The Many Faces of Strategy Complexity

James C. A. Main[1]
Based on joint work with
Michal Ajdarów[2] Petr Novotný[2] Mickaël Randour[1]

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

[2]Masaryk University, Brno, Czech Republic

March 5, 2025

Talk overview

Strategies are at the center of game-theoretic approaches to reactive synthesis.

Goal of this talk

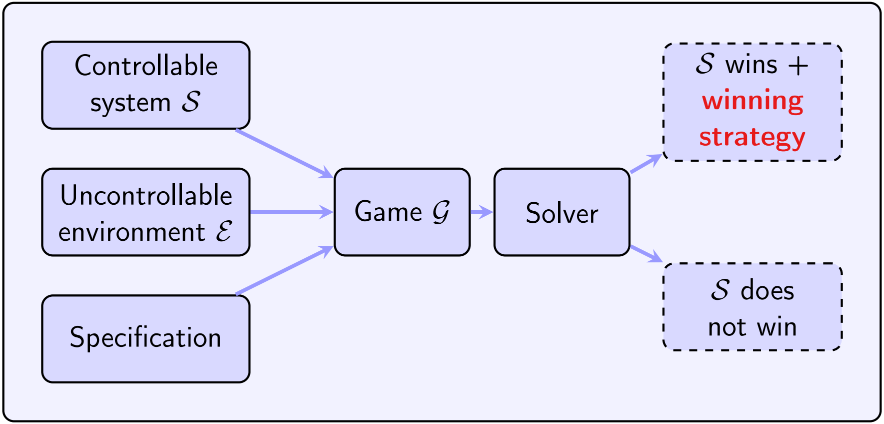
Motivate and explain a **multifaceted vision** of strategy complexity.

In the second part of this talk, we will focus on:

- **randomised** strategies;
- **alternative** representations of strategies.

Table of contents

Reactive synthesis through game theory

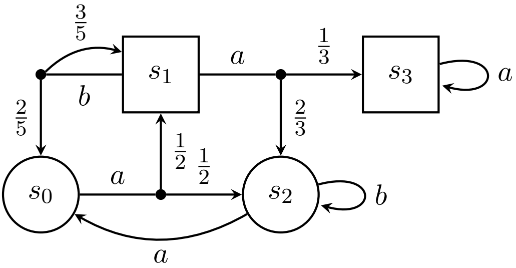


A **strategy** is a formal blueprint for a **controller of the system**.

Turn-based stochastic games

Turn-based stochastic game $\mathcal{G}$

- **Finite or countable** state space $S = S_1 \uplus S_1$.
- **Finite** action space A.
- **Randomised** transition function $\delta\colon S \times A \to \mathcal{D}(S)$.



Plays are sequences in $(SA)^\omega$ coherent with transitions.
$\rightsquigarrow$ **Example**: $s_0 a s_1 b s_1 \ldots$

Strategies

A **history** is a prefix h of a play ending in a **state**.

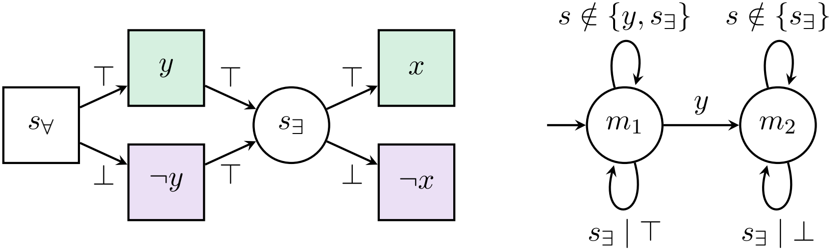
Strategy

A **behavioural strategy** of $\mathcal{P}_i$ is a function $\sigma_i \colon \mathsf{Hist}_i(\mathcal{G}) \to \mathcal{D}(A^{(i)})$.

- Two strategies σ_1, σ_2 and an initial state $s \rightsquigarrow$ **distribution** $\mathbb{P}_s^{\sigma_1,\sigma_2}$ over plays.
- A strategy σ_i is **pure** if $\sigma_i \colon \mathsf{Hist}_i(\mathcal{G}) \to A$.
- A strategy is **memoryless** if $\sigma_i \colon S_i \to \mathcal{D}(A)$.

