Stochastic Games

(in Formal Verification)

Antonín Kučera

Masaryk University Brno

Game theory

Game theory studies the behavior of rational "players" who can make choice and attempt to achieve a certain objective. A player's success depends on the choices of the other players.

- stochastic games:
  - the impact of players' choices in uncertain;
  - the players' choice can be randomized.
- games in computer science:
  - formal semantics;
  - communication protocols;
  - Internet auctions;
  - ...many other things.

Stochastic games in formal verification

Our setting:

- state space: discrete
- players: controller, environment
- objectives: antagonistic
- choice: turn-based, randomized
- information: perfect

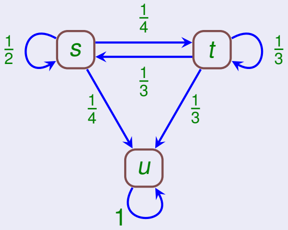
Is there a strategy for the controller such that the system satisfies a certain property no matter what the environment does?

Outline

- Preliminaries.
  - Games, strategies, objectives.
- Stochastic games with reachability objectives.
  - The (non)existence of optimal strategies.
  - Algorithms for finite-state games.
- Stochastic games with branching-time objectives.
- Stochastic games with time.

Markov chains

Definition 1 (Markov chain)

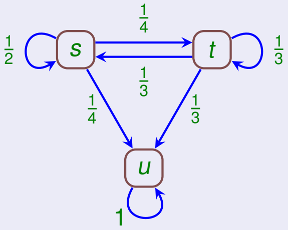


$\mathcal{M} = (S, \rightarrow, Prob)$

- S is at most countable set of *states*;
- $\rightarrow \subseteq S \times S$ is a *transition relation*;
- $Prob$ is a *probability assignment*.

Markov chains

Definition 1 (Markov chain)



$\mathcal{M} = (S, \rightarrow, Prob)$

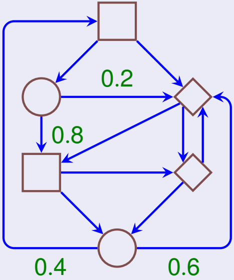
- S is at most countable set of *states*;
- $\rightarrow \subseteq S \times S$ is a *transition relation*;
- $Prob$ is a *probability assignment*.

We want to measure the probability of certain subsets of $Run(s)$.

- For every finite path w initiated in s, we define the probability of $Run(w)$ in the natural way.
- This assignment can be uniquely extended to the (Borel) σ-algebra $\mathcal{F}$ generated by all $Run(w)$.
- Thus, we obtain the probability space $(Run(s), \mathcal{F}, \mathcal{P})$.

Turn-based stochastic games

Definition 2 (Turn-based stochastic game)



$G = (V, E, (V_\Box, V_\Diamond, V_\bigcirc), Prob)$

- *the set V is at most countable;*
- *each vertex has a successor;*
- *$Prob$ is positive;*
- *G is a Markov decision process (MDP) if $V_\Diamond = \emptyset$ or $V_\Box = \emptyset$.*

Strategies

Definition 3 (Strategy)

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.

- A strategy for player $\Diamond$ is defined analogously.
- We can classify strategies according to
  - memory requirements: 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

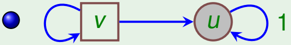
Definition 4 (Play)

Let $G = (V, E, (V_\Box, V_\Diamond, V_\bigcirc), Prob)$ be a game. Each pair (σ, π) of strategies for player $\Box$ and player $\Diamond$ determines a unique play $G^{(\sigma,\pi)}$, which is a Markov chain where V^+ is the set of states and transitions are defined accordingly.

- Plays are infinite trees.
- For a pair of memoryless strategies (σ, π), the play $G^{(\sigma,\pi)}$ can be depicted as a Markov chain with the set of states V.

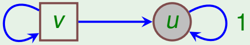
Plays (2)

Example 5 (A game and its play)



Plays (2)

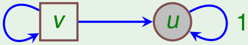
Example 5 (A game and its play)

- 

- Is there a strategy σ such that $v \models \mathcal{G}^{>0}(v)$ in G^σ ?

Plays (2)

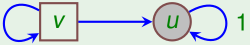
Example 5 (A game and its play)

- 

- Is there a strategy σ such that $v \models \mathcal{G}^{>0}(v)$ in G^σ ?
- Is there a strategy σ such that $v \models \mathcal{G}^{>0}(v \wedge \mathcal{F}^{>0}u)$ in G^σ ?

Plays (2)

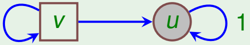
Example 5 (A game and its play)

- 

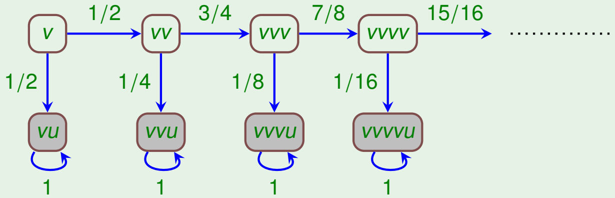
- Is there a strategy σ such that $v \models \mathcal{G}^{>0}(v)$ in G^{σ} ?
- Is there a strategy σ such that $v \models \mathcal{G}^{>0}(v \wedge \mathcal{F}^{>0}u)$ in G^{σ} ?
  - Obviously, there is no such MR (or even FR) strategy.

