Stochastic Games with Time

Antonín Kučera

Masaryk University Brno

Outline

- Event-driven stochastic processes.
- Games over event-driven stochastic processes.
  - Strategies, objectives, determinacy.
- Existing results for games with reachability objectives.
  - Games over discrete-time Markov chains.
  - Games over continuous-time Markov chains.
  - Games over event-driven stochastic processes.
- Conclusions.

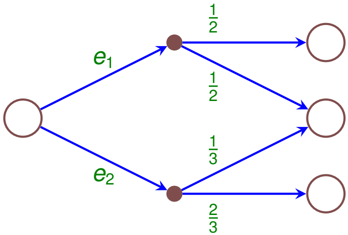
Stochastic games in formal verification

In general:

- the arena (game board) corresponds to the state-space of a given system;
- a state reacts to some events whose impact in uncertain;
- there are special control states where two players, controller and environment can choose some action whose impact may also be uncertain;

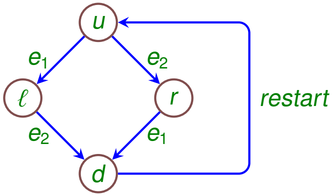
Is there a strategy for the controller such that the system satisfies a certain property no matter what the environment does?

Event-driven stochastic processes



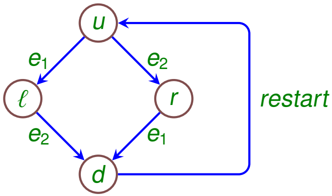
- state-space can be (countably) infinite;
- each event is either discrete-time or continuous-time;
- this model is closely related to real-time probabilistic processes [Alur, Courcoubetis, Dill] and generalized semi-Markov processes.

Event-driven stochastic processes (2)



- Suppose that e_1 and e_2 have densities f_1 and f_2 and *restart* is a discrete-time event which takes zero time.
- The event e_2 awaited in ℓ has actually been fired in u. How do we capture this formally? What is the semantics of a given event-driven stochastic process G ?

Event-driven stochastic processes (3)



- A fully rigorous approach: define the associated Markov process M_G (with uncountable state-space).
  - Usually, the state-space of M_G is formed by tuples of the form $(s, t_1, \ldots, t_n)$. This is not appropriate in our setting.
  - Alternatively, the state-space of M may consist of computational histories of G.
- A "lightweight" approach: define a suitable probability space over the runs in G.

Event-driven stochastic processes (4)

- A run of a given event-driven stochastic process is an infinite sequence

$$(s_0, t_0, e_0), (s_1, t_1, e_1), (s_2, t_2, e_2), \ldots$$

where t_i is the time spent in s_i and e_i is the triggering event.

- A basic cylinder determined by a finite sequence

$$(s_0, I_0, e_0), \ldots (s_n, I_n, e_n)$$

consists of all runs of the form

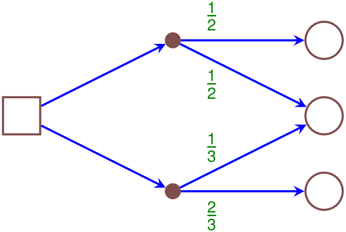
$$(s_0, t_0, e_0), \ldots (s_n, t_n, e_n), \ldots$$

where $t_i \in I_i$ for all $0 \leq i \leq n$.

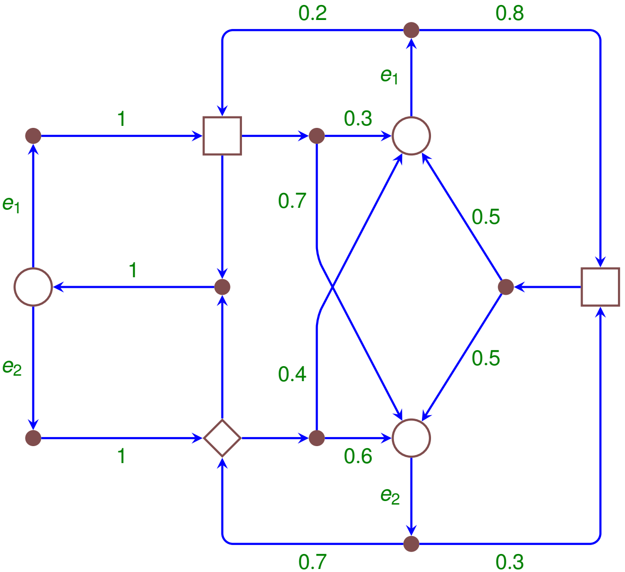
- We define the probability of basic cylinders in the natural way. Thus, we obtain the probability space $(Run, \mathcal{F}, \mathcal{P})$.

