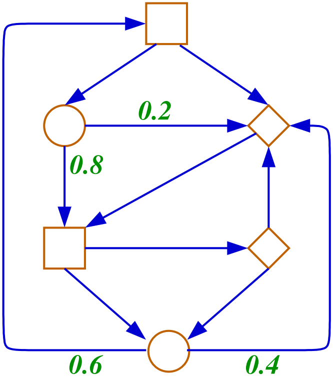
Stochastic Games with Branching Time Winning Objectives

Antonín Kučera

(a joint work with Tomáš Brázdil and Vojtěch Forejt)

Faculty of Informatics, Masaryk University, Brno

Simple stochastic games



- $G = (V, E, (V_\Box, V_\Diamond, V_\bigcirc), Prob)$, each vertex has a successor
- Markov chains: $V_\Box = V_\Diamond = \emptyset$
- Markov decision processes: $V_\Diamond = \emptyset$

Strategies

Let $G = (V, E, (V_\Box, V_\Diamond, V_\bigcirc), Prob)$ be a game.

- A strategy for player $\Box$ is a function σ which to every $wv \in V^* V_\Box$ assigns a probability distribution over the set of outgoing edges of v.
- Memory: history-dependent (H), finite-memory (F), memoryless (M)
- Randomization: randomized (R), deterministic (D)
- Thus, we obtain the classes of MD, MR, FD, FR, HD, and HR strategies.

Plays and Winning Objectives

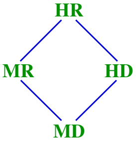
- Each pair (σ, π) of strategies for player $\Box$ and player $\Diamond$ determine a unique play $G^{(\sigma,\pi)}$, which is a Markov chain where V^+ is the set of states and transitions are defined accordingly (if σ, π are memoryless, the set of states of $G^{(\sigma,\pi)}$ can be just V).
- A winning objective is some property P of states in Markov chains that is to be achieved by player $\Box$ and spoilt by player $\Diamond$.
- A strategy σ is winning for player $\Box$ in a vertex v iff P is valid in the state v of $G^{(\sigma,\pi)}$ for every strategy π of player $\Diamond$. Similarly, we define a winning strategy for player $\Diamond$.

Winning objectives

- Linear-time objectives
  - ⋆ Qualitative/quantitative Büchi, co-Büchi, Rabin, Street, Muller, etc.
  - ⋆ These games are determined, have equilibria, winning/optimal strategies are memoryless or require only a finite memory, etc.
- Branching-time objectives
  - ⋆ Specified by formulae of branching-time logics that are interpreted over Markov chains.
  - ⋆ $\mathcal{G}^{=1}(p \Rightarrow \mathcal{F}^{\geq 0.1} q)$
  - ⋆ Properties of stochastic games with branching-time objectives are quite different from the ones with linear-time objectives.

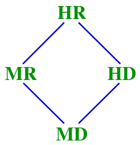
Properties of games with branching-time objectives (I)

- Memory and randomization help:

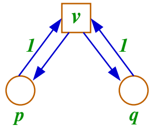


Properties of games with branching-time objectives (I)

- Memory and randomization help:



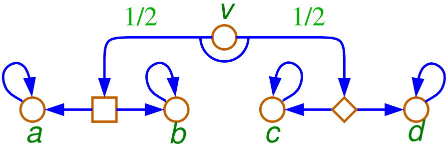
- Consider the following game:



- $\mathcal{X}^{=1} p \wedge \mathcal{F}^{=1} q$. Requires memory.
- $\mathcal{X}^{>0} p \wedge \mathcal{X}^{>0} q$. Requires randomization.
- $\mathcal{X}^{>0} p \wedge \mathcal{X}^{>0} q \wedge \mathcal{F}^{=1} \mathcal{G}^{=1} q$. Requires both memory and randomization.
- In some cases, infinite memory is required.

Properties of games with branching-time objectives (II)

- The games are not determined (for any strategy type).
- $\mathcal{F}^{=1}(a \vee c) \;\vee\; \mathcal{F}^{=1}(b \vee d) \;\vee\; \big(\mathcal{F}^{>0}c \wedge \mathcal{F}^{>0}d\big)$



Who wins the game (MD strategies) ?

Theorem

The existence of a winning MD strategy for player $\Box$ is $\Sigma_2 = \mathbf{NP}^{\mathbf{NP}}$-complete.

Who wins the game (MD strategies) ?

Theorem

The existence of a winning MD strategy for player $\square$ is $\Sigma_2 = \boldsymbol{NP}^{\boldsymbol{NP}}$-complete.