Plays (2)

Example 5 (A game and its play)

- 

- Is there a strategy σ such that $v \models \mathcal{G}^{>0}(v)$ in G^σ ?
- Is there a strategy σ such that $v \models \mathcal{G}^{>0}(v \wedge \mathcal{F}^{>0}u)$ in G^σ ?
  - Obviously, there is no such MR (or even FR) strategy.
  - Let $\sigma(wv) \;=\; v \xrightarrow{1/2^{|wv|}} u,\ v \xrightarrow{1-1/2^{|wv|}} v$



A taxonomy of objectives

- Each play of a game G is assigned a (numerical) yield. The goal of player $\Box/\Diamond$ is to maximize/minimize the yield.
- Win-lose objectives assign either 1 or 0 to each play.
  - $P^{\bowtie\varrho}\varphi$, where φ is an LTL formula.
  - PCTL or PCTL* objectives.
- Objectives specified by Borel measurable payoffs.
  - $yield(G^{\sigma,\pi}) = \mathbb{E}(f^{\sigma,\pi})$, where $f : Run(G) \rightarrow \mathbb{R}$ is measurable.
  - Qualitative payoffs assign either 1 or 0 to each run
    - Büchi, parity, Rabin, Street, Muller, etc.
  - Quantitative payoffs
    - Mean payoff: $MP(w) = \lim_{n\rightarrow\infty} \frac{\sum_{i=0}^{n} rew(w(i))}{n}$
    - Discounted payoff: $DP(w) = \sum_{i=0}^{\infty} \lambda^i \cdot rew(w(i))$

The problems of interest

- Win-lose objectives
  - Determinacy: does one of the two players always have a winning strategy? If so, what type of strategy?
  - Can we effectively determine the winner and compute a winning strategy for her?
- Objectives specified by Borel measurable payoffs
  - Is there an equilibrium value?
  - If so, do the players have optimal strategies? And of what type?
  - Can we compute the value and (ε-) optimal strategies?

The existence of an equilibrium value

Theorem 6 (Martin, 1998; Maitra & Sudderth, 1998)

Let $G = (V, E, (V_\square, V_\Diamond, V_\bigcirc), Prob)$ be a game, $v \in V$, and $f : Run(G) \to \mathbb{R}$ a bounded Borel measurable payoff. Then

$$\sup_\sigma \inf_\pi \mathbb{E}(f_v^{\sigma,\pi}) \quad = \quad \inf_\pi \sup_\sigma \mathbb{E}(f_v^{\sigma,\pi})$$

- Thm. 6 does not impose any restrictions on G. The set of vertices and the branching degree of G can be infinite.
- References:

  - D.A. Martin. *The Determinacy of Blackwell Games.* The Journal of Symbolic Logic, Vol. 63, No. 4 (Dec., 1998), pp. 1565–1581.
  - A. Maitra and W. Sudderth. *Finitely Additive Stochastic Games with Borel Measurable Payoffs.* International Journal of Game Theory, Vol. 27 (1998), pp. 257–267.

Optimal strategies

Definition 7

Let $G = (V, E, (V_\Box, V_\Diamond, V_\bigcirc), Prob)$ be a game, $v \in V$, and $f : Run(G) \to \mathbb{R}$ a bounded Borel measurable payoff. Let $\varepsilon \in [0, 1]$.

- *An ε-optimal maximizing strategy is a strategy σ for player $\Box$ such that for every strategy π of player $\Diamond$ we have that $\mathbb{E}(f_v^{\sigma,\pi}) \geq val_f(v) - \varepsilon$.*
- *An ε-optimal minimizing strategy is a strategy π for player $\Diamond$ such that for every strategy σ of player $\Box$ we have that $\mathbb{E}(f_v^{\sigma,\pi}) \leq val_f(v) + \varepsilon$.*

An optimal maximizing/minimizing strategy is a 0-optimal maximizing/minimizing strategy.

- According to Thm. 6, ε-optimal maximizing/minimizing strategies exist for every $\varepsilon > 0$.
- . . . and we cannot say much more in the general setting.

Reachability objectives

Now we examine the properties of our interest for reachability objectives in greater detail (and reveal some surprising facts).

- Let $G = (V, E, (V_\Box, V_\Diamond, V_\bigcirc), Prob)$ be a game, $T \subseteq V$ a set of target vertices.
- Let $Reach(T)$ be the set of all runs that visit T.
- The goal of player $\Box/\Diamond$ is to maximize/minimize the probability of $Reach(T)$.

Reachability games have a value (1)

Theorem 8

Let $G = (V, E, (V_\Box, V_\Diamond, V_\bigcirc), Prob)$ be a game, $T \subseteq V$ target vertices. For every $v \in V$ we have that

$$\sup_\sigma \inf_\pi \mathcal{P}_v^{\sigma,\pi}(Reach(T)) \quad = \quad \inf_\pi \sup_\sigma \mathcal{P}_v^{\sigma,\pi}(Reach(T))$$

Reachability games have a value (2)

Proof sketch.

- Let $\Gamma : [0,1]^{|V|} \to [0,1]^{|V|}$ be a (monotonic) function defined by

$$\Gamma(\alpha)(v) = \begin{cases} 1 & \text{if } v \in T; \\ \sup \{\alpha(v') \mid (v,v') \in E\} & \text{if } v \notin T \text{ and } v \in V_{\Box}; \\ \inf \{\alpha(v') \mid (v,v') \in E\} & \text{if } v \notin T \text{ and } v \in V_{\Diamond}; \\ \sum_{(v,v') \in E} Prob(v,v') \cdot \alpha(v') & \text{if } v \notin T \text{ and } v \in V_{\bigcirc}. \end{cases}$$

Reachability games have a value (2)

Proof sketch.

- Let $\Gamma : [0,1]^{|V|} \to [0,1]^{|V|}$ be a (monotonic) function defined by

$$\Gamma(\alpha)(v) = \begin{cases} 1 & \text{if } v \in T; \\ \sup\{\alpha(v') \mid (v,v') \in E\} & \text{if } v \notin T \text{ and } v \in V_\Box; \\ \inf\{\alpha(v') \mid (v,v') \in E\} & \text{if } v \notin T \text{ and } v \in V_\Diamond; \\ \sum_{(v,v')\in E} Prob(v,v') \cdot \alpha(v') & \text{if } v \notin T \text{ and } v \in V_\bigcirc. \end{cases}$$

- $\mu\Gamma(v) \;\leq\; \sup_\sigma \inf_\pi \mathcal{P}_v^{\sigma,\pi}(Reach(T)) \;\leq\; \inf_\pi \sup_\sigma \mathcal{P}_v^{\sigma,\pi}(Reach(T))$

Reachability games have a value (2)

Proof sketch.

