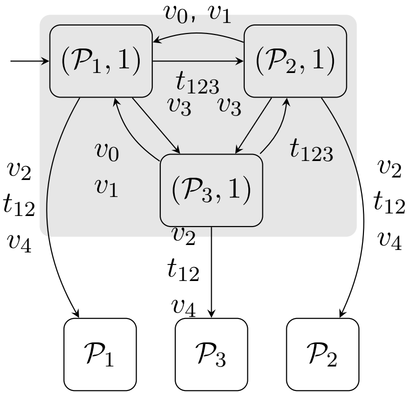
When to punish?

The key is to not punish all deviations: we tolerate deviations that do not exit the current segment.

An example with a single segment

- Let $T_1 = T_2 = \{t_{12}, t_{123}\}$ and $T_3 = \{t_{123}\}$.
- Considered NE outcome $v_0v_1v_3t_{123}^{\omega}$: focus on $v_0v_1v_3t_{123}$.





Punishing strategy against $\mathcal{P}_1$

General result

Theorem (shortest-path games)

For all NE outcomes π from v_0 in a shortest-path game, there exists a finite-memory NE σ from v_0 with strategies of memory at most $n^2 + 2n$ such that $\mathsf{SPath}_w^{T_i}(\mathsf{Out}(\sigma, v_0)) \leq \mathsf{SPath}_w^{T_i}(\pi)$ for all $i \leq n$.

In reachability games, we can refine the memory bounds.

Theorem (Reachability games)

For all NE outcomes π from v_0 in a reachability game, there exists a finite-memory NE σ with strategies of memory at most n^2 such that the same targets are visited in π and in $\mathsf{Out}(\sigma, v_0)$.

Thank you for your attention.

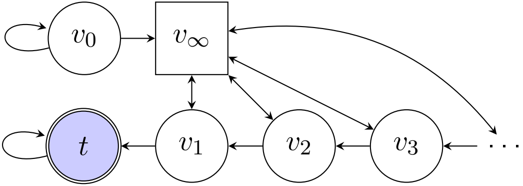
References I

Brihaye, Thomas et al. "On relevant equilibria in reachability games". In: *J. Comput. Syst. Sci.* 119 (2021), pp. 211–230. DOI: 10.1016/j.jcss.2021.02.009. URL: https://doi.org/10.1016/j.jcss.2021.02.009.

Mazala, René. "Infinite Games". In: *Automata, Logics, and Infinite Games: A Guide to Current Research [outcome of a Dagstuhl seminar, February 2001]*. Ed. by Erich Grädel, Wolfgang Thomas, and Thomas Wilke. Vol. 2500. Lecture Notes in Computer Science. Springer, 2001, pp. 23–42. DOI: 10.1007/3-540-36387-4_2. URL: https://doi.org/10.1007/3-540-36387-4_2.

No optimal strategy in zero-sum shortest-path games

- In a zero-sum shortest-path game, $\mathcal{P}_1$ has a memoryless uniform optimal strategy.
- However, $\mathcal{P}_2$ does not have an optimal strategy in general.



- In this game, $\mathsf{val}(v_j) = j$ for all $j \in \mathbb{N}_0 \cup \{\infty\}$.
- However, no matter the strategy of $\mathcal{P}_2$ from v_∞, $\mathcal{P}_2$ cannot prevent a visit to t.

Obtaining finite-memory NEs

- Finite-memory strategy profiles have ultimately periodic outcomes in finite arenas.
- We therefore have to simplify NE outcomes for them to result from a finite-memory strategy profile.

How do we proceed?

1. We rely on a characterisation of plays that can result from NEs.
2. We use the characterisation to derive from any outcome another that results from a finite-memory NE.

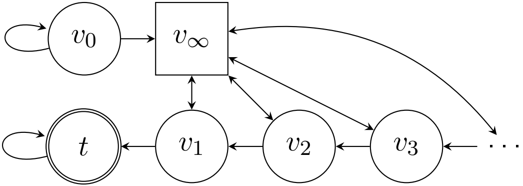
When does a play result from an NE?

- In finite arenas, we have the following characterisation of NE outcomes in shortest-path games [BBGT21][2]:

Theorem ([BBGT21])

Let $\pi = v_0 v_1 \dots$ be a play and let $(T_i)_{i=1}^n$ be the targets. Then π is an outcome of an NE from v_0 in $\mathcal{G}$ if and only for all $1 \leq i \leq n$, $\ell \leq r_i$, it holds that $\mathsf{SPath}_w^{T_i}(\pi_{\geq \ell}) \leq \mathsf{val}_i(v_\ell)$ where $r_i = \inf\{r \in \mathbb{N} \mid v_r \in T_i\}$.

