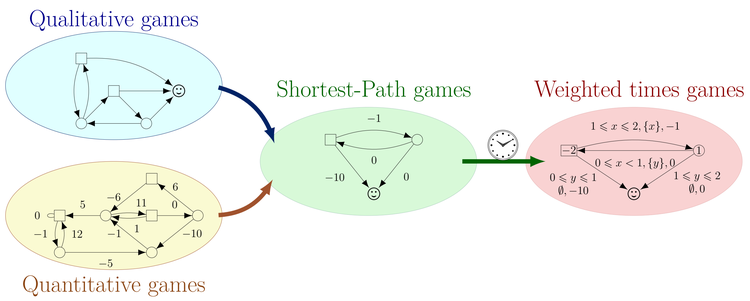
Playing Stochastically in Weighted Timed Games to Emulate Memory

Benjamin Monmege Julie Parreaux Pierre-Alain Reynier

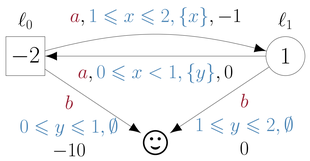
Benjamin.Monmege@univ-amu.fr Julie.Parreaux@univ-amu.fr Pierre-Alain.Reynier@univ-amu.fr

Aix–Marseille University, France

Weighted Timed Games



How to play in a weighted timed game?



When it is its turn, □ or ○ chooses an edge and a delay such that clocks satisfy the edge's guard.

$$\rho : (\ell_1, \begin{pmatrix}0\\0\end{pmatrix}) \xrightarrow[1 \times 0.5 + 0]{a,\ 0.5} (\ell_0, \begin{pmatrix}0.5\\0\end{pmatrix}) \xrightarrow[-2 \times 1.25 - 1]{a,\ 1.25} (\ell_1, \begin{pmatrix}0\\1.25\end{pmatrix}) \xrightarrow[1 \times \frac{1}{3} + 0]{b,\ 1/3} (☺, \begin{pmatrix}1/3\\19/12\end{pmatrix}) \rightsquigarrow -\frac{8}{3}$$

Objective of players

For all plays ρ, we define

$$\mathbf{SP}(\rho) = \begin{cases} \sum_{i=0}^{n-1} (w(e_i) + t\, w(\ell_i)) & \text{if } n \text{ is the smallest index s.t. } \ell_n = ☺ \\ +\infty & \text{if } \rho \text{ does not reach } ☺ \end{cases}$$

and the objective of each player is

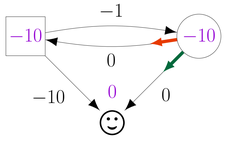
□ Maximizes **SP** ○ Minimizes **SP**

○ needs memory to play (ε-)optimally.

Switching strategy

A memory combination between two memoryless strategies

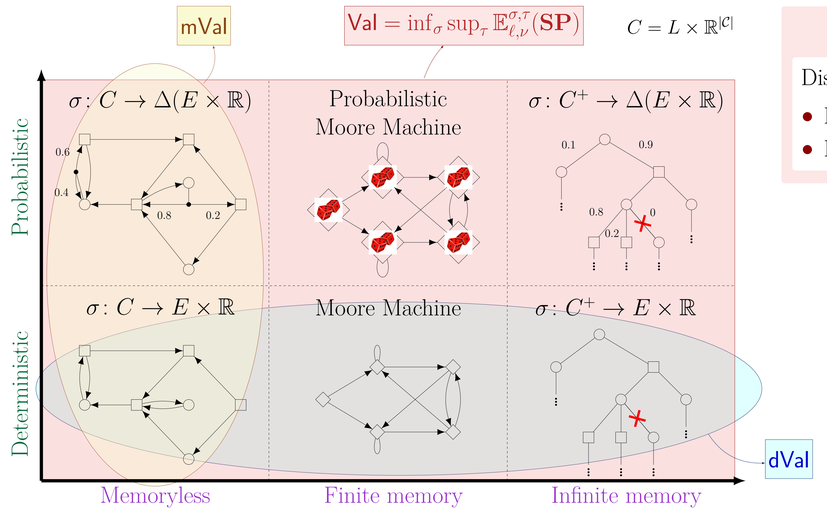
- σ_1: reach negative cycle
- σ_2: reach ☺
- K: number of turns before switch



Existence of a switching strategy?

Requires divergent weighted timed game : an extension, in negative case, of non-Zeno cost weighted timed game

Strategies and their value



A probabilistic strategy

Distribution over possible choices:

- Edge **b**: discrete distribution
- Delay for **b**: continuous distribution

Existence of Val?

