Expected Window Mean-Payoff

Benjamin Bordais[1], Shibashis Guha[2], Jean-François Raskin[2]

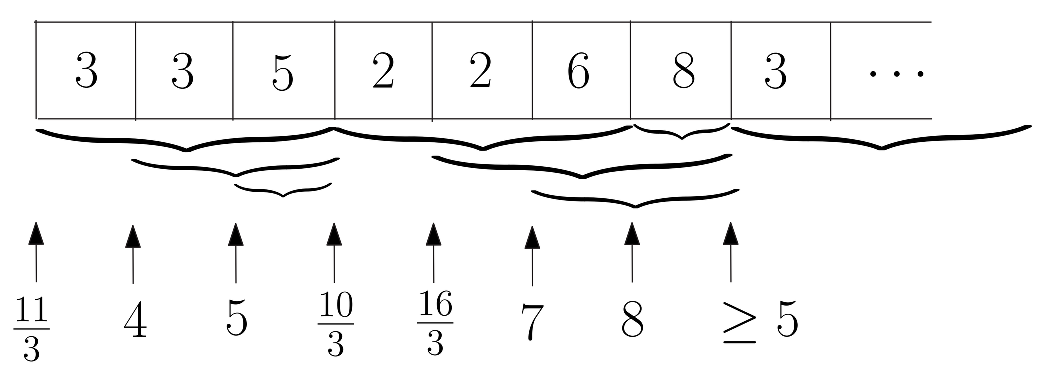
[1]École Normale Supérieure de Rennes

[2]Université libre de Bruxelles

Window Mean-Payoff

- **mean-payoff**: Does not guarantee local stability
- **window mean-payoff**: payoffs over a local finite length window sliding along the infinite sequence
- **good window property**: the mean-payoff must always reach a given threshold within the window length ℓ
- **strengthening of the classical mean-payoff objective**: for all lengths ℓ, and all infinite sequences π, payoffs that satisfy the window mean-payoff objective for threshold λ imply that π has a mean-payoff value of at least λ

$\ell = 3$



$$WMP(\pi, \ell) = \max_{k \in [\ell]} \{\frac{1}{k}\sum_{i=1}^{k} w(e_i)\}$$

$$WMP(\pi, \ell) \geq \lambda \Rightarrow \pi \in \mathbf{GW}(\lambda, \ell)$$

- **Good Window Objective**: Given ℓ and a threshold $\lambda \in \mathbb{Q}$
  - $\mathbf{GW}(\lambda, \ell)$: $\mathbf{GW}(\frac{11}{3}, 3)$.
  - $\pi \in \mathbf{GW}(\lambda, \ell) \implies \pi \in \mathbf{GW}(\lambda, \ell')$ for $\ell' \geq \ell$.
  - $\mathbf{GW}(\lambda, \ell) \supseteq \mathbf{GW}(\lambda', \ell)$ for $\lambda' \geq \lambda$.
- **Direct (Prefix-dependent) Window Objectives**
  - **Direct Fixed Window Mean-Payoff Objective**: $\mathbf{DirWMP}(\lambda, \ell)$: Here $\pi \in \mathbf{DirWMP}(\frac{10}{3}, 3)$
  - **Direct Bounded Window Mean-Payoff Objective**: $\mathbf{DirBWMP}(\lambda)$: "does there exist some window length"?
  - Here $\pi \in \mathbf{DirBWMP}(5 - \varepsilon)$ for every small ϵ.
- **Non-direct (Prefix-independent) Window Objectives**
  - **Fixed Window Mean-Payoff Objective**: $\mathbf{WMP}(\lambda, \ell)$: Here $\pi \in \mathbf{WMP}(5, 3)$
  - **Bounded Window Mean-Payoff Objective**: $\mathbf{BWMP}(\lambda)$: Here $\pi \in \mathbf{BWMP}(5)$