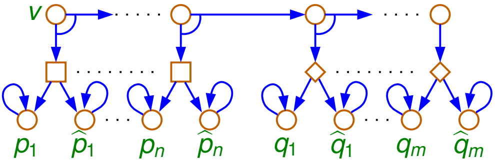
- The membership to Σ_2 follows easily.

Who wins the game (MD strategies) ?

Theorem

The existence of a winning MD strategy for player $\Box$ is $\Sigma_2 = \mathbf{NP}^{\mathbf{NP}}$-complete.

- The membership to Σ_2 follows easily.
- The Σ_2-hardness can be established as follows:
  - ⋆ Let $\exists x_1, \cdots, x_n \forall y_1, \cdots, y_m\, B$ be a Σ_2 formula.
  - ⋆ Consider the following game:



  - ⋆ Let φ be the PCTL formula obtained from B by substituting each occurrence of x_i, $\neg x_i$, y_j, and $\neg y_j$ with $\mathcal{F}^{>0} p_i$, $\mathcal{F}^{>0} \widehat{p}_i$, $\mathcal{F}^{>0} q_j$, and $\mathcal{F}^{>0} \widehat{q}_j$, respectively.

Who wins the game (MR strategies) ?

Theorem

*The existence of a winning MR strategy for player $\Box$ is Σ_2-hard and in **EXPTIME**. For the qualitative fragment of PCTL, the problem is Σ_2-complete.*

Who wins the game (HD, HR, FD, FR strategies) ?

Theorem

The existence of a winning HD (or HR) strategy for player $\Box$ in MDPs is highly undecidable (and Σ_1^1-complete). Moreover, the existence of a winning FD (or FR) strategy is also undecidable.

Who wins the game (HD, HR, FD, FR strategies) ?

Theorem

The existence of a winning HD (or HR) strategy for player $\Box$ in MDPs is highly undecidable (and Σ_1^1-complete). Moreover, the existence of a winning FD (or FR) strategy is also undecidable.

- The result holds for the $\mathcal{L}(\mathcal{F}^{=1/2}, \mathcal{F}^{=1}, \mathcal{F}^{>0}, \mathcal{G}^{=1})$ fragment of PCTL (the role of $\mathcal{F}^{=1/2}$ is crucial).
- The proof is obtained by reduction of the problem whether a given non-deterministic Minsky machine has an infinite recurrent computation.

The undecidability proof

- A non-deterministic Minsky machine $\mathcal{M}$ with two counters c_1, c_2:

$$1 : ins_1, \; \cdots, \; n : ins_n$$

where each ins_i takes one of the following forms:

  - $c_j := c_j + 1;\; goto\ k$
  - $if\ c_j{=}0\ then\ goto\ k\ else\ c_j := c_j - 1;\; goto\ m$
  - $goto\ \{k\ or\ m\}$

The undecidability proof

- A non-deterministic Minsky machine $\mathcal{M}$ with two counters c_1, c_2:

$$1 : ins_1, \; \cdots, \; n : ins_n$$

  where each ins_i takes one of the following forms:

  - ⋆ $c_j := c_j + 1;\ goto\ k$
  - ⋆ $if\ c_j{=}0\ then\ goto\ k\ else\ c_j := c_j - 1;\ goto\ m$
  - ⋆ $goto\ \{k\ or\ m\}$

- The problem whether a given non-deterministic Minsky machine with two counters initialized to zero has an infinite computation that executes ins_1 infinitely often is Σ_1^1-complete.

The undecidability proof

- A non-deterministic Minsky machine $\mathcal{M}$ with two counters c_1, c_2:

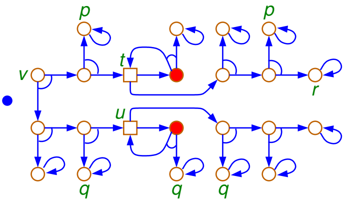
$$1: ins_1, \ \cdots, \ n: ins_n$$

  where each ins_i takes one of the following forms:

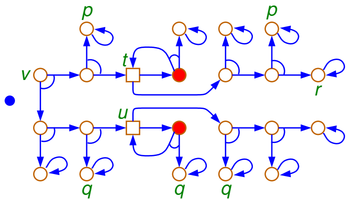
  - $\star$ $c_j := c_j + 1; \ goto\ k$
  - $\star$ $if\ c_j{=}0\ then\ goto\ k\ else\ c_j := c_j - 1; \ goto\ m$
  - $\star$ $goto\ \{k\ or\ m\}$

