The Many Faces of Strategy Complexity

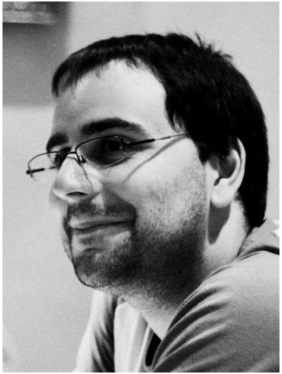
James C. A. Main
Thesis realised under the supervision of Mickael Randour

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

UMONS fnrs LA LIBERTÉ DE CHERCHER

Public PhD Defence – September 23, 2025

One does not journey alone



The wonderful
Mickael Randour

My **co-authors**, to whom I am grateful.

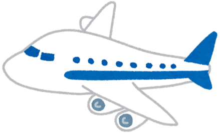
- Jeremy Sproston
- Michal Ajdarów
- Petr Novotný
- Thomas Brihaye
- Aline Goeminne

Table of contents

Motivations: a world of computing

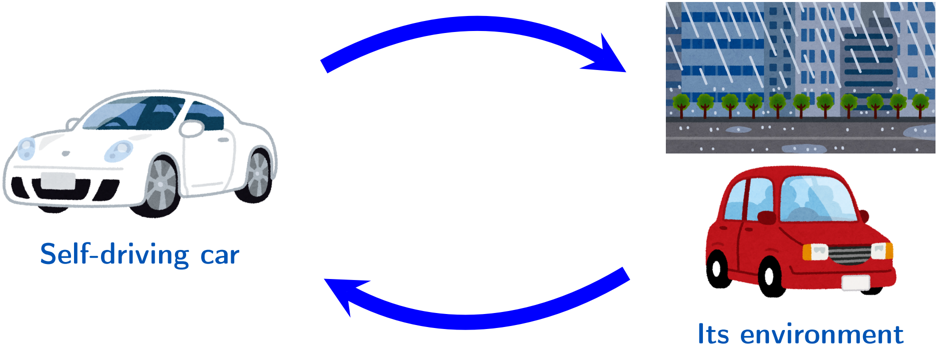
We are surrounded by **computer systems**.

Bugs should not occur in **safety-critical** systems.



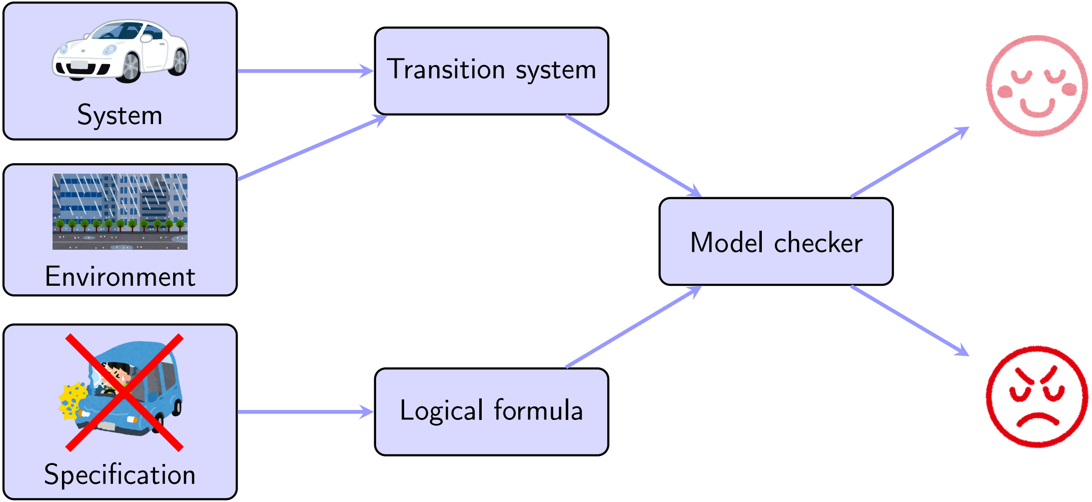
Reactive systems

A **reactive system** is a system that constantly interacts with its **environment**.