Functions of Interest

$$f^{\ell}_{WMP}(\pi) = \sup\{\lambda \mid \pi \in WMP(\lambda, \ell)\}$$
$$f^{\ell}_{DirWMP}(\pi) = \sup\{\lambda \mid \pi \in DirWMP(\lambda, \ell)\}$$
$$f_{DirBWMP}(\pi) = \sup\{\lambda \mid \pi \in DirBWMP(\lambda)\}$$
$$f_{BWMP}(\pi) = \sup\{\lambda \mid \pi \in BWMP(\lambda)\}$$

Theorem: For every path π, $f_{DirBWMP}(\pi) = f_{BWMP}(\pi)$

Given an MDP, compute the best expected value for each of these functions.

Direct Fixed Window Objective

	MDP		
	Complexity	**Memory**	**Hardness**
DirWMP	exponential	exponential	PP

Computing $\mathbb{E}^{\Gamma}_{s_{init}}(f^{\ell}_{DirWMP})$

- Step 1: Construct a new MDP Γ' whose state space is $S' = S \times ([W]_0)^{\ell-1} \times \Lambda$ with $\Lambda = \{\frac{p}{q} \mid q \in [\ell], p \in [q \cdot W]\}$; a state keeps track of last $\ell - 1$ weights encountered and the minimum window mean-payoff seen so far.
- Step 2: Compute $\mathbb{E}^{\Gamma'}_{s'_{init}}(f_{Mean})$.

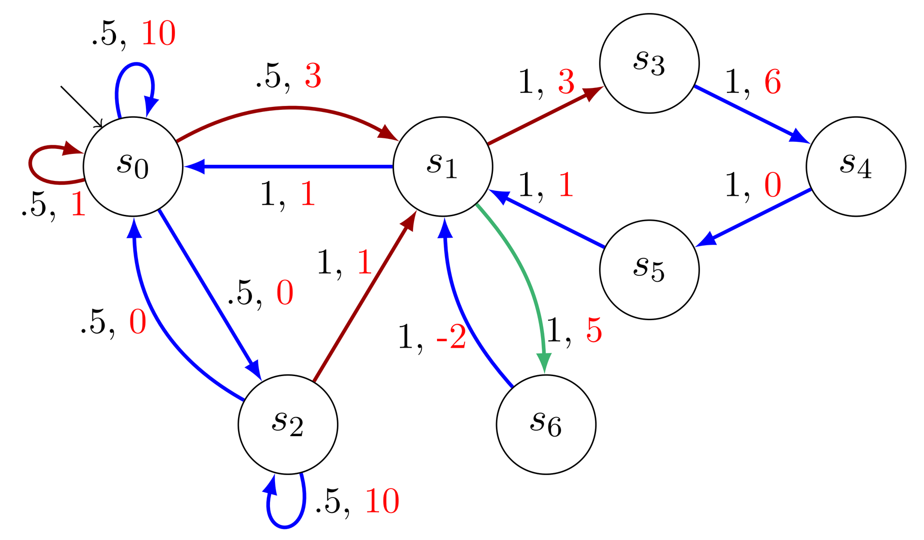
Complexity

$\mathbb{E}^{\Gamma}_{s_{init}}(f^{\ell}_{DirWMP})$ can be computed in time $O(poly(|S| \cdot W^{\ell} \cdot \ell^2))$.

Theorem: (Hardness)

Given an MDP Γ with an initial state s_{init}, a window length ℓ and a $\lambda \in \mathbb{Q}$, deciding whether $\mathbb{E}^{\Gamma}_{s_{init}}(f^{\ell}_{DirWMP}) \geq \lambda$ is PP-Hard.

Example: Non-direct (Prefix-independent) Objectives



- $\mathbb{E}^{M}_{s_0}(f^{\ell}_{WMP}) = 2$ (in s_1, actions brown and green);
- $\mathbb{E}^{M}_{s_0}(f_{BWMP}) = 2.5$ (in s_1, action blue);
- $\mathbb{E}^{M}_{s_0}(f_{Mean}) = 5$ (in s_0 and s_2, action blue).