Requires measurability conditions on strategies

Objectives

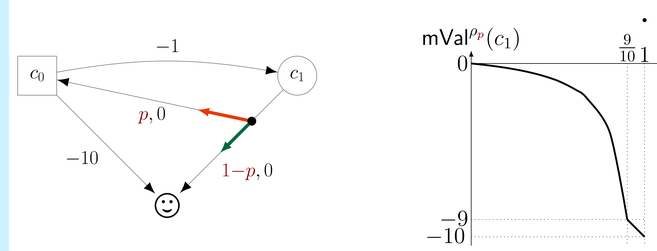
- Trading memory with probabilities
- Define probabilistic strategies in weighted timed games and their value
- Prove the equality between the probabilistic, memoryless and deterministic values

Randomisation emulate memory

Let $\langle \sigma_1, \sigma_2, K \rangle$ be an ε-optimal switching strategy, for all $p \in (0,1)$,

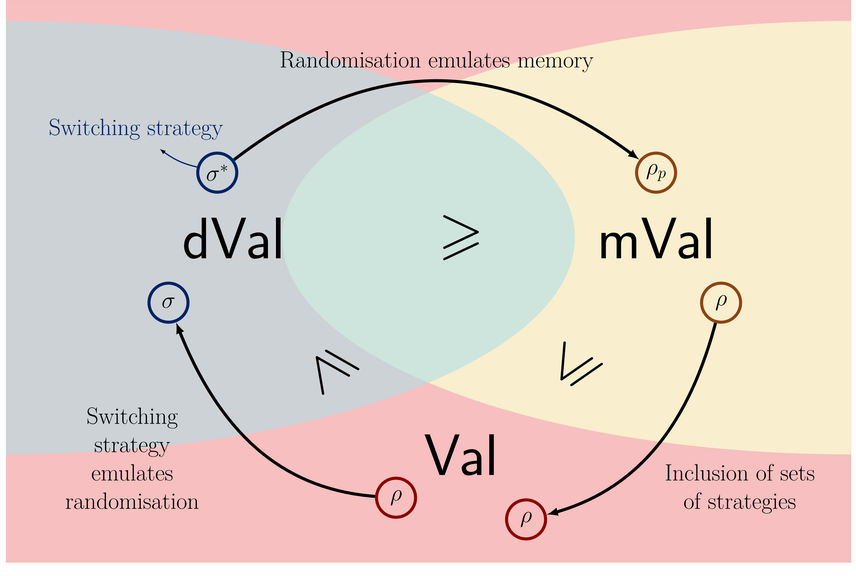
$$\rho_p = p \times \sigma_1 + (1-p) \times \sigma_2$$

Computation of $\mathsf{mVal}^{\rho_p}(c_1)$

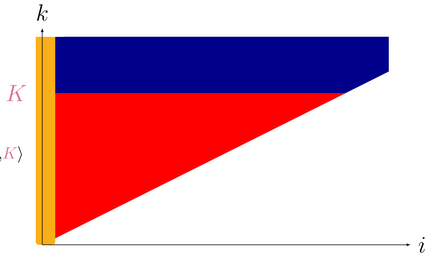


In this example : $\lim_{p \to 1} \mathsf{mVal}^{\rho_p}(c_1) = \mathsf{dVal}(c_1)$ with $p \neq 1$

Main result



Proof that $\mathsf{mVal}^{\rho_p}(c_1) \leqslant \mathsf{dVal}(c_1)$



Yellow zone: All plays conforming to σ_1, $\mathbf{SP}(\rho) \leqslant \mathsf{dVal}^{\langle \sigma_1, \sigma_2, K \rangle}$

Blue zone: Plays with many negative cycles $\mathbf{SP}(\rho) \leqslant \mathsf{dVal}^{\langle \sigma_1, \sigma_2, K \rangle}$

Red zone : Rest of plays such that $\lim_{p \to 1} \mathbb{P} = 0$

References

[1] Bertrand, N., Bouyer, P., Brihaye, T., Menet, Q., Baier, C., Größer, M., and Jurdzinski, M. (2014). Stochastic timed automata. *Log. Methods Comput. Sci.*, 10(4).

[2] Busatto-Gaston, D., Monmege, B., and Reynier, P. (2017). Optimal reachability in divergent weighted timed games. In *Foundations of Software Science and Computation Structures - 20th International Conference, FOSSACS 2017*.

[3] Chatterjee, K., Henzinger, T. A., and Prabhu, V. S. (2008). Trading infinite memory for uniform randomness in timed games. In *Hybrid Systems: Computation and Control, 11th International Workshop, HSCC 2008*.