Table of contents

Strategies and memory



Representation of strategies via Mealy machines with randomisation

- Set of **memory states** M;
- initial **memory distribution** μ_{init};
- **next-move** function $\mathsf{nxt}_{\mathcal{M}} \colon M \times S_i \to \mathcal{D}(A)$;
- memory **update** function $\mathsf{up}_{\mathcal{M}} \colon M \times S \times A \to \mathcal{D}(M)$.

Strategy complexity via memory

- The complexity of strategies is traditionally measured by the size of their **memory**.
- Memory requirements for optimal strategies in games have been thoroughly studied.

A glimpse into known results on memory

- Characterisations and one-to-two player lifts (e.g., [GZ05; Bou+22]).
- Refining memory bounds/computing optimal bounds (e.g., [Bou+23; Mai24]).
- Trading memory for **randomisation** (e.g., [CAH04; CRR14]).

Strategy complexity in general

- Memory size does **not fully describe** the complexity of strategies.
- Other aspects also play a role in the complexity of strategies.
- **Major question**: what makes a strategy complex?

Our vision

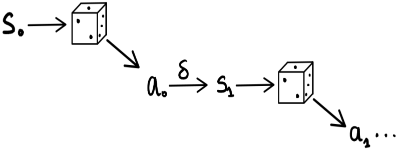
Strategy complexity is **multifaceted**: various factors contribute to the complexity of a strategy.

- **Next step**: a brief look into **randomisation**.

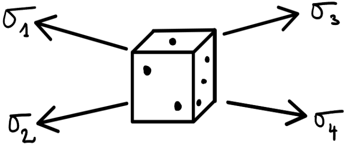
Table of contents

Mixed and behavioural strategies

- There exist different definitions of **randomised strategies**.



Behavioural strategies



Mixed strategies

- In general, these two classes of strategies are **not comparable**.
- Kuhn's theorem [Aum64]: in **games of perfect recall** any mixed strategy has an equivalent behavioural strategy and vice-versa.

What happens with finite-memory strategies?

Are all models of **finite-memory randomised strategies equivalent**?

Randomisation and finite memory [MR24]

A class of Mealy machines is denoted by **XYZ** where X, Y, Z $\in$ {D, R} where D stands for deterministic and R for random, and

- X characterises initialisation,
- Y characterises the next-move function,
- Z characterises updates.

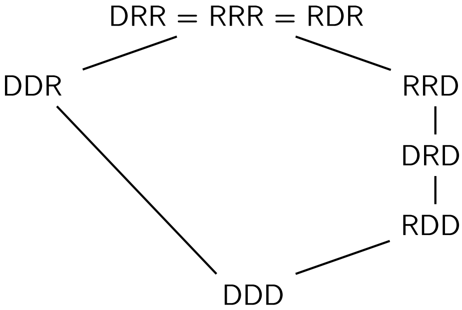


Table of contents

Random strategies and multiple objectives

- We study one-player games, i.e., **Markov decision processes**, with **multiple payoffs**.
- In general, the satisfaction of multi-objective queries requires **randomised strategies**.

Main questions

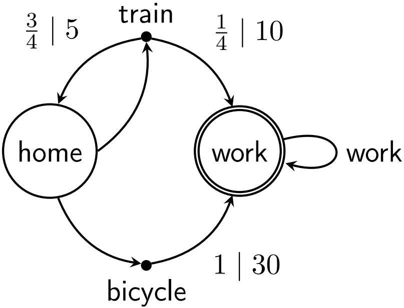
- What is the relationship between expected payoffs of **pure strategies** and expected payoffs of **general strategies**?
- What **type of randomisation** do we need for multi-objective queries?

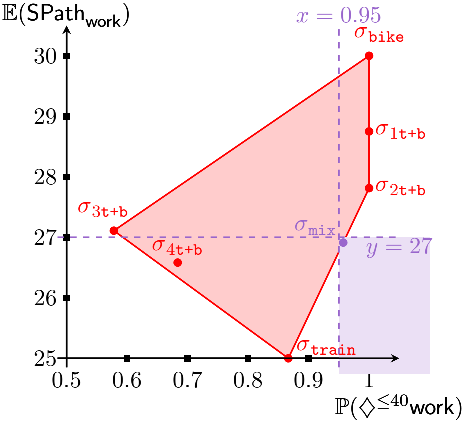
→ **Goal**: results for the **broadest possible class of payoffs**.

Multi-objective Markov decision processes

We consider **two goals**:

- reaching work under 40 minutes with **high probability**;
- minimising the **expected** time to reach work.