- Let $\Gamma : [0,1]^{|V|} \to [0,1]^{|V|}$ be a (monotonic) function defined by

$$\Gamma(\alpha)(v) = \begin{cases} 1 & \text{if } v \in T; \\ \sup\,\{\alpha(v') \mid (v,v') \in E\} & \text{if } v \notin T \text{ and } v \in V_\Box; \\ \inf\,\{\alpha(v') \mid (v,v') \in E\} & \text{if } v \notin T \text{ and } v \in V_\Diamond; \\ \sum_{(v,v')\in E} Prob(v,v') \cdot \alpha(v') & \text{if } v \notin T \text{ and } v \in V_\bigcirc. \end{cases}$$

- $\mu\Gamma(v) \;\leq\; \sup_\sigma \inf_\pi \mathcal{P}_v^{\sigma,\pi}(Reach(T)) \;\leq\; \inf_\pi \sup_\sigma \mathcal{P}_v^{\sigma,\pi}(Reach(T))$
  - the second inequality holds for all Borel objectives;
  - the tuple of all $\sup_\sigma \inf_\pi \mathcal{P}_v^{\sigma,\pi}(Reach(T))$ is a fixed-point of Γ.

Reachability games have a value (2)

Proof sketch.

- Let $\Gamma : [0,1]^{|V|} \to [0,1]^{|V|}$ be a (monotonic) function defined by

$$\Gamma(\alpha)(v) = \begin{cases} 1 & \text{if } v \in T; \\ \sup \{\alpha(v') \mid (v,v') \in E\} & \text{if } v \notin T \text{ and } v \in V_{\Box}; \\ \inf \{\alpha(v') \mid (v,v') \in E\} & \text{if } v \notin T \text{ and } v \in V_{\Diamond}; \\ \sum_{(v,v') \in E} Prob(v,v') \cdot \alpha(v') & \text{if } v \notin T \text{ and } v \in V_{\bigcirc}. \end{cases}$$

- $\mu\Gamma(v) \;\leq\; \sup_{\sigma} \inf_{\pi} \mathcal{P}_v^{\sigma,\pi}(Reach(T)) \;\leq\; \inf_{\pi} \sup_{\sigma} \mathcal{P}_v^{\sigma,\pi}(Reach(T))$
  - the second inequality holds for all Borel objectives;
  - the tuple of all $\sup_{\sigma} \inf_{\pi} \mathcal{P}_v^{\sigma,\pi}(Reach(T))$ is a fixed-point of Γ.
- It cannot be that $\mu\Gamma(v) \;<\; \inf_{\pi} \sup_{\sigma} \mathcal{P}_v^{\sigma,\pi}(Reach(T))$
  - For all $\varepsilon > 0$ and $v \in V$, there is a strategy $\hat{\pi}$ such that $\sup_{\sigma} \mathcal{P}_v^{\sigma,\hat{\pi}}(Reach(T)) \;\leq\; \mu\Gamma(v) + \varepsilon$. □

Minimizing strategies (1)

Definition 9 (Locally optimal minimizing strategy)

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

- *An edge* $(v, v') \in E$ *is* ***value minimizing*** *if*

$$val(v') = \min \left\{ val(\hat{v}) \in V \mid (v, \hat{v}) \in E \right\}$$

- *A* ***locally optimal minimizing*** *strategy is a strategy which in every play selects only value minimizing edges.*

Minimizing strategies (2)

Theorem 10

Every locally optimal min. strategy is an optimal min. strategy.

Proof.

Let $v \in V$ be an initial vertex, and $u \in V$ a target vertex.

(1) After playing k rounds according to a locally optimal minimizing strategy, player $\Diamond$ can switch to ε-optimal minimizing strategies in the current vertices of the play. Thus, we always (for every k and $\varepsilon > 0$) obtain an ε-optimal minimizing strategy for v.

Minimizing strategies (2)

Theorem 10

Every locally optimal min. strategy is an optimal min. strategy.

Proof.

Let $v \in V$ be an initial vertex, and $u \in V$ a target vertex.

(1) After playing k rounds according to a locally optimal minimizing strategy, player $\diamond$ can switch to ε-optimal minimizing strategies in the current vertices of the play. Thus, we always (for every k and $\varepsilon > 0$) obtain an ε-optimal minimizing strategy for v.

(2) Let π be a locally optimal min. strategy which is not optimal.

- Then there is a strategy σ of player $\Box$ such that $\mathcal{P}_v^{\sigma,\pi}(Reach(T)) = val(v) + \delta$, where $\delta > 0$.
- This means that there is $k \in \mathbb{N}$ such that $\mathcal{P}_v^{\sigma,\pi}(Reach^k(T)) > val(v) + \frac{\delta}{2}$.
- Hence, if player $\diamond$ switches to $\frac{\delta}{4}$-optimal minimizing strategy after playing k rounds according to π, we do not obtain a $\frac{\delta}{4}$-optimal minimizing strategy for v. □

Minimizing strategies (3)

Corollary 11 (Properties of minimizing strategies.)

In every finitely-branching game, there is an optimal minimizing MD strategy.

Minimizing strategies (3)

Corollary 11 (Properties of minimizing strategies.)

In every finitely-branching game, there is an optimal minimizing MD strategy.

Theorem 12

Every optimal min. strategy is a locally optimal min. strategy. Hence, if player $\Diamond$ has some optimal minimizing strategy, then she also has an MD optimal minimizing strategy.

Minimizing strategies (3)

Corollary 11 (Properties of minimizing strategies.)

In every finitely-branching game, there is an optimal minimizing MD strategy.

Theorem 12

Every optimal min. strategy is a locally optimal min. strategy. Hence, if player $\Diamond$ has some optimal minimizing strategy, then she also has an MD optimal minimizing strategy.

Proof.

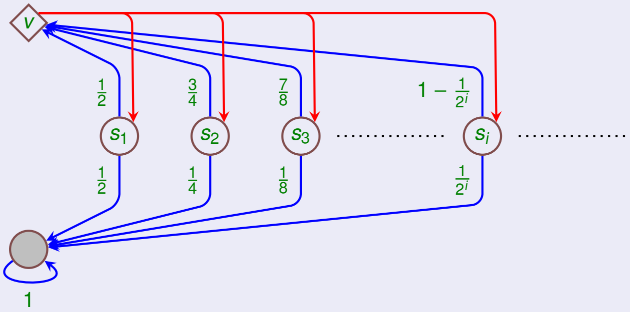
This is WRONG. Optimal minimizing strategies may require infinite memory. □

Minimizing strategies (4)

Theorem 13

Optimal minimizing strategies do not necessarily exist, and (ε-) optimal minimizing strategies may require infinite memory.

