CONCUR'19

Life is Random, Time is Not: MDPs with Window Objectives

T. Brihaye, F. Delgrange, Y. Oualhadj, M. Randour

⚠ Focus on intuit°, not tech details

1) Window objectives?

→ Introduced by Chatterjee, Doyen, R., Raskin in ATVA'13.

Context: 2-p turn-based games on graphs

Goals

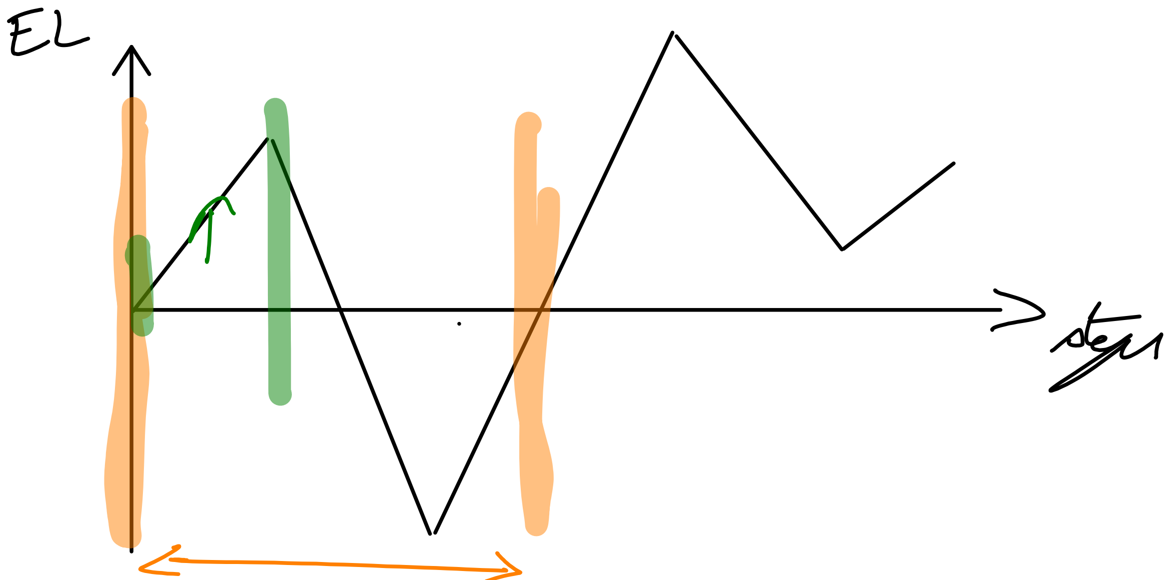
1. Strengthen classical objectives (MP, TP) with time bounds
2. Bypass complexity barriers (NP ∩ coNP in 1-dim, undec in k-dim for TP)

⟹ Window objectives
‖
Conservative approximations

Illustration for MP

↳ Weighted game

$$\overline{MP}(\pi) = \limsup \frac{1}{n} \sum^{n} w(s_i, s_{i+1})$$



λ max window size

⟶ Window slides along π.

Several variants:

- Prefix independent or not
- λ fixed or $\exists ? \lambda$

↳ Formal definition for MDPs incoming.

Landscape for games

	parity		MP/TP	
	compl.	mem.	compl.	mem.
DFW	P-c.*	poly	P-c.*	poly
FW				
BW		memoryless	NP $\cap$ coNP	memoryless

* For BW $\rightsquigarrow$ finitary parity

* In λ also (assumed to be in unary)

Good news: complexity
Bad news: memory

Parity results from Bruyère, Hautem, R. (GandALF'16)