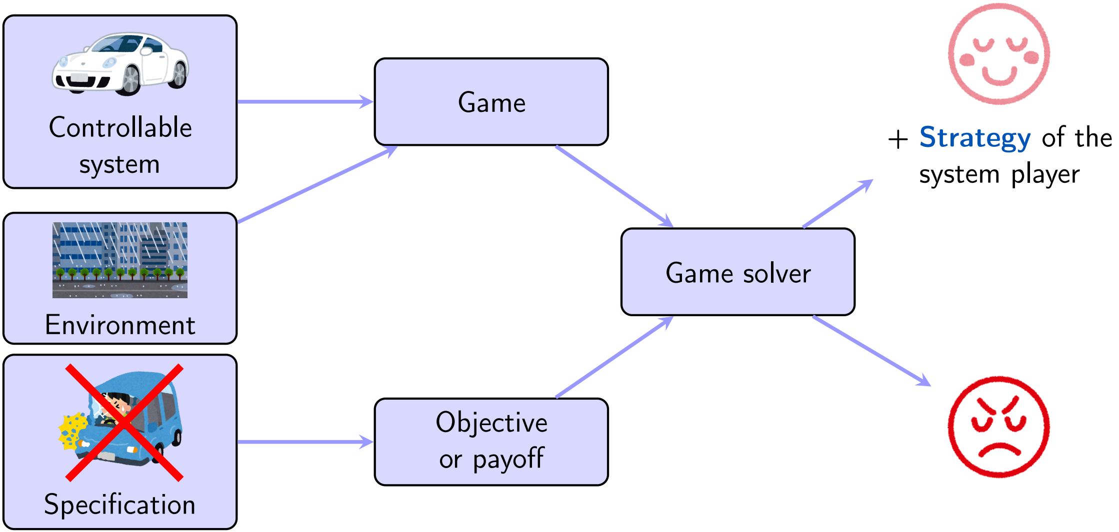


→ **Testing does not guarantee the absence of bugs.**

Model checking



Reactive synthesis



What is a game?

> A **game** is a mathematical model of the **interaction** between entities called **players**.

There exist **many game variants**:

- one-shot games or **sequential games**;
- deterministic or with **randomness**;
- with **perfect** or imperfect information.

We focus on **Markov decision processes**: **one player versus randomness**.

Table of contents

Markov decision processes

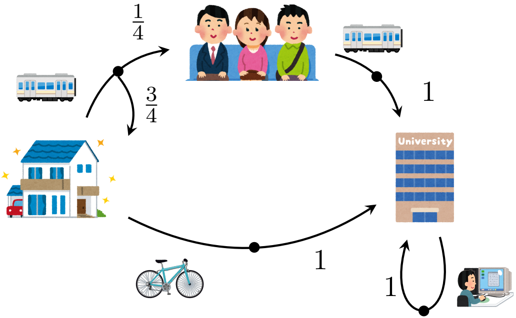
A **Markov decision process** (or **MDP**) models the interaction of a **player** with a **stochastic environment**.
It is described by

- a set of **states** S,
- a set of **actions** A and
- a **randomised transition** function $\delta\colon S \times A \to \mathcal{D}(S)$.

A **play** of an MDP is an **infinite path** along transitions.

Ex.  . . .

MDP example: commuting to work

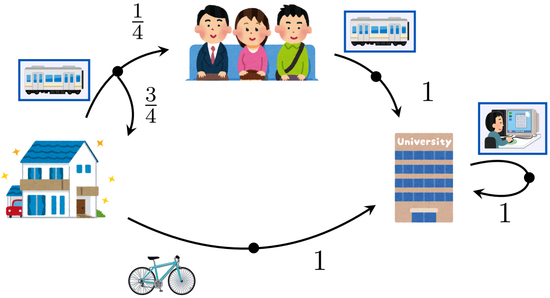


Strategies

A **strategy** describes the decisions to be made **in all scenarios**.