Games over event-driven stochastic processes

We add special control states $V_\square \uplus V_\Diamond$ where player $\square$ and player $\Diamond$ can choose successor states. The impact of this choice may be uncertain in general. Players' decisions are timeless.



Games over event-driven stochastic processes (2)



Games over event-driven stochastic processes (3)

- A history of a game G is a finite sequence of the form

$$(s_0, t_0, e_0, v_0), \ldots, (s_n, t_n, e_n, v_n)$$

- A strategy of player $\odot$, where $\odot \in \{\Box, \Diamond\}$, is a measurable function which to every history

$$(s_0, t_0, e_0, v_0), \ldots, (s_n, t_n, e_n, v_n)$$

such that $v_n \in V_\odot$ assigns a probability distribution over the set of actions enabled in v_n.

- Let μ_0 be an initial probability distribution over the set of states of G. Then each pair σ, π of strategies for player $\Box$ and player $\Diamond$ determines a unique play of G, denoted by $G^{\sigma,\pi}$, which is an event-driven stochastic process.

Games over event-driven stochastic processes (4)

- A run in a game G is an infinite sequence of the form

$$(s_0, t_0, e_0, v_0), \ldots, (s_n, t_n, e_n, v_n), \ldots$$

- One can define (Borel) σ-algebra over the runs of G which is the least σ-algebra containing all basic cylinders.
- A Borel objective is Borel set of runs in G.
- Various Borel objectives are definable by timed automata and linear-time logics.

Games over event-driven stochastic processes (5)

Theorem 1

Games over event-driven stochastic processes with Borel objectives have a value. That is,

$$\sup_{\sigma} \inf_{\pi} P^{\sigma,\pi}(R) \quad = \quad \inf_{\pi} \sup_{\sigma} P^{\sigma,\pi}(R)$$

where $R \subseteq Run$ is Borel.

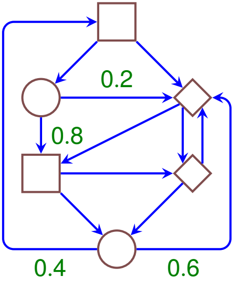
- Thm. 1 follows directly from the result of Maitra & Sudderth [1998] (which relies on Martin's determinacy result for Blackwell games).
- Thm. 1 implies the existence of ε-optimal strategies for both players, but not the existence of optimal strategies.
- One can use various formalisms (e.g., timed automata) to construct finite representations of time-dependant strategies.

Reachability objectives

- Let G be a game and T a set of target nodes.
  - $reach(T)$ consists of all runs that visit a target node.
  - $reach^{\leq t}(T)$ consists of all runs that visit a target node before time t.
  - The goal of player $\Box/\Diamond$ is to maximize/minimize the probability of $reach^{\leq t}(T)$ (or $reach(T)$).
- The problems of our interest.
  - Do the players have optimal strategies? And of what type?
  - Can we compute the value and ε-optimal strategies?
  - How about win-lose objectives of the form $\mathcal{P}^{\geq \varrho}(reach(T))$? Are such games determined? If so, what is the type of winning strategies?

Reachability games over Markov chains

- Also known as simple stochastic games.
- There is only one discrete-time event e with delay 1.
- Recall that the state-space can be infinite.