Payoffs

- A **payoff** is a measurable function $f\colon \mathsf{Plays}(\mathcal{M}) \to \bar{\mathbb{R}}$.
- We let $\mathbb{E}_s^\sigma(f) = \int_{\pi \in \mathsf{Plays}(\mathcal{M})} f(\pi) \mathrm{d}\mathbb{P}_s^\sigma(\pi)$.

Which payoffs f are relevant?

- f is **good** if $\mathbb{E}_s^\sigma(f)$ is well-defined for all strategies σ and all $s \in S$.
- f is **universally integrable** payoffs: $\mathbb{E}_s^\sigma(|f|) \in \mathbb{R}$ if for all strategies σ and all $s \in S$.

For a **multi-dimensional payoff** $\bar{f} = (f_1, \ldots, f_d)$ and $s \in S$, we let:

- $\mathsf{Pay}_s(\bar{f}) = \{\mathbb{E}_s^\sigma(\bar{f}) \mid \sigma \text{ strategy}\}$;
- $\mathsf{Pay}_s^{\mathsf{pure}}(\bar{f}) = \{\mathbb{E}_s^\sigma(\bar{f}) \mid \sigma \text{ pure strategy}\}$.

Universally integrable payoffs

In the introductory example, we had $\mathsf{Pay}_{\mathsf{home}}(\bar{f}) = \mathrm{conv}(\mathsf{Pay}^{\mathsf{pure}}_{\mathsf{home}}(\bar{f}))$.

When does this generalise?

Theorem ((M., Randour))

*Let $\bar{f} = (f_1, \ldots, f_d)$ be **universally integrable**. Then, for all states s,*

$$\mathsf{Pay}_s(\bar{f}) = \mathrm{conv}(\mathsf{Pay}^{\mathsf{pure}}_s(\bar{f})).$$

In particular, to match the expected payoff of any strategy, it suffices to:

- ***mix $d+1$ pure strategies**;*
- *consider strategies use **randomisation at most d** along any play.*

Sequel: proof of a weaker result

If $\bar{f}$ is universally integrable, then $\mathrm{cl}(\mathsf{Pay}_s(\bar{f})) = \mathrm{cl}(\mathrm{conv}(\mathsf{Pay}^{\mathsf{pure}}_s(\bar{f})))$.

Universally integrable payoffs

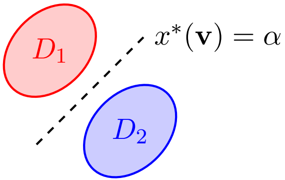
A simpler proof

Non-direct inclusion: $\mathsf{Pay}_s(\bar{f}) \subseteq \mathrm{cl}(\mathrm{conv}(\mathsf{Pay}_s^{\mathsf{pure}}(\bar{f})))$.
Let σ be a strategy and $\mathbf{q} = \mathbb{E}_s^\sigma(\bar{f})$. Assume $\mathbf{q} \notin \mathrm{cl}(\mathrm{conv}(\mathsf{Pay}_s^{\mathsf{pure}}(\bar{f})))$.

Main idea: reduction to a **one-dimensional** payoff.

Theorem (Hyperplane separation theorem)

*Let D_1, $D_2 \subseteq \mathbb{R}^d$ be **disjoint convex** sets. If D_1 is **closed** and D_2 is **compact**, then there exists a **linear form** $x^* : \mathbb{R}^d \to \mathbb{R}$ and $\varepsilon > 0$ such that for all $\mathbf{p}_1 \in D_1$ and $\mathbf{p}_2 \in D_2$, $x^*(\mathbf{p}_1) + \varepsilon < x^*(\mathbf{p}_2)$.*



Universally integrable payoffs

A simpler proof

Non-direct inclusion: $\mathsf{Pay}_s(\bar{f}) \subseteq \mathrm{cl}(\mathrm{conv}(\mathsf{Pay}_s^{\mathsf{pure}}(\bar{f})))$.

Let σ be a strategy and $\mathbf{q} = \mathbb{E}_s^\sigma(\bar{f})$. Assume $\mathbf{q} \notin \mathrm{cl}(\mathrm{conv}(\mathsf{Pay}_s^{\mathsf{pure}}(\bar{f})))$.

Main idea: reduction to a **one-dimensional** payoff.