Non-Direct Window Objective

	MDP		
	Complexity	**Memory**	**Hardness**
WMP	polynomial	polynomial	2-p DirWMP
BWMP	UP ∩ coUP	memoryless	2-p Mean-payoff

Computing $\mathbb{E}^{\Gamma}_{s_{init}}(f^{\ell}_{WMP})$ (resp. $\mathbb{E}^{\Gamma}_{s_{init}}(f_{BWMP})$)

- Step 1: Decompose Γ into MECs;
- Step 2: For each MEC M, compute the expected value λ^{ℓ}_{M} inside M by solving several **two-player games** for the **DirWMP** (resp. **MP**) objective (λ^{ℓ}_{M} is equal to the maximum that can be achieved among all starting states in M);
- Step 3: Construct a new MDP Γ^{MEC} where the weight appearing in an MEC M are replaced by λ^{ℓ}_{M};
- Step 4: Compute $\mathbb{E}^{\Gamma^{MEC}}_{s_{init}}(f_{Mean})$.

Complexity

- $\mathbb{E}^{\Gamma}_{s_{init}}(f^{\ell}_{WMP})$: $O(poly(|\Gamma|, \ell, \log_2 W))$ (with W maximum weight);
- $\mathbb{E}^{\Gamma}_{s_{init}}(f_{BWMP})$: UP ∩ coUP.

Theorem: (Relative Hardness)

Given a two-player game G and a window length ℓ, we can construct in log-space an MDP Γ_G (resp. Γ'_G) such that the outcome of the two-player game for the **DirWMP** (resp. **MP**) objective in G is equal to $\mathbb{E}^{\Gamma_G}_{s_{init}}(f^{\ell}_{WMP})$ (resp. $\mathbb{E}^{\Gamma'_G}_{s'_{init}}(f_{BWMP})$).

Window Mean-Payoff Objectives for Markov Chain

	Markov chain: complexity
WMP	polynomial
BWMP	polynomial
DirWMP	pseudopolynomial

WMP

Same as MDPs.

BWMP

- Algorithm: As for MDPs, with BSCCs instead of MECs, and, in each BSCC, we compute the mean over the minimum mean-cycle.
- Complexity: $\mathbb{E}^{\mathcal{M}}_{s_{init}}(f_{BWMP}) \geq \lambda$ can be computed in polynomial time.

DirWMP

- Algorithm:
  - Step 1 For every 'interesting' value λ, that is or all $\lambda \in \Lambda$, we construct $\mathcal{M}^{\lambda}$ so that $\Pr(\pi \in \mathsf{Paths}^{\mathcal{M}} \mid (f^{\ell}_{WMP})(\pi) \geq \lambda)) = \Pr(\pi \in \mathsf{Paths}^{\mathcal{M}^{\lambda}} \mid \pi \models \neg\Diamond\{trap\})$. We only need to remember the location of the largest window that is still open, as well as the 'amount of payoff' that is required to close it. Hence, in the Markov chain $\mathcal{M}^{\lambda_i}_{\ell}$, the state space $S' = (S \times [\ell - 1]_0 \times [W \cdot (\ell - 1)]_0) \cup \{trap\}$.
  - Step 2 Then, we can compute $\Pr(\pi \in \mathsf{Paths}^{\mathcal{M}} \mid (f^{\ell}_{WMP})(\pi) = \lambda))$. Computing $\mathbb{E}^{\mathcal{M}}_{s_{init}}(f^{\ell}_{DirWMP})$ follows;
- Complexity: $\mathbb{E}^{\mathcal{M}}_{s_{init}}(f^{\ell}_{DirWMP}) \geq \lambda$ can be computed in pseudopolynomial (in binary representation of W) time.