Reachability games have a value

Theorem 2

Let $G = (V, E, (V_\square, V_\diamond, V_\bigcirc), Prob)$ 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

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

- $\mu\Gamma(v) \;\leq\; \sup_{\sigma} \inf_{\pi} \mathcal{P}_v^{\sigma,\pi}(reach(T)) \;\leq\; \inf_{\pi} \sup_{\sigma} \mathcal{P}_v^{\sigma,\pi}(reach(T))$
  - 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 3 (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 4

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 5 (Properties of minimizing strategies.)

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

Theorem 6

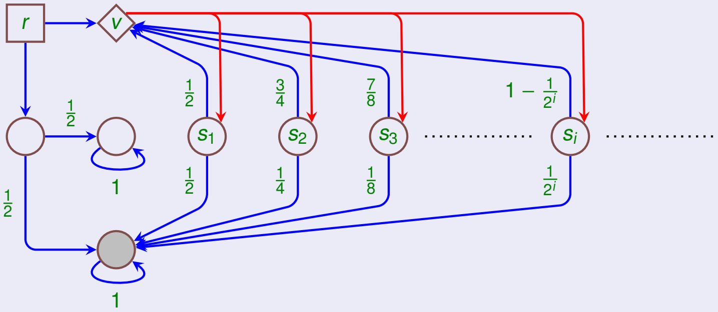
Every optimal min. strategy is a locally optimal min. strategy. Hence, if player $\Diamond$ has some optimal minimizing strategy, then she also has an MD optimal minimizing strategy.

Minimizing strategies (4)

Theorem 7

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

Proof.



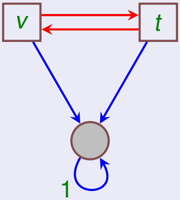
□

Maximizing strategies (1)

Observation 8

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

Proof.



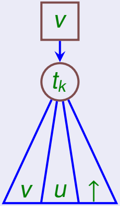
□

Maximizing strategies (2)

Theorem 9

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.



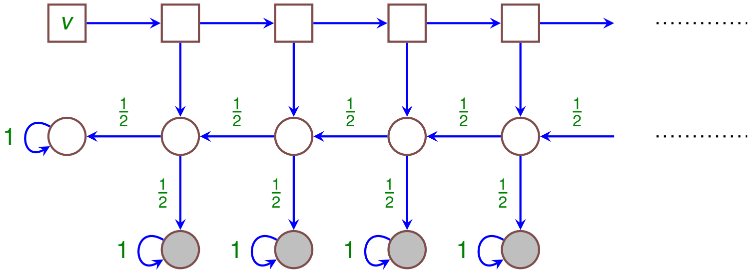
- $V_{t_k}^{(\sigma,\pi)} = \begin{cases} \frac{\mathcal{P}(u)}{\mathcal{P}(u)+\mathcal{P}(\uparrow)} & \text{if } \mathcal{P}(u)+\mathcal{P}(\uparrow) > 0; \\ 0 & \text{otherwise;} \end{cases}$
- $V_{t_k}^{\sigma} = \inf_\pi V_{t_k}^{(\sigma,\pi)}$
- $V_{t_k} = \sup_\sigma V_{t_k}^{\sigma}$
- There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

□

Maximizing strategies (3)

Theorem 10

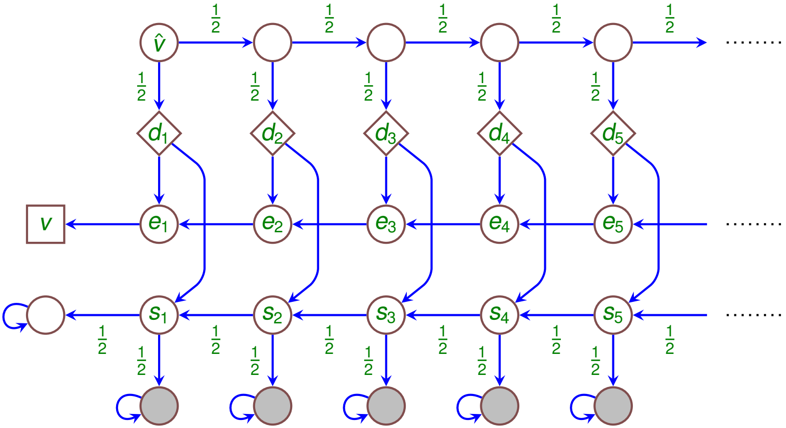
Optimal maximizing strategies may not exist, even in finitely-branching MDPs.