- There exists a linear form x^* such that, for all **pure strategies** τ,

$$x^*(\mathbb{E}_s^\tau(\bar{f})) < x^*(\mathbf{q})$$

- By linearity, we obtain that for all pure strategies τ,

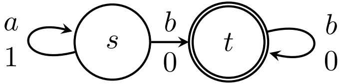
$$\mathbb{E}_s^\tau(x^*(\bar{f})) < \mathbb{E}_s^\sigma(x^*(\bar{f}))$$

Lemma

*Let f be **universally integrable**. For all strategies σ, there exists a **pure** strategy τ such that $\mathbb{E}_s^\sigma(f) \leq \mathbb{E}_s^\tau(f)$.*

Beyond universally integrable payoffs

Example



Payoffs

1. reaching $t \rightsquigarrow f_1 = \mathbb{1}_{\Diamond t}$;
2. sum of weights $\rightsquigarrow f_2 = \sum_{\ell=0}^{\infty} w(c_\ell)$.

- $\mathbb{E}_s^{\sigma_a}(f_2) = +\infty \implies f_2$ is **not universally integrable**.
- $\mathsf{Pay}_s^{\mathsf{pure}}(\bar{f}) = \{(0,+\infty)\} \cup \{(1,\ell) \mid \ell \in \mathbb{N}\}$.
  $\implies \mathrm{conv}(\mathsf{Pay}_s^{\mathsf{pure}}(\bar{f})) = (\{1\} \times \mathbb{R}_{\geq 0}) \cup ([0,1[\times \{+\infty\})$.
- We have $(1,+\infty) \in \mathsf{Pay}_s(\bar{f})$ via σ such that for all $\ell \in \mathbb{N}$:

$$\sigma(s(as)^\ell)(a) = \begin{cases} \frac{1}{2} & \text{if } \ell \in 2^{\mathbb{N}} \\ 1 & \text{if } \ell \notin 2^{\mathbb{N}} \end{cases}$$

$\rightarrow$ **The theorem and the key lemma do not generalise.**

Beyond universally integrable payoffs

Theorem (M., Randour)

*Let $\bar{f}$ be a good payoff and $s \in S$. Let $\mathbf{q} \in \mathsf{Pay}_s(\bar{f})$.
All neighbourhoods of $\mathbf{q}$ (in $\bar{\mathbb{R}}$) intersect $\mathrm{conv}(\mathsf{Pay}_s^{\mathsf{pure}}(\bar{f}))$. In other words, $\mathbf{q}$ can be approximated by* ***finite-support mixed strategies*** *.*

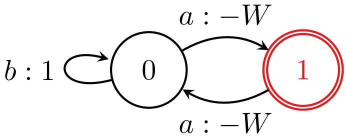
Table of contents

Memory does not tell the whole story (1/2)

Counter-based strategies

Memory and **randomisation** do **not fully reflect** the complexity of a strategy.

- We consider a game with an **energy-Büchi** objective [CD12], where $W \in \mathbb{N}$.

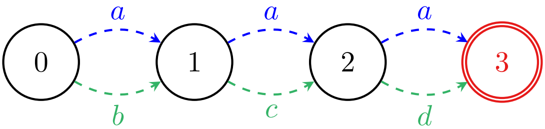


- Need memory **exponential** in the binary encoding of W to satisfy the energy-Büchi objective.
- **Polynomial** representation with a **counter**-based approach.

Memory does not tell the whole story (2/2)

Action choices influence simplicity

Memory and **randomisation** do **not fully reflect** the complexity of a strategy.



$\rightarrow$ Strategy σ_1 is **simpler to represent** than σ_2

- The **action choices** can impact how concise the strategy can be made.

Related challenge

How to represent and analyse **memoryless strategies** when the state space is **infinite**?

Memoryless strategies in one-counter MDPs

- We study **one-counter Markov decision processes**.
- We consider counter-based strategies with a **compact representation** that we call **interval strategies**.

Our contribution (Ajdarów, M., Novotný, Randour)

- PSPACE **verification** algorithms for interval strategies.
- PSPACE **realisability** algorithms for **structurally-constrained** interval strategies.

- Our algorithms are based on a **finite abstraction** of an **infinite system**.

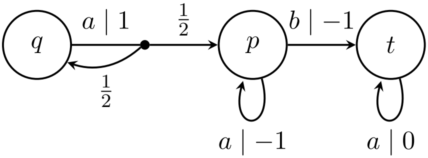
One-counter Markov decision processes

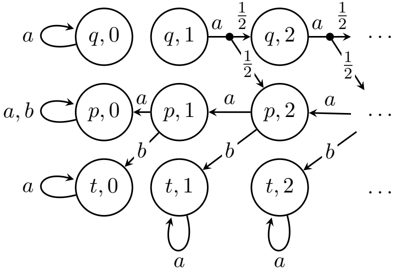
One-counter MDP (OC-MDP) $\mathcal{Q}$

- **Finite** MDP (Q, A, δ).
- **Weight function** $w\colon Q \times A \to \{-1, 0, 1\}$.