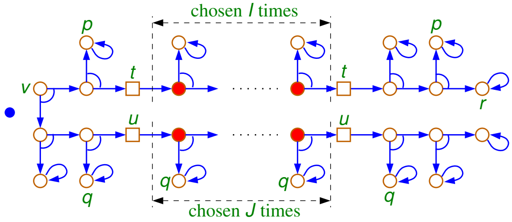
- The problem whether a given non-deterministic Minsky machine with two counters initialized to zero has an infinite computation that executes ins_1 infinitely often is Σ_1^1-complete.
- For a given machine $\mathcal{M}$, we construct a finite-state MDP $G(\mathcal{M})$ and a formula $\varphi \in \mathcal{L}(\mathcal{F}^{=1/2}, \mathcal{F}^{=1}, \mathcal{F}^{>0}, \mathcal{G}^{=1})$ such that $\mathcal{M}$ has an infinite recurrent computation iff player $\Box$ has a winning HD (or HR) strategy for φ in a distingushed vertex v of $G(\mathcal{M})$.

The construction of $G(\mathcal{M})$ and φ

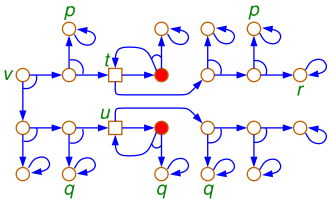


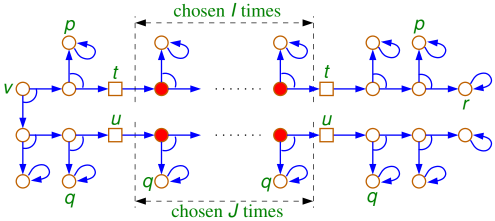
The construction of $G(\mathcal{M})$ and φ





The construction of $G(\mathcal{M})$ and φ

- 

- 

- $I = J < \omega$ iff $v \models \mathcal{F}^{>0} r \wedge \mathcal{F}^{=1/2}(p \vee q)$
- The probability of $\mathcal{F}(p \vee q)$: $0.01\underbrace{0\cdots0}_{I}01 + 0.001\underbrace{1\cdots1}_{J}1$

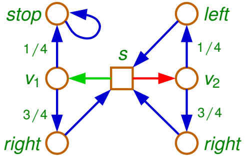
Positive results

- We restrict ourselves to qualitative fragments of probabilistic branching time logics.

Positive results

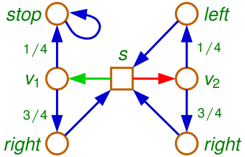
- We restrict ourselves to qualitative fragments of probabilistic branching time logics.
- Even MDPs with qualitative PCTL objectives may require infinite memory.

Positive results

- We restrict ourselves to qualitative fragments of probabilistic branching time logics.
- Even MDPs with qualitative PCTL objectives may require infinite memory.
- 

$$\mathcal{G}^{>0}(\neg stop \wedge \mathcal{F}^{>0} stop)$$

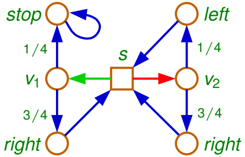
Positive results

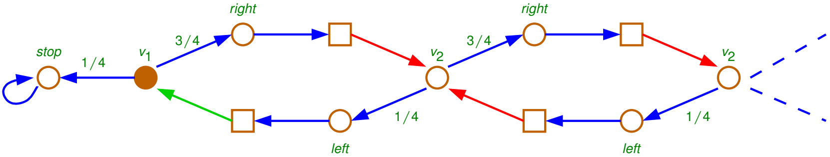
- We restrict ourselves to qualitative fragments of probabilistic branching time logics.
- Even MDPs with qualitative PCTL objectives may require infinite memory.
- 

$$\mathcal{G}^{>0}(\neg stop \wedge \mathcal{F}^{>0} stop)$$

- A winning strategy: if $\#left < \#right$ use the red transition, otherwise use the green one.

Positive results

- We restrict ourselves to qualitative fragments of probabilistic branching time logics.
- Even MDPs with qualitative PCTL objectives may require infinite memory.
- 

  $\mathcal{G}^{>0}(\neg \mathit{stop} \wedge \mathcal{F}^{>0} \mathit{stop})$
- A winning strategy: if $\#\mathit{left} < \#\mathit{right}$ use the red transition, otherwise use the green one.
- 

The Result

Theorem

- *The existence of a winning HD (or HR) strategy for player $\Box$ in MDPs with qualitative PECTL* objectives is decidable in time which is polynomial in the size of MDP and doubly exponential in the size of the formula.*
- *Moreover, iff there is a winning HD (or HR) strategy, there is also a one-counter winning strategy and one can effectively construct a one-counter automaton which implements this strategy (the associated complexity bounds are the same as above).*

Open problems

- Exact complexity bounds (e.g., the existence of a winning MR strategy for MDPs or stochastic games with PCTL objectives).
- How about stochastic games with qualitative branching-time objectives ?
- How about infinite-state MDPs and stochastic games?