Maximizing strategies (4)

Theorem 11

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.

Reachability as a win-lose objective (1)

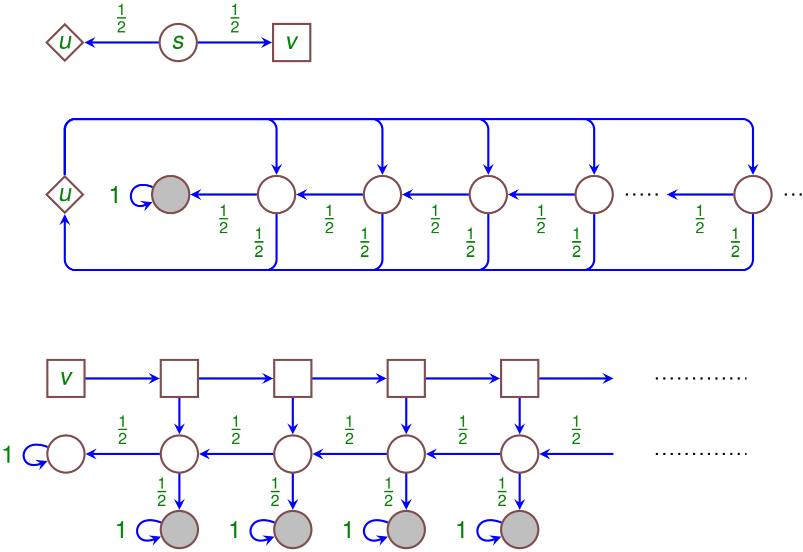
- Let $\varrho \in [0,1]$.
- A strategy $\sigma \in \Sigma$ is $(\geq\varrho)$-winning in v if for every $\pi \in \Pi$ we have that $\mathcal{P}_v^{(\sigma,\pi)}(\mathit{reach}(T) \geq \varrho)$.
- A strategy $\pi \in \Pi$ is $(<\varrho)$-winning if for every $\sigma \in \Sigma$ we have that $\mathcal{P}_v^{(\sigma,\pi)}(\mathit{reach}(T)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

Reachability as a win-lose objective (2)

Theorem 12

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 MDP 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 by linear programming;
- 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** [Gimbert, Horn, 2008]);
- There are also algorithms for certain classes of infinite-state games.

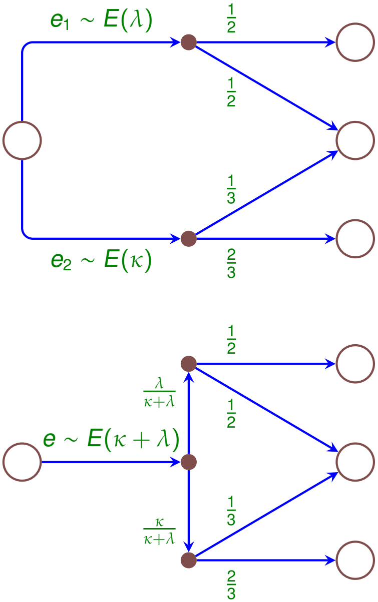
References

- 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.
- M.L. Puterman. *Markov Decision Processes*, Wiley, 1994.
- T. Brázdil, V. Brožek, V. Forejt, A. Kučera. *Reachability in recursive Markov decision processes.*
- A. Kučera. *Turn-based Stochastic Games.* In Lectures in Game Theory for Computer Scientists, Cambridge. To appear.
- 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.