Proof.



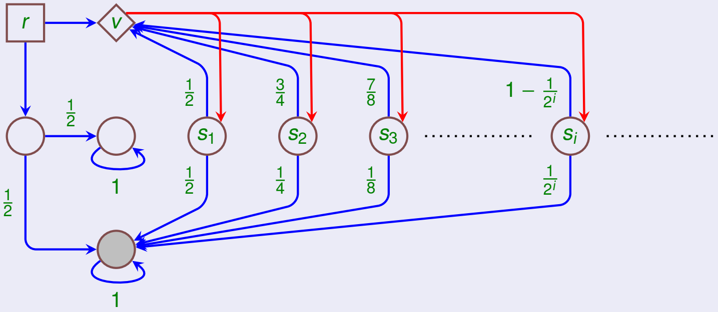
□

Minimizing strategies (4)

Theorem 13

Optimal minimizing strategies do not necessarily exist, and (ε-) optimal minimizing strategies may require infinite memory.

Proof.



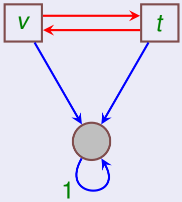
□

Maximizing strategies (1)

Observation 14

A locally optimal maximizing strategy is not necessarily an optimal maximizing strategy. This holds even for finite-state MDPs.

Proof.



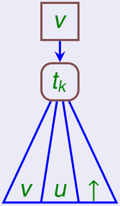
□

Maximizing strategies (2)

Theorem 15

Let $v \in V_\Box$ be a vertex with finitely many successors $t_1, \ldots, t_n$. Then there is $1 \leq i \leq n$ such that $val(v)$ does not change if all edges (v, t_j), where $i \neq j$, are deleted from the game.

Proof.



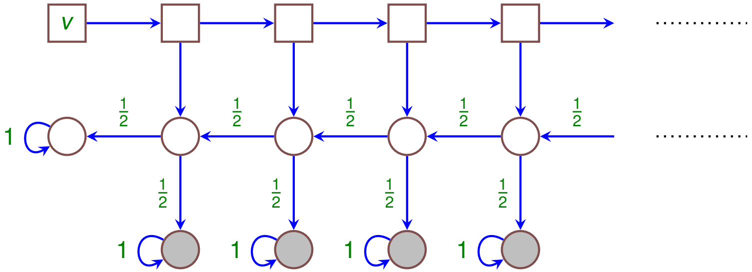
- $V_{t_k}^{(\sigma,\pi)} = \begin{cases} \frac{\mathcal{P}(u)}{\mathcal{P}(u)+\mathcal{P}(\uparrow)} & \text{if } \mathcal{P}(u) + \mathcal{P}(\uparrow) > 0; \\ 0 & \text{otherwise;} \end{cases}$
- $V_{t_k}^{\sigma} = \inf_\pi V_{t_k}^{(\sigma,\pi)}$
- $V_{t_k} = \sup_\sigma V_{t_k}^{\sigma}$
- There must be some k such that $V_{t_k} = val(v)$.
- We put $i = k$.

□

Maximizing strategies (3)

Theorem 16

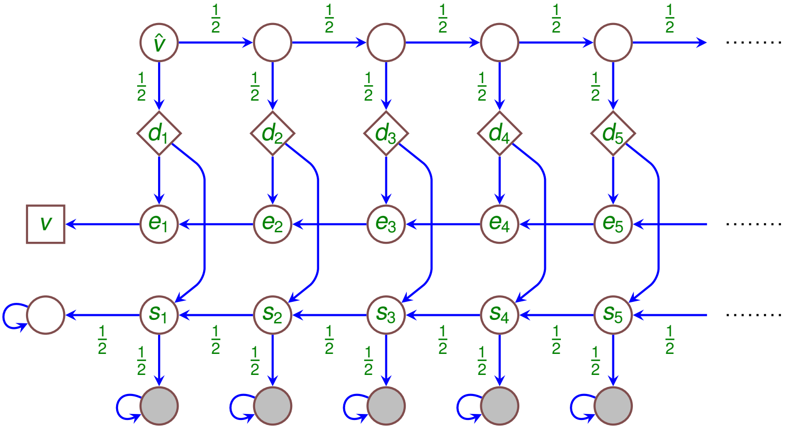
Optimal maximizing strategies may not exist, even in finitely-branching MDPs.



Maximizing strategies (4)

Theorem 17

Optimal maximizing strategies may require infinite memory, even in finitely-branching games.



Summary

Minimizing strategies:

- Optimal minimizing strategies may not exist. Optimal and ε-optimal minimizing strategies may require infinite memory.
- In finitely-branching games, there are MD optimal minimizing strategies.

Maximizing strategies:

- Optimal maximizing strategies may not exist, even in finitely-branching games. Optimal maximizing strategies may require infinite memory.
- In finite-state games, there are MD optimal maximizing strategies.

References:

- M.L. Puterman. *Markov Decision Processes*, Wiley, 1994.
- T. Brázdil, V. Brožek, V. Forejt, A. Kučera. *Reachability in recursive Markov decision processes.* Information and Computation, vol. 206, pp. 520–537, 2008.

Reachability as a win-lose objective (1)

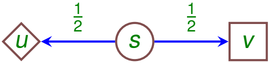
- Let $\varrho \in [0,1]$.
- A strategy $\sigma \in \Sigma$ is $(\geq\varrho)$-winning in v if for every $\pi \in \Pi$ we have that $\mathcal{P}_v^{(\sigma,\pi)}(Reach(T) \geq \varrho)$.
- A strategy $\pi \in \Pi$ is $(<\varrho)$-winning if for every $\sigma \in \Sigma$ we have that $\mathcal{P}_v^{(\sigma,\pi)}(Reach(T) < \varrho)$.
- Is there a winning strategy for one of the two players?