+ many more related papers

Our goal

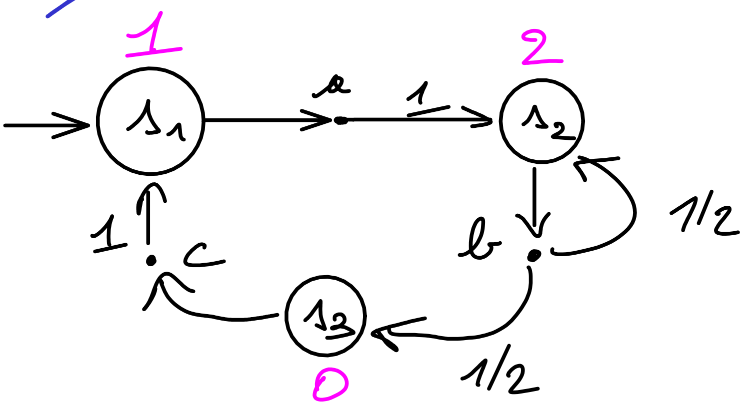
Lift to MDPs

$\Longrightarrow$ Treshold probability problem

$\exists^? \sigma,\ \mathbb{P}^{\sigma}[\Omega] \geq \alpha$

$\hookrightarrow$ encompasses many other problems (e.g., $\mathbb{E}$ via reaching the good ECs)

Example



→ states, actions

→ priorities

↳ Here, Markov chain (MC)

↳ min parity

⟹ any run is winning

$s_1 (s_2^* s_3 s_1)^* s_2^\omega \Downarrow 2$

$(s_1 s_2^* s_3)^\omega \Downarrow 0$

⟹ We win not only AS, — $\mathbb{P} = 1$

but surely — $\forall \pi$

⟹ Now, consider window parity

Intuitively, the min parity inside a window of size $\leq \lambda$ must be even, with this window sliding along the run.

$\Longrightarrow$ For any $\lambda \in \mathbb{N}_0$, $\mathbb{P} > 0$ to falsify this at every visit of s_1.

$(1/2^{\lambda-1}) \Longrightarrow$ bad window

But s_1 seen infinitely often with $\mathbb{P} = 1$ (because BSCC)

$\Longrightarrow \mathbb{P}(\text{window parity}) = 0$

⚠ Striking $\neq$ between parity and window parity due to time bounds.

often desired in applications

Here, either parameter or $\exists$?

II) Our contribution

↳ unified view of all window objectives (here MP and Parity)

↳ generic approach.

	parity		MP/TP	
	compl.	mem.	compl.	mem.
DFW	P-c.	poly	EXP-c./PSPACE-h	pseudo-poly
FW			P-c.	poly.
BW		memoryless	NP ∩ coNP	memoryless

Formal defs

Threshold probability problem

$$GW_{MP}(\lambda) = \{\rho \in Runs(\mathcal{M}) \mid \exists \ell < \lambda, MP(\rho[0, \ell+1]) \geq 0\}$$

$$GW_{par}(\lambda) = \{ \text{''} \mid \exists \ell < \lambda, (p(\rho[\ell]) \bmod 2 = 0 \ \wedge \ \forall k < \ell, p(\rho[\ell]) < p(\rho[k]))\}$$

$\Omega \in \{MP, par\}$

$$DFW_\Omega(\lambda) = \{ \text{''} \mid \forall j \geq 0, \rho[j, \infty] \in GW_\Omega(\lambda)\}$$

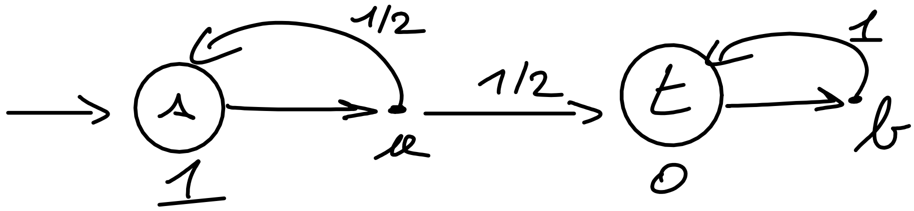
$$FW_\Omega(\lambda) = \{ \text{''} \mid \exists i \geq 0, \rho[i, \infty] \in DFW_\Omega(\lambda)\}$$

$$BW_\Omega = \{ \text{''} \mid \exists \lambda > 0, \rho \in FW_\Omega(\lambda)\}$$

↳ Need not be uniform over all runs.

Comparison with games

→ Clear ≠ in behaviors
E.g., uniform bound in games, not in MDPs



Almost all runs are "bounded" hence $\mathbb{P}(\mathrm{DBW}_{par}) = 1$ but for all λ, $\mathbb{P}(\mathrm{DFW}_{par}(\lambda)) < 1$

→ Despite that, almost identical results complexity-wise

↳ ≠ in DFW_{MP}

P-c in games

PSPACE-h. here

⇓

Because we can emulate shortest path problems on MDPs.