MDP $\mathcal{M}^{\leq\infty}(\mathcal{Q})$ induced by $\mathcal{Q}$

- **Countable** MDP over $S = Q \times \mathbb{N}$.
- State transitions via δ.
- Counter updates via w.





Interval strategies

We study a restricted class of **memoryless strategies** of $\mathcal{M}^{\leq\infty}(\mathcal{Q})$.

Open-ended interval strategies (OEIS)

σ is an OEIS if $\exists\, k_0 \in \mathbb{N}$ s.t. $\forall\, q \in Q$ and $\forall\, k \geq k_0$, $\sigma(q, k) = \sigma(q, k_0)$.

$\mathbb{N}_0$	1	2	$\cdots$	$k_0 - 1$	k_0	$k_0 + 1$	$\cdots$
Q	σ_1	σ_2	$\cdots$	σ_{k_0-1}	σ_{k_0}	σ_{k_0}	$\cdots$

constant

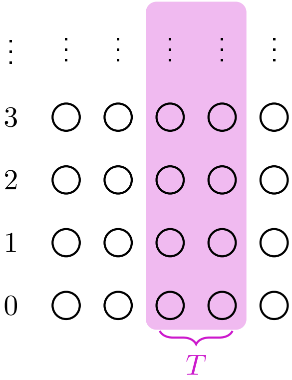
Group counter values in intervals

Inter.	I_1	I_2	$\cdots$	$I_d = [\![k_0, \infty]\!]$
Q	τ_1	τ_2	$\cdots$	$\tau_d = \sigma_{k_0}$

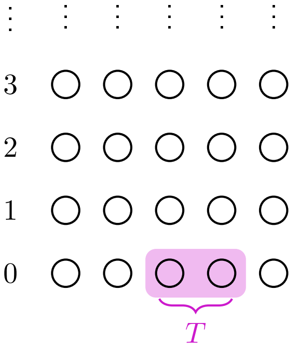
$\rightsquigarrow$ **Finite partition** of $\mathbb{N}_0$ into **intervals**

Objectives

- An **objective** is a measurable set of plays.
- Let $T \subseteq Q$ **be a target**.
- We study variants of **reachability objectives**.



State reachability $\mathsf{Reach}(T)$



Selective termination $\mathsf{Term}(T)$

Interval strategy verification problem

Interval strategy verification problem

Decide whether $\mathbb{P}^{\sigma}_{\mathcal{M}^{\leq\infty}(\mathcal{Q}),s_{\mathsf{init}}}(\Omega) \geq \theta$ given an **OEIS** σ, **an objective** $\Omega \in \{\mathsf{Reach}(T), \mathsf{Term}(T)\}$, a **threshold** $\theta \in \mathbb{Q} \cap [0,1]$ and an **initial configuration** $s_{\mathsf{init}} \in Q \times \mathbb{N}$.

- We construct a finite **compressed Markov chain** $\mathcal{C}^{\sigma}_{\mathcal{I}}$.
- We have formulae (in the signature $\{0, 1, +, -, \cdot, \leq\}$):
  - $\Phi^{\mathcal{I}}_{\delta}(\mathbf{x}, \mathbf{z}^{\sigma})$ for **transition probabilities** of $\mathcal{C}^{\sigma}_{\mathcal{I}}$;
  - $\Phi^{\mathcal{I}}_{\Omega}(\mathbf{x}, \mathbf{y})$ for **termination probabilities** from configurations of $\mathcal{C}^{\sigma}_{\mathcal{I}}$.
- We can solve the verification problem by checking if

$$\mathbb{R} \models \forall\,\mathbf{x}\forall\,\mathbf{y}\,(\Phi^{\mathcal{I}}_{\delta}(\mathbf{x}, \mathbf{z}^{\sigma}) \wedge \Phi^{\mathcal{I}}_{\Omega}(\mathbf{x}, \mathbf{y})) \implies y_{s_{\mathsf{init}}} \geq \theta.$$

Conclusion

Strategy complexity can be analysed through different approaches:

- memory requirements;
- randomisation requirements;
- the existence of small strategy representations.

In a nutshell

We are interested in developing deeper insight on **strategy complexity** and studying **alternative strategy models**.