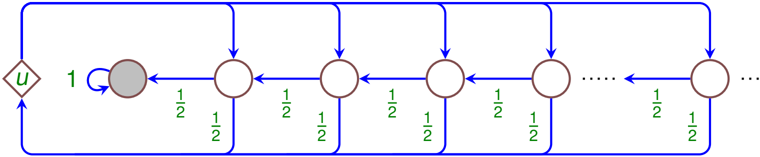
Reachability as a win-lose objective (2)

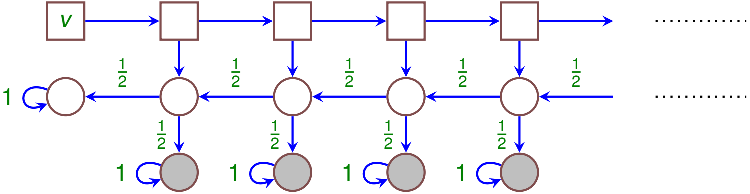
Theorem 18

Turn-based stochastic games with reachability objectives are not necessarily determined. However, finitely-branching games are determined.

Reachability as a win-lose objective (3)







Algorithms for finite-state MDPs and games

We show how to compute the values and optimal strategies for reachability objectives in finite-state games and MDPs.

- For finite-state MDPs we have that
  - the values and optimal strategies are computable in polynomial time;
- For finite-state games we have that
  - the values and optimal strategies are computable in polynomial space (for a fixed number of randomized vertices, the problem is in **P**);

Finite-state MDPs (1)

Theorem 19

Let $G = (V, E, (V_\Box, V_\bigcirc), Prob)$ be a finite-state MDP. Then

- $\mathcal{V}^{=0} = \{v \in V \mid val(v) = 0\}$
- $\mathcal{V}^{=1} = \{v \in V \mid val(v) = 1\}$

are computable in polynomial time.

Finite-state MDPs (1)

Theorem 19

Let $G = (V, E, (V_\Box, V_\bigcirc), Prob)$ be a finite-state MDP. Then

- $\mathcal{V}^{=0} = \{v \in V \mid val(v) = 0\}$
- $\mathcal{V}^{=1} = \{v \in V \mid val(v) = 1\}$

are computable in polynomial time.

Proof.

It suffices to realize that $\mathcal{V}^{=1}$ is exactly the greatest $S \subseteq V$ satisfying the following conditions:

- If $v \in S$, then there is a finite path from v to the target vertex which visits only the vertices of S.
- If $v \in S \cap V_\bigcirc$, then all successors of v belong to S.

Hence, $\mathcal{V}^{=1}$ is computable in polynomial time. The set $\mathcal{V}^{=0}$ can be computed similarly. Note that the sets $\mathcal{V}^{=1}$ and $\mathcal{V}^{=0}$ depend only on the "topology" of G. □

Finite-state MDPs (2)

Theorem 20

Let $G = (V, E, (V_\Box, V_\bigcirc), Prob)$ be a finite-state MDP where $Prob$ is rational. The values $val(v)$, $v \in V$, are rational and computable in polynomial time. An optimal maximizing strategy is also constructible in polynomial time.

Proof.

Let $V = \{v_1, \ldots, v_n\}$, where v_n is the (only) target vertex.

minimize $x_1 + \cdots + x_n$
subject to
$x_n = 1$
$x_i \geq x_j$ for all $(v_i, v_j) \in E$ where $v_i \in V_\Box$ and $i < n$
$x_i = \sum_{(v_i,v_j)\in E} Prob(v_i, v_j) \cdot x_j$ for all $v_i \in V_\bigcirc$, $i < n$
$x_i \geq 0$ for all $i \in \{1, \ldots, n\}$

An optimal strategy can be constructed by successively removing the ougoing edges of every $v \in V_\Box$ untill only one such edge is left. □

Finite-state games (1)

Theorem 21

Let $G = (V, E, (V_\square, V_\diamond, V_\bigcirc), Prob)$ be a finite-state game. Then

- $\mathcal{V}^{=0} = \{v \in V \mid val(v) = 0\}$
- $\mathcal{V}^{=1} = \{v \in V \mid val(v) = 1\}$

are computable in polynomial time.

Finite-state games (1)

Theorem 21

Let $G = (V, E, (V_\Box, V_\Diamond, V_\bigcirc), Prob)$ *be a finite-state game. Then*

- $\mathcal{V}^{=0} = \{v \in V \mid val(v) = 0\}$
- $\mathcal{V}^{=1} = \{v \in V \mid val(v) = 1\}$

are computable in polynomial time.

Proof.

- $\mathcal{V}^{>0} = \mu\Gamma$, where $\Gamma : 2^V \to 2^V$ is defined as follows:

$$\begin{array}{rcl}\Gamma(A) & = & T \cup \{v \in V_\Box \cup V_\bigcirc \mid \exists (v, v') \in E \text{ s.t. } v' \in A\} \\ & \cup & \{v \in V_\Diamond \mid \forall (v, v') \in E \text{ we have that } v' \in A\}\end{array}$$

Finite-state games (1)

Theorem 21

Let $G = (V, E, (V_\Box, V_\Diamond, V_\bigcirc), Prob)$ *be a finite-state game. Then*

- $\mathcal{V}^{=0} = \{v \in V \mid val(v) = 0\}$
- $\mathcal{V}^{=1} = \{v \in V \mid val(v) = 1\}$

are computable in polynomial time.

Proof.

- $\mathcal{V}^{>0} = \mu\Gamma$, where $\Gamma : 2^V \to 2^V$ is defined as follows:

$$\begin{array}{rcl} \Gamma(A) & = & T \cup \{v \in V_\Box \cup V_\bigcirc \mid \exists (v, v') \in E \text{ s.t. } v' \in A\} \\ & \cup & \{v \in V_\Diamond \mid \forall (v, v') \in E \text{ we have that } v' \in A\} \end{array}$$

- $\mathcal{V}^{=0} = V \setminus \mathcal{V}^{>0}$

Finite-state games (1)

Theorem 21

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

- $\mathcal{V}^{=0} = \{v \in V \mid val(v) = 0\}$
- $\mathcal{V}^{=1} = \{v \in V \mid val(v) = 1\}$

are computable in polynomial time.

Proof.