⇓

AS case collapses to P.

→ In games, complexity of window objectives lower than classical ones.

Here, main interest is modeling power since classical objectives already in P.

⚠ Could still prove more efficient in practice.

III) Technical overview

A) DFW

↔ Reduction to safety over well-chosen unfoldings (sink = seeing a bad window)

⟹ Poly for parity
Pseudo-poly for MP

⟹ Almost tight complexities (PSPACE-h even for acyclic MDPs)

⚠ No upper bound on λ

≠ games

B) FW and BW

→ We first use similar ω-Büchi reduction to prove that FM strategies suffice

⇒ We can do better complexity-wise.

→ Study of End Components

crucial for all prefix-ind. objs in MDPs

→ Classificat° based on 2-p game interpretat° of MECs

uses FM strategies

games algos as black boxes

good ECs **OR** bad ECs

$\exists \sigma$ AS$^{\sigma}$(obj) $\quad$ $\nexists \sigma$ $\mathbb{P}$(obj) > 0

Zero-one law

→ Lift to general MDPs

↳ Reach (good ECs)

→ complexity dominated by the classificat° (solving 2-p games).

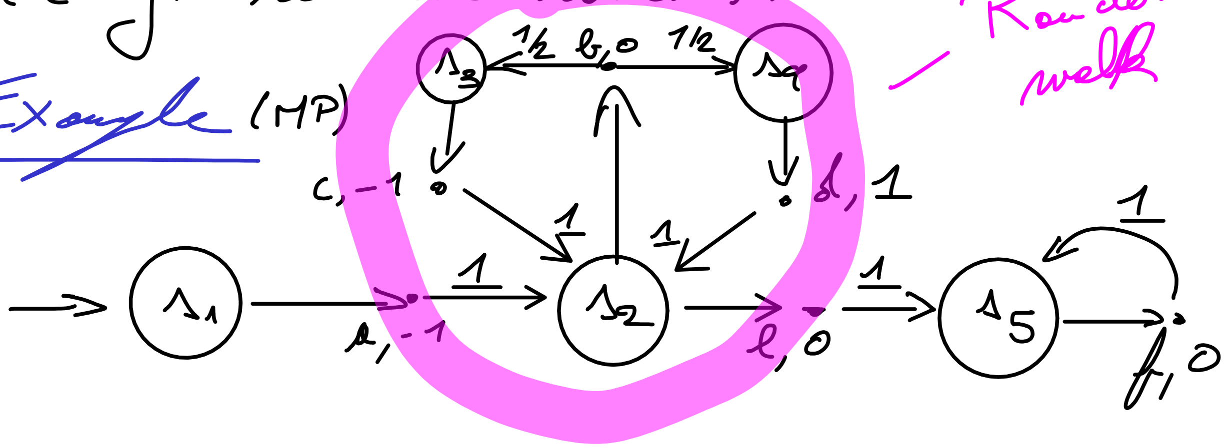
→ Pure strategies always suffice.

c) DBW

→ Very strange behavior in MDPs

→ Left out because not natural and encodes many complex models (e.g. random walks)

Example (MP)



↳ Infinite memory strategy needed to ensure $AS(DBW_{MP})$

↳ Even for qualitative question, actual probabilities matter!

D | Comments and future work

$\star$ MCs

Still PP-h. for DFW_{MP} so not much to gain.

$\star$ $\mathbb{E}$

$$\mathbb{E}^{\sigma}_{\mathcal{M},s,\Omega}(\lambda) = \sum_{\lambda > 0}^{\infty} \lambda \cdot \mathbb{P}^{\sigma}_{\mathcal{M},s}\left[FW_\Omega(\lambda) \setminus FW_\Omega(\lambda-1)\right]$$

↳ Easy to do: binary search, etc.

$\star$ Multi-objective

⟶ Effortless extension (replace the black-boxes in the classificat°)

$\star$ Tool

On the way (based on Storm).

THANKS