- However, it does not hold as is in infinite arenas.
- Counterexample: the play v_0^ω, assuming $T_1 = \{t\}$, $T_2 = V$.



[2] Brihaye et al., "On relevant equilibria in reachability games".

Characterising NE outcomes

- In infinite games, we must consider the winning regions in the reachability game.

Theorem

Let $\pi = v_0 v_1 \dots$ be a play and let $(T_i)_{i=1}^n$ be the targets. Then π is the outcome of an NE from v_0 in $\mathcal{G}$ iff for all $1 \leq i \leq n$ and $\ell \in \mathbb{N}$, we have

1. *if T_i does not occur in π, then $v_\ell \notin W_i(\mathsf{Reach}(T_i))$ and*
2. *if T_i occurs in π, then $\ell \leq r_i$, implies that $\mathsf{SPath}_w^{T_i}(\pi_{\geq \ell}) \leq \mathsf{val}_i(v_\ell)$ where $r_i = \min\{r \in \mathbb{N} \mid v_r \in T_i\}$.*

Proof idea. ($\Longleftarrow$) We construct an NE $(\sigma_i)_{i=1}^n$ from v_0 by letting for all $1 \leq i \leq n$:

- if h is a prefix $v_0 \dots v_k$ of π, $\sigma_i(h) = v_{k+1}$ and
- otherwise, if h is not a prefix of π and $\mathcal{P}_j$ is responsible for deviating, let $\sigma_i(h) = \sigma_{-j}(\mathsf{last}(h))$ for some $\mathcal{P}_j$-punishing memoryless strategy.

□

Simplifying NE outcomes

Lemma

Let $\rho = v_0 v_1 v_2 \dots$ be an NE outcome in an n-player shortest-path game. There exists an NE outcome π from v_0 that can be decomposed as $h_1 \cdot \ldots \cdot h_k \cdot \pi'$ such that

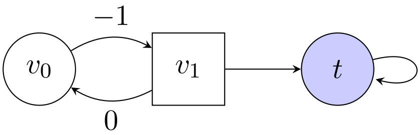
1. *h_j is a simple history ending in the jth visited target;*
2. *π' is a simple play or of the form hc^ω with hc a simple history;*
3. *for all $j \leq k$, $w(h_j)$ is minimum among histories sharing their first and last vertices with h_j that traverse a subset of the vertices of h_j;*
4. *for all $i \leq n$, $\mathsf{SPath}_w^{T_i}(\pi) \leq \mathsf{SPath}_w^{T_i}(\rho)$.*

Proof idea. We apply the following steps.

- Decompose ρ similarly to condition 1, replace each obtained history by a simple one of minimal weight.
- Change the suffix to loop in the first cycle along it if applicable.
- The resulting π is an NE outcome by the characterisation.

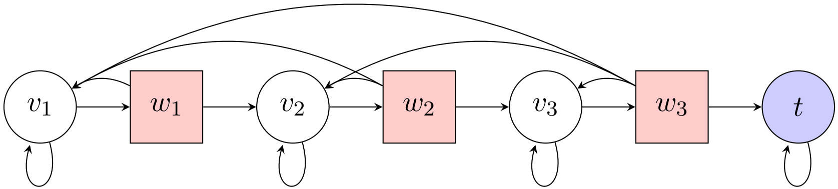
Negative weights

- In a setting with negative weights, in the presence of a negative cycle, there can be NE cost profiles that require an arbitrarily large memory size.
- If $T_1 = \{t\}$ and $T_2 = V$, for all $n \in \mathbb{N}_0$, the play $(v_0 v_1)^n t^\omega$ is an NE outcome that requires a memory of size n and gives a cost of $-n$ for $\mathcal{P}_1$.



Beyond reachability games

- A Büchi objective for $T \subseteq V$ requires that T is visited infinitely often.
- It is not possible to obtain arena-independent memory bounds for Büchi objectives, e.g., below with $T_1 = \{t\}$ and $T_2 = \{w_1, w_2, w_3\}$.
- Below, 3 memory states are needed and it generalises for all $k \geq 1$.



Theorem

In any Büchi game, from any NE we can build a finite-memory NE such that the same objectives are satisfied, even in infinite arenas.