- $\mathcal{V}^{>0} = \mu\Gamma$, where $\Gamma : 2^V \to 2^V$ is defined as follows:

$$\begin{aligned} \Gamma(A) &= T \cup \{v \in V_\Box \cup V_\bigcirc \mid \exists (v, v') \in E \text{ s.t. } v' \in A\} \\ &\cup \{v \in V_\Diamond \mid \forall (v, v') \in E \text{ we have that } v' \in A\} \end{aligned}$$

- $\mathcal{V}^{=0} = V \setminus \mathcal{V}^{>0}$
- $\mathcal{V}^{<1} = \mu\Gamma$, where $\Gamma : 2^V \to 2^V$ is defined as follows:

$$\begin{aligned} \Gamma(A) &= \mathcal{V}^{=0} \cup \{v \in V_\Diamond \cup V_\bigcirc \mid \exists (v, v') \in E \text{ s.t. } v' \in A\} \\ &\cup \{v \in V_\Box \mid \forall (v, v') \in E \text{ we have that } v' \in A\} \end{aligned}$$

Finite-state games (1)

Theorem 21

Let $G = (V, E, (V_\Box, V_\Diamond, V_\bigcirc), Prob)$ *be a finite-state game. Then*

- $\mathcal{V}^{=0} = \{v \in V \mid val(v) = 0\}$
- $\mathcal{V}^{=1} = \{v \in V \mid val(v) = 1\}$

are computable in polynomial time.

Proof.

- $\mathcal{V}^{>0} = \mu\Gamma$, where $\Gamma : 2^V \to 2^V$ is defined as follows:

$$\begin{array}{rcl}\Gamma(A) & = & T \cup \{v \in V_\Box \cup V_\bigcirc \mid \exists (v,v') \in E \text{ s.t. } v' \in A\} \\ & \cup & \{v \in V_\Diamond \mid \forall (v,v') \in E \text{ we have that } v' \in A\}\end{array}$$

- $\mathcal{V}^{=0} = V \setminus \mathcal{V}^{>0}$
- $\mathcal{V}^{<1} = \mu\Gamma$, where $\Gamma : 2^V \to 2^V$ is defined as follows:

$$\begin{array}{rcl}\Gamma(A) & = & \mathcal{V}^{=0} \cup \{v \in V_\Diamond \cup V_\bigcirc \mid \exists (v,v') \in E \text{ s.t. } v' \in A\} \\ & \cup & \{v \in V_\Box \mid \forall (v,v') \in E \text{ we have that } v' \in A\}\end{array}$$

- $\mathcal{V}^{=1} = V \setminus \mathcal{V}^{<1}$ □

Finite-state games (2)

Theorem 22 (Anne Condon, 1992)

Let $G = (V, E, (V_\Box, V_\Diamond, V_\bigcirc), Prob)$ be a finite-state game. The problem whether $val(v) > \frac{1}{2}$ for a given $v \in V$ is in **NP** $\cap$ **coNP**.

Proof.

Since both players have optimal MD strategies, it suffices to

- "guess" an optimal MD strategy for player $\Box$ (or player $\Diamond$);
- compute the value in the resulting MDP by solving the associated linear program.

□

Obviously, $val(v)$ and the optimal strategies for both players are computable by exhaustive search.

Finite-state games (3)

Theorem 23 (Gimbert, Horn, 2008)

The values and MD optimal strategies in a finite-state game $G = (V, E, (V_\Box, V_\Diamond, V_\bigcirc), Prob)$ are computable in

$$\mathcal{O}\Big(|V_\bigcirc|! \cdot (\log(|V|)\,|E| + |p|)\Big)$$

time, where $|p|$ is the maximal bit-length of an edge probability.

Remark 24

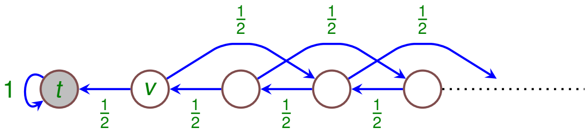
The question whether finite-state stochastic games are solvable in **P** is a longstanding open problem in algorithmic game theory.

References:

- A. Condon. *The Complexity of Stochastic Games.* Information and Computation, 96(2):203–224, 1992.
- L.S. Shapley. *Stochastic games.* Proceedings of the National Academy of Sciences USA, 39:1095–1100, 1953.
- H. Gimbert, F. Horn. *Simple Stochastic Games with Few Random Vertices Are Easy to Solve.* Proc. FoSSaCS 2008, pp. 5–19, LNCS 4962, Springer, 2008.

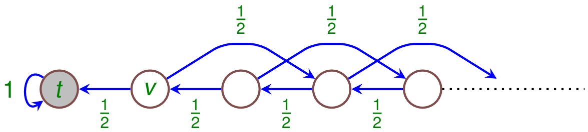
Infinite-state games

- Interesting classes of infinite-state stochastic games are obtained by extending non-deterministic computational devices with randomized choice. So far, most of the results consider
  - pushdown automata (recursive state machines);
  - lossy channel systems.
- There are some "new" problems:
  - The value can be irrational



Infinite-state games

- Interesting classes of infinite-state stochastic games are obtained by extending non-deterministic computational devices with randomized choice. So far, most of the results consider
  - pushdown automata (recursive state machines);
  - lossy channel systems.
- There are some "new" problems:
  - The value can be irrational



$val(v)$ is the least solution of $x = \frac{1}{2} + \frac{1}{2}x^3$ in $[0,1]$, i.e., $\frac{\sqrt{5}-1}{2}$

  - Even MD strategies may not be finitely representable.