Mathematically, a strategy is a **function** $\sigma\colon (SA)^*S \to A$.

Once an initial state and a strategy are chosen, we obtain a **stochastic process** known as a **Markov chain**.



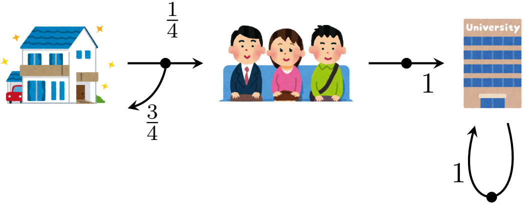
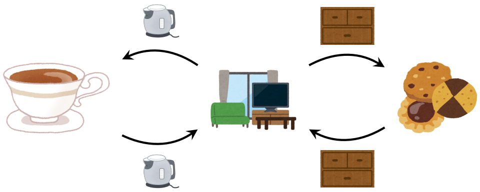


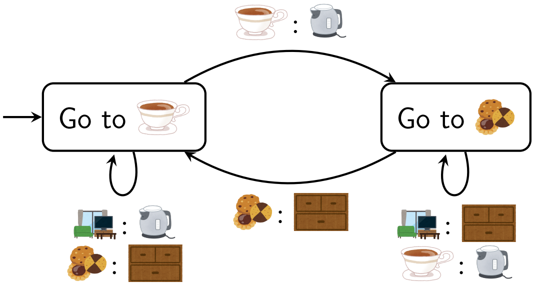
Table of contents

Memory in strategies

Memoryless strategies are simpler strategies that **make decisions based only on the current state**, i.e., they disregard the past.

Some goals may require **memory** to be satisfied.

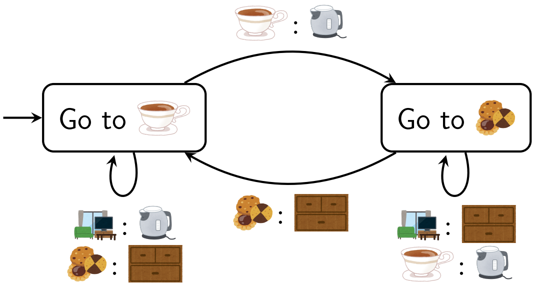




Finite-memory strategies

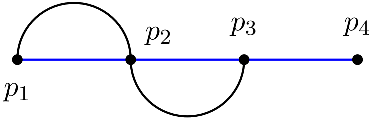
A strategy has finite memory if it can be encoded by a Mealy machine (i.e., a finite automaton with outputs).

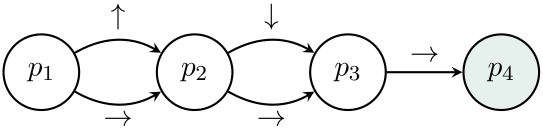
The memory of a strategy provides a measure of its complexity.



The complexity of strategies

What makes a strategy complex?





→ All **memoryless** strategies lead to the target, but the **constant one** is simplest.

My contribution in a nutshell: studying the **complexity of strategies** via different angles.

Table of contents

The need for richer strategies

We aim for a good performance in the **worst case**.

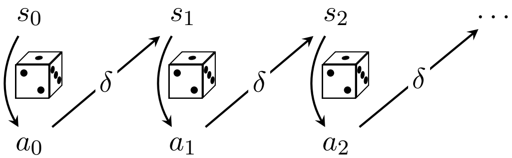
Some applications require **richer strategies**.

Simple example: **rock paper scissors**.

- ▷ If I choose **rock**, my opponent will play **paper**.
- ▷ If I choose **paper**, my opponent will play **scissors**.
- ▷ If I choose **scissors**, my opponent will play **rock**.

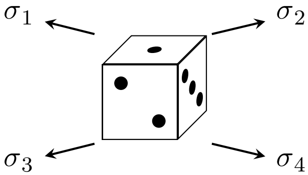
By using **randomisation**, I can improve my chances.

What is a randomised strategy?