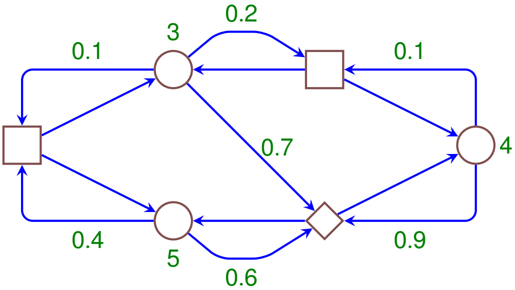
Games over cont.-time Markov chains

- All events are continuous and exponentially distributed, i.e., $\mathcal{P}(d_e \leq t) = 1 - \mathbf{e}^{-\lambda t}$ where $\lambda > 0$ is a rate of the event e.
- For simplicity, we assume that the impact of players' choice is determined (i.e., the distributions associated to the available actions are Dirac).
- Some facts about exponential distribution:
  - Let $X \sim E(\lambda)$. Then $\mathcal{P}(X \geq t + t' \mid X \geq t') = \mathcal{P}(X \geq t)$.
  - Let $X \sim E(\lambda)$ and $Y \sim E(\kappa)$. Then $\min(X, Y) \sim E(\lambda + \kappa)$ and $\mathcal{P}(X < Y) = \lambda/(\lambda + \kappa)$.

Games over cont.-time Markov chains (2)



Games over cont.-time Markov chains (3)



Games over cont.-time Markov chains (4)

Stochastic games with time-bounded reachability objectives have been so far studied mainly for time abstract strategies.

Theorem 13

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

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

where σ and π range over time abstract strategies.

Games over cont.-time Markov chains (5)

Proof sketch.

- Let $\mathcal{H}$ be the set of all histories $\mathbf{i} : \mathcal{R}(G) \to \mathbb{N}_0$ where $\sum_{a \in \mathcal{R}(G)} \mathbf{i}(a) < \infty$.
- Let $\Gamma : (\mathcal{H} \times V \to [0,1]) \to (\mathcal{H} \times V \to [0,1])$ be a (monotonic) function defined by

$$\Gamma(H)(\mathbf{i}, v) = \begin{cases} \mathbf{F}_{\mathbf{i}}(t) & v \in T \\ \sup_{a \in E(v)} \sum_{u \in V} P(a)(u) \cdot H(\mathbf{i} + \mathbf{1}_{Rate(a)}, u) & v \in V_{\Box} \setminus T \\ \inf_{a \in E(v)} \sum_{u \in V} P(a)(u) \cdot H(\mathbf{i} + \mathbf{1}_{Rate(e)}, u) & v \in V_{\Diamond} \setminus T \end{cases}$$

- Let $\mu\Gamma$ be the least fixed-point of Γ. The value of a given vertex v is equal to $\mu\Gamma(\mathbf{0}, v)$.
- Observation: as $\sum_{a \in \mathcal{R}(G)} \mathbf{i}(a)$ increases, $\mathbf{F}_{\mathbf{i}}(t)$ approaches zero (assuming the rates are bounded). Hence, Γ allows to compute ε-optimal strategies.

Games over cont.-time Markov chains (6)

- In general, optimal strategies do not exist.
- In finitely-branching games, player $\diamond$ is guaranteed to have an optimal CD strategy.
- In finitely-branching games with bounded rates, player $\square$ is guaranteed to have an optimal CD strategy.
- In finitely-branching uniform games, both players have BCD optimal strategies that are effectively computable.

Games over cont.-time Markov chains (7)

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.

Games over event-driven stochastic processes

- Assuming that all events are continuous and the objective is encoded as a deterministic timed automaton, one can decide if player $\Box$ has an almost-sure winning strategy and compute a finite description of this strategy [Brázdil et al., Concur 2010].
- It is not easy to extend this result to the general model with both continuous and discrete events.

Conclusions, open problems

- Games over event-driven stochastic processes can model concurrent systems with stochastic delays that are not necessarily exponentially distributed.
- One can rely on rich theory of discrete-time stochastic games and Markov processes with general state-space.
- Almost everything is open. New theoretical results can also bring efficient algorithms for solving the considered problems.