Stochastic BPA games (1)

Definition 25

A stochastic BPA game is a tuple $\Delta = (\Gamma, \hookrightarrow, (\Gamma_\Box, \Gamma_\Diamond, \Gamma_\bigcirc), Prob)$ *where*

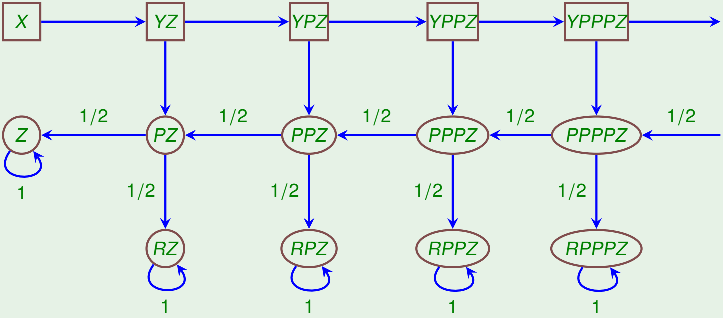
- Γ *is a finite stack alphabet,*
- $\hookrightarrow \subseteq \Gamma \times \Gamma^{\leq 2}$ *is a finite set of rules,*
- $(\Gamma_\Box, \Gamma_\Diamond, \Gamma_\bigcirc)$ *is a partition of* Γ*,*
- $Prob$ *is a probability assignment which to each* $X \in \Gamma_\bigcirc$ *assigns a rational positive probability distribution on the set of all rules of the form* $X \hookrightarrow \alpha$*.*

Stochastic BPA games (2)

Example 26

Let $\Gamma = \{X, Y, Z, P, R\}$, where $\Gamma_\Box = \{X, Y\}$, $\Gamma_\Diamond = \emptyset$, $\Gamma_\bigcirc = \{P, R, Z\}$, and

$$X \hookrightarrow YZ,\ Y \hookrightarrow YP,\ Y \hookrightarrow P,\ P \overset{1/2}{\hookrightarrow} R,\ P \overset{1/2}{\hookrightarrow} \varepsilon,\ R \overset{1}{\hookrightarrow} R,\ Z \overset{1}{\hookrightarrow} Z$$



BPA MDPs with reachability objectives (1)

- Let $\Delta = (\Gamma, \hookrightarrow, (\Gamma_\Box, \Gamma_\bigcirc), Prob)$ be a BPA Markov decision process, and $T \subseteq \Gamma^*$ a regular set of target configurations.
- Consider the sets
  - $\mathcal{W}^{>0} = \{\alpha \in \Gamma^* \mid \exists \sigma : \mathcal{P}^\sigma_\alpha(Reach(T)) > 0\}$
  - $\mathcal{W}^{=0} = \{\alpha \in \Gamma^* \mid \exists \sigma : \mathcal{P}^\sigma_\alpha(Reach(T)) = 0\}$
  - $\mathcal{W}^{=1} = \{\alpha \in \Gamma^* \mid \exists \sigma : \mathcal{P}^\sigma_\alpha(Reach(T)) = 1\}$
  - $\mathcal{W}^{<1} = \{\alpha \in \Gamma^* \mid \exists \sigma : \mathcal{P}^\sigma_\alpha(Reach(T)) < 1\}$

  These sets are regular and the associated finite-state automata are computable in polynomial time. The corresponding winning strategies are regular and computable in polynomial time.
- Similar results hold for BPA games.

BPA MDPs with reachability objectives (2)

References:

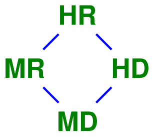
- K. Etessami, M. Yannakakis. *Efficient Qualitative Analysis of Classes of Recursive Markov Decision Processes and Simple Stochastic Games.* Proc. STACS 2006, pp. 634–645, LNCS 3884, Springer 2006.
- T. Brázdil, V. Brožek, A. Kučera, and J. Obdržálek. *Qualitative Reachability in Stochastic BPA Games.* Proc. STACS 2009, pp. 207–218, 2009.

Branching-time winning objectives

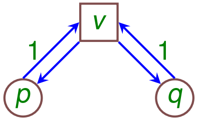
- Specified by formulae of branching-time logics that are interpreted over Markov chains (such as PCTL or PCTL*).
- $\mathcal{G}^{=1}(p \Rightarrow \mathcal{F}^{\geq 0.1} q)$
- The aim of player $\Box$ and player $\Diamond$ is to satisfy and falsify a given formula, respectively.

Properties of games with b.-t. objectives (I)

- Memory and randomization help:



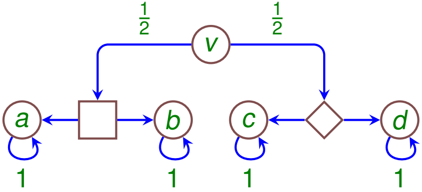
- Consider the following game:



- $\mathcal{X}^{=1}p \wedge \mathcal{F}^{=1}q$. 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 b.-t. objectives (II)

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



Who wins the game (MD strategies) ?

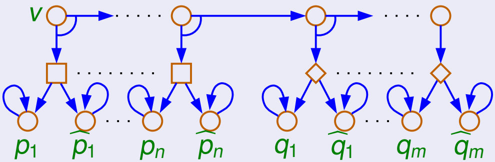
Theorem 27 (Brázdil, Brožek, Forejt, K., 2006)

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

Proof.

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. $\Box$

Who wins the game (MR strategies) ?

Theorem 28 (Brázdil, Brožek, Forejt, K., 2006)

*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.*

Proof.

- The Σ_2-hardness is established similarly as for MD strategies.
- The membership to **EXPTIME** is obtained by encoding the condition into Tarski algebra.
- The membership to Σ_2 for the qualitative PCTL follows easily.

$\Box$

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

Theorem 29 (Brázdil, Brožek, Forejt, K., 2006)

The existence of a winning HD (or HR) strategy for player □ 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.

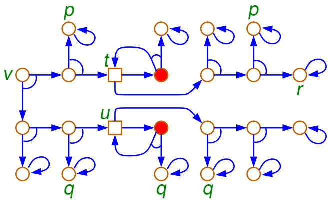
The undecidability proof

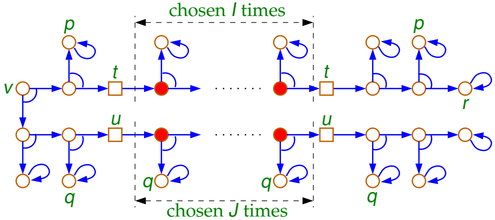
- A non-deterministic Minsky machine $\mathcal{M}$ with two counters c_1, c_2:

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

  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 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 $G(\mathcal{M})$.