Behavioural strategy

$\sigma\colon (SA)^*S \to \mathcal{D}(A)$



Mixed strategy

$\mathcal{D}(\sigma\colon (SA)^*S \to A)$

Kuhn's theorem implies that mixed and behavioural strategies have the **same expressiveness** in MDPs and games with perfect information.

Memory and randomness

How can we implement **randomisation** in **Mealy machines**?

Idea 1: mixing Initial memory distribution	<	**Idea 2: random actions choices** Randomised outputs	<	**Idea 3: randomised memory** Randomised updates

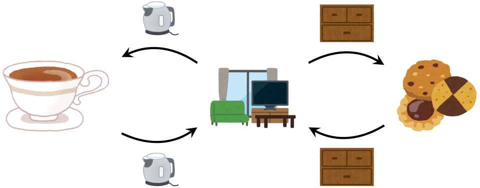
In (M., Randour, Inf. Comp. 2024), we have compared the **expressiveness** of all variants of **stochastic Mealy machines** and provided a **full classification**.

Memory and randomness

What is the relationship between **memory** and **randomisation**?

To visit  and  infinitely often, we can:

- ▷ use a pure strategy with **two memory states**;
- ▷ **toss a coin** to select an action in all rounds ⇝ **memoryless strategy**.



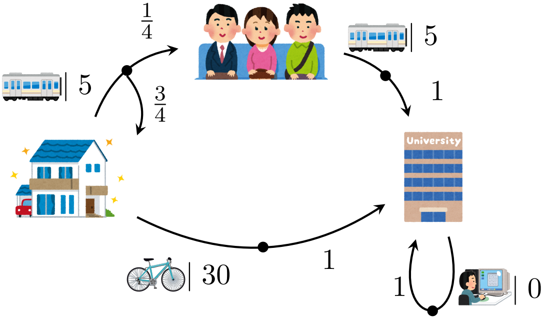
There can be **trade-offs** between **memory** and **randomness**.

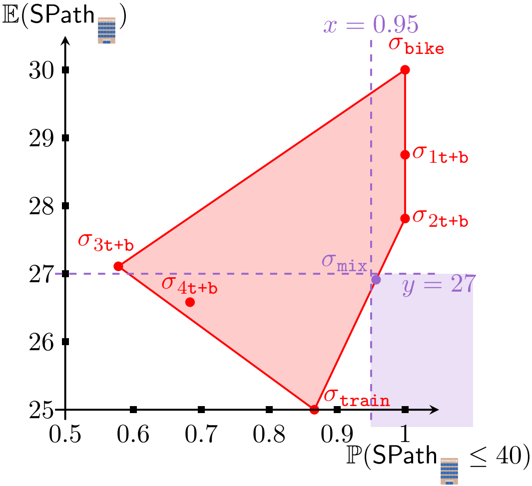
Table of contents

Example: optimising your commute

Optimising two goals:

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



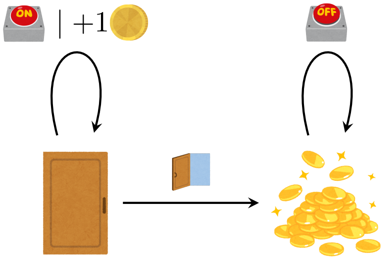


Another example

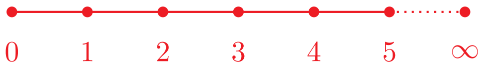
Optimising the number of coins obtained:

- add coins by pressing a button in a closed room;
- collect them by exiting the room.
- no more coins can be obtained if the door is opened.

> Need more complex randomisation.



Possible payoffs:



Randomisation requirements

General framework: **Markov decision processes** with **multiple payoff functions**.

> **Payoff functions**: $f\colon \mathsf{Plays}(\mathcal{M}) \to \bar{\mathbb{R}}$; quantify the quality of plays.

What can we say about **randomisation requirements** in multi-objective MDPs?