References I

[Aum64] Robert J . Aumann. “Mixed and Behavior Strategies in Infinite Extensive Games”. In: *Advances in Game Theory. (AM-52), Volume 52*. Ed. by Melvin Dresher, Lloyd S. Shapley, and Albert William Tucker. Princeton University Press, 1964, pp. 627–650. DOI: `doi:10.1515/9781400882014-029`.

[Bou+22] Patricia Bouyer et al. “Games Where You Can Play Optimally with Arena-Independent Finite Memory”. In: *Log. Methods Comput. Sci.* 18.1 (2022). DOI: `10.46298/lmcs-18(1:11)2022`. URL: `https://doi.org/10.46298/lmcs-18(1:11)2022`.

References II

[Bou+23] Patricia Bouyer et al. "How to Play Optimally for Regular Objectives?" In: *50th International Colloquium on Automata, Languages, and Programming, ICALP 2023, July 10-14, 2023, Paderborn, Germany*. Ed. by Kousha Etessami, Uriel Feige, and Gabriele Puppis. Vol. 261. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023, 118:1–118:18. DOI: 10.4230/LIPICS.ICALP.2023.118.

[CAH04] Krishnendu Chatterjee, Luca de Alfaro, and Thomas A. Henzinger. "Trading Memory for Randomness". In: *1st International Conference on Quantitative Evaluation of Systems (QEST 2004), 27-30 September 2004, Enschede, The Netherlands*. IEEE Computer Society, 2004, pp. 206–217. DOI: 10.1109/QEST.2004.1348035.

References III

[CD12] Krishnendu Chatterjee and Laurent Doyen. "Energy parity games". In: *Theor. Comput. Sci.* 458 (2012), pp. 49–60. DOI: 10.1016/J.TCS.2012.07.038. URL: https://doi.org/10.1016/j.tcs.2012.07.038.

[CRR14] Krishnendu Chatterjee, Mickael Randour, and Jean-François Raskin. "Strategy synthesis for multi-dimensional quantitative objectives". In: *Acta Informatica* 51.3-4 (2014), pp. 129–163. DOI: 10.1007/S00236-013-0182-6.

[EWY10] Kousha Etessami, Dominik Wojtczak, and Mihalis Yannakakis. "Quasi-Birth-Death Processes, Tree-Like QBDs, Probabilistic 1-Counter Automata, and Pushdown Systems". In: *Performance Evaluation* 67.9 (2010), pp. 837–857. DOI: 10.1016/J.PEVA.2009.12.009. URL: https://doi.org/10.1016/j.peva.2009.12.009.

References IV

[GZ05] Hugo Gimbert and Wieslaw Zielonka. "Games Where You Can Play Optimally Without Any Memory". In: *CONCUR 2005 - Concurrency Theory, 16th International Conference, CONCUR 2005, San Francisco, CA, USA, August 23-26, 2005, Proceedings*. 2005, pp. 428–442. DOI: 10.1007/11539452_33. URL: https://doi.org/10.1007/11539452_33.

[KEM06] Antonín Kucera, Javier Esparza, and Richard Mayr. "Model Checking Probabilistic Pushdown Automata". In: *Logical Methods in Computer Science* 2.1 (2006). DOI: 10.2168/LMCS-2(1:2)2006. URL: https://doi.org/10.2168/LMCS-2(1:2)2006.

References V

[Mai24] James C. A. Main. "Arena-Independent Memory Bounds for Nash Equilibria in Reachability Games". In: *41st International Symposium on Theoretical Aspects of Computer Science, STACS 2024, March 12-14, 2024, Clermont-Ferrand, France*. Ed. by Olaf Beyersdorff et al. Vol. 289. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2024, 50:1–50:18. DOI: `10.4230/LIPICS.STACS.2024.50`.

[MR24] James C. A. Main and Mickael Randour. "Different strokes in randomised strategies: Revisiting Kuhn's theorem under finite-memory assumptions". In: *Information and Computation* 301 (2024), p. 105229. DOI: `10.1016/J.IC.2024.105229`. URL: `https://doi.org/10.1016/j.ic.2024.105229`.

References VI

[Tiw92] Prasoon Tiwari. "A problem that is easier to solve on the unit-cost algebraic RAM". In: *Journal of Complexity* 8.4 (1992), pp. 393–397. DOI: 10.1016/0885-064X(92)90003-T. URL: https://doi.org/10.1016/0885-064X(92)90003-T.