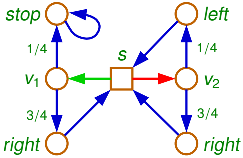
The construction of $G(\mathcal{M})$ and φ

- 

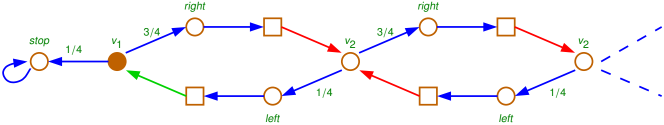
- 

- $I = J < \omega$ 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 (1)

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

$$\mathcal{G}^{>0}(\neg stop \wedge \mathcal{F}^{>0} stop)$$
$$\wedge\ \mathcal{G}^{=1}(s \Rightarrow (X^{=1} v_1 \vee X^{=1} v_2))$$

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

Positive result (2)

Theorem 30 (Brázdil, Forejt, K., 2008)

- *The existence of a winning HD (or HR) strategy for player □ 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. The problem is **2-EXPTIME**-hard.*
- *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).*

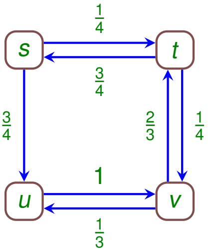
References:

- T. Brázdil, V. Brožek, V. Forejt, and A. Kučera. *Stochastic Games with Branching-Time Winning Objectives.* Proc. of LICS 2006, pp. 349-358, 2006.
- T. Brázdil, V. Forejt, and A. Kučera. *Controller Synthesis and Verification for Markov Decision Processes with Qualitative Branching Time Objectives.* Proc. of ICALP 2008, pp. 148-159, volume 5126 of LNCS. Springer, 2008.

Games with time

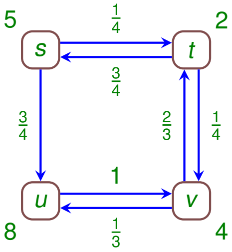
- Games over continuous-time stochastic processes such as
  - continuous-time Markov chains;
  - semi-Markov processes;
  - generalized semi-Markov processes.
- Time-dependent objectives such as
  - time-bounded reachability;
  - properties expressible in temporal logics with time;
  - properties encoded by timed automata.

Continuous-time Markov chains (1)

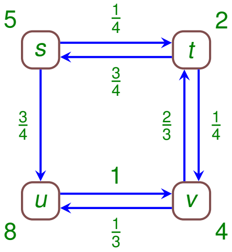


- The probability that a transition occurs in a state s before time $t > 0$ is equal to $1 - e^{-\lambda_s t}$.
- A timed run is an infinite sequence $s_0, t_0, s_1, t_1, \ldots$ where $s_0, s_1, \ldots$ is a run and $t_i \in \mathbb{R}^{\geq 0}$.

Continuous-time Markov chains (1)



Continuous-time Markov chains (1)



- The probability that a transition occurs in a state s before time $t > 0$ is equal to $1 - e^{-\lambda_s t}$.
- A timed run is an infinite sequence $s_0, t_0, s_1, t_1, \ldots$ where $s_0, s_1, \ldots$ is a run and $t_i \in \mathbb{R}^{\geq 0}$.

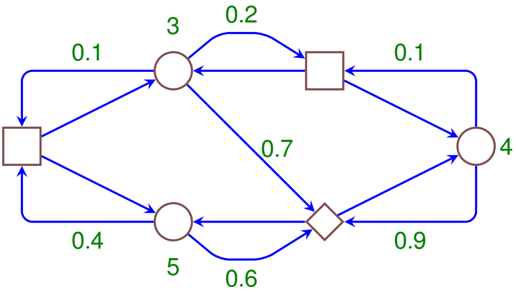
Continuous-time Markov chains (2)

- For every timed cylinder $w = s_0, I_0, \ldots, s_{n-1}, I_{n-1}, s_n$ we put

$$\mathcal{P}(w) = \prod_{i=0}^{n-1} Prob(s_i, s_{i+1}) \cdot \int_{I_i} \lambda_{s_i} e^{-\lambda_{s_i} x} dx$$

- This assignment can be uniquely extended to the (Borel) σ-algebra $\mathcal{F}$ generated by all timed cylinders.
- Thus, we obtain the probability space $(TRun(s), \mathcal{F}, \mathcal{P})$.

Continuous-time stochastic games



- A strategy of player $\odot$ assigns to each timed history wv (where $v \in V_\odot$) a probability distribution over the outgoing edges of v.
- In general, a play is a Markov process with uncountable state-space.
- Time abstract strategies do not depend on time stamps, and the corresponding play is a continuous-time Markov chain.

Time-bounded reachability objectives (1)

- The objective of player □/◇ is to maximize/minimize the probability of reaching a target vertex before a time bound t.
- Continuous-time stochastic games with time-bounded reachability objectives have a value (w.r.t. time abstract strategies), i.e.,

$$\sup_{\sigma} \inf_{\pi} \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}^{\leq t}(T)) \quad = \quad \inf_{\pi} \sup_{\sigma} \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}^{\leq t}(T))$$

- An optimal strategy for player ◇ exists in finitely-branching CTSGs.
- An optimal strategy for player □ exists in finitely-branching CTSGs with bounded rates.
- In finite uniform CTGs, both players have FD optimal strategies which are effectively computable.

Time-bounded reachability objectives (2)

References:

- C. Baier, H. Hermanns, J.-P. Katoen, and B.R. Haverkort. *Efficient computation of time-bounded reachability probabilities in uniform continuous-time Markov decision processes.* Theoretical Computer Science, 345:2–26, 2005.
- M. Neuhäußer, M. Stoelinga, and J.-P. Katoen. *Delayed nondeterminism in continuous-time Markov decision processes.* In Proceedings of FoSSaCS 2009, volume 5504 of LNCS, pages 364–379. Springer, 2009.
- T. Brázdil, V. Forejt, J. Krčál, J. Křetínský, and A. Kučera. *Continuous-time stochastic games with time-bounded reachability.* In Proceedings of FST&TCS 2009, pages 61–72, 2009.
- M. Rabe and S. Schewe. *Optimal time-abstract schedulers for CTMDPs and Markov games.* In Eighth Workshop on Quantitative Aspects of Programming Languages, 2010.
- T. Brázdil, J. Krčál, J. Křetínský, A. Kučera, and V. Řehák. *Stochastic Real-Time Games with Qualitative Timed Automata Objectives.* In Proceedings of Concur 2010, to appear.

Open problems

MANY…

- Games with continuous time.
- Hybrid games.
- Games with multiple players (cooperative games).
- Games over infinite-state computational models.