> **Theorem (M., Randour, 2025).** If all payoffs $f_1, \dots, f_d$ have **finite expectations under all strategies**:
> - mixing at most $d + 1$ **many pure strategies** is sufficient to **obtain** any expected payoff vector;
> - $\mathsf{Pay}_s(\bar{f}) = \mathsf{conv}(\mathsf{Pay}_s^{\mathsf{pure}}(\bar{f}))$.

Randomisation requirements

General framework: **Markov decision processes** with **multiple payoff functions**.

Payoff functions: $f\colon \mathsf{Plays}(\mathcal{M}) \to \bar{\mathbb{R}}$; quantify the quality of plays.

What can we say about **randomisation requirements** in multi-objective MDPs?

Theorem (M., Randour, 2025). If all payoffs $f_1, \ldots, f_d$ have a **well-defined expectations under all strategies**:

- mixing at most $d+1$ **many pure strategies** is sufficient to **approximate** any expected payoff vector;
- $\mathsf{Pay}_s(\bar{f}) \subseteq \mathsf{cl}(\mathsf{conv}(\mathsf{Pay}_s^{\mathsf{pure}}(\bar{f})))$.

Table of contents

Memory and randomness do not tell the whole story

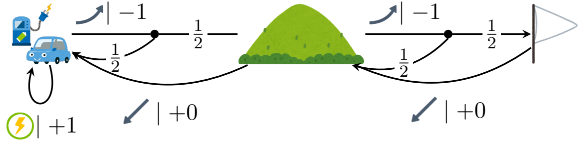
There is **more to strategy complexity** than only memory and randomness.

Some **strategies** can admit **small representations**.

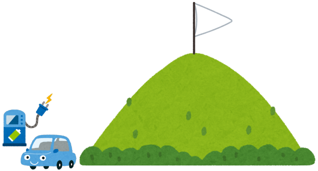
We focus on a setting with **counters**.

One-counter MDPs

A **one-counter MDP** (OC-MDP) is an MDP with weights in $\{-1, 0, 1\}$ on its transitions.

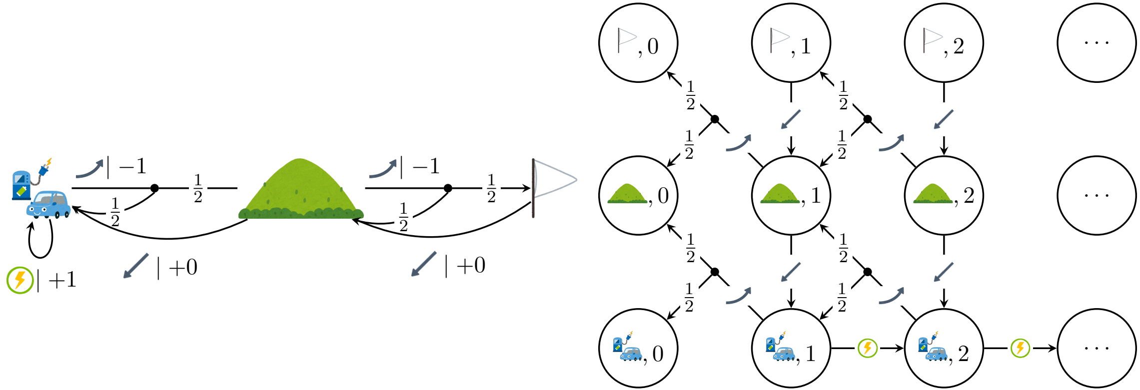


Example: going up a slippery hill.



One-counter MDPs

A **one-counter MDP** (OC-MDP) is an MDP with weights in $\{-1, 0, 1\}$ on its transitions.



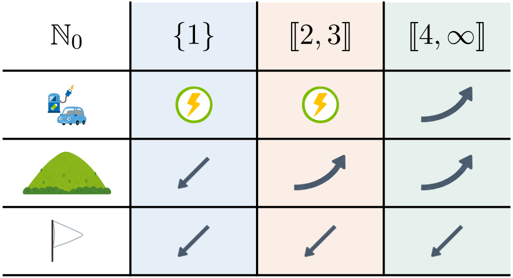
Strategies over configurations

A **memoryless strategy** over configurations can be seen as an **infinite table**.

$\mathbb{N}_0$	1	2	3	4	5	$\cdots$
	⚡	⚡	⚡	↗	↗	$\cdots$
	↙	↗	↗	↗	↗	$\cdots$
	↙	↙	↙	↙	↙	$\cdots$

Strategies over configurations

A **memoryless strategy** over configurations can be seen as an **infinite table**.



We need **specific representations** for this context.

Interval strategies

An **interval strategy** is a strategy that can be described by a **finite interval partition** of $\mathbb{N}_{>0}$ and **memoryless strategies** for each interval.

$\mathbb{N}_0$	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 →

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

Verification of interval strategies

Verification problem. When following a given interval strategy, do we **reach a target state** with probability greater than or equal to some given threshold?

Challenges

- **Infinite** Markov chain.
- Compressed Markov chains have **irrational** or **very precise** probabilities.

Solutions

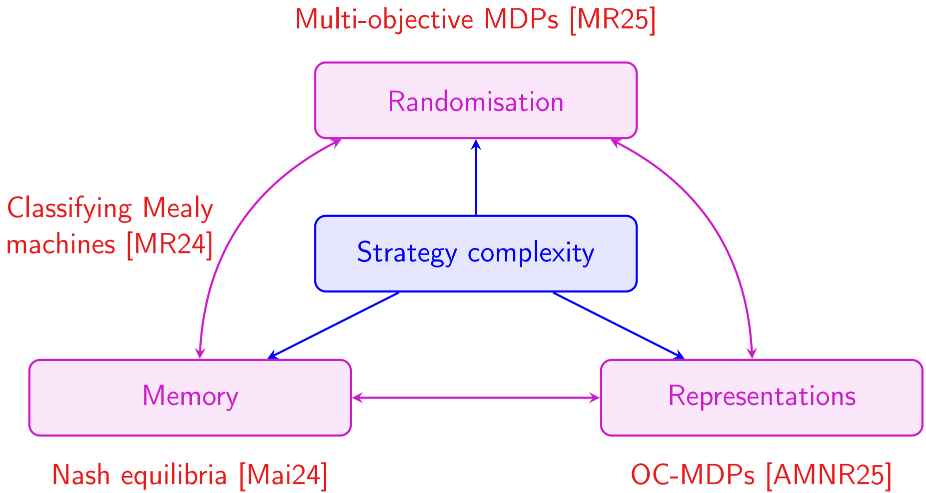
- **Compression** to finite Markov chain.
- Transition probabilities can be represented by small **logical formulae**.

Algorithm (Ajdarów, M., Novotný, Randour). Construct a **universal logical formula** and check if it is satisfied in the **theory of the reals**.

We have also built on these logical formulae to design **synthesis algorithms**.

Table of contents

Summary


Multi-objective MDPs [MR25]
Randomisation
Classifying Mealy machines [MR24]
Strategy complexity
Memory
Representations
Nash equilibria [Mai24]
OC-MDPs [AMNR25]


THANK YOU

References I

Illustration images used in these slides originate from irasutoya.com.

[Ajd+25] Michal Ajdarów et al. "Taming Infinity one Chunk at a Time: Concisely Represented Strategies in One-Counter MDPs". In: *Proceedings of the 52th International Colloquium on Automata, Languages, and Programming, ICALP 2025, July 8–11, 2025, Aarhus, Denmark*. Ed. by Keren Censor-Hillel et al. Vol. 334. LIPIcs. Schloss Dagstuhl –Leibniz-Zentrum für Informatik, 2025.

[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. URL: https://doi.org/10.4230/LIPIcs.STACS.2024.50.

References II

[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.

[MR25] James C. A. Main and Mickael Randour. "Mixing Any Cocktail with Limited Ingredients: On the Structure of Payoff Sets in Multi-Objective MDPs and its Impact on Randomised Strategies". In: *CoRR* abs/2502.18296 (2025). DOI: 10.48550/arXiv.2502.18296.