Continuous-Time Stochastic Games with Time-Bounded Reachability

T. Brázdil, V. Forejt, J. Krčál, J. Křetínský, A. Kučera

Faculty of Informatics, Masaryk University, Botanická 78a, 70200 Brno, Czech Republic
{brazdil,forejt,krcal,xkretins,kucera}@fi.muni.cz

MEMICS, 2009

accepted to FSTTCS 2009
Discrete-time stochastic game – example

- 2 players – □ and ◊
- in each vertex – a set of enabled actions
- for each action a probability distribution on the set of vertices

goals of the players
- player □ – reach the target vertex
- player ◊ – avoid it
Discrete-time stochastic game – example

- 2 players – □ and ◊
- in each vertex – a set of enabled actions
- for each action a probability distribution on the set of vertices

goals of the players

- player □ – reach the target vertex
- player ◊ – avoid it
Discrete-time stochastic game – example

- 2 players – □ and ◊
- in each vertex – a set of enabled actions
- for each action a probability distribution on the set of vertices

**goals of the players**
- player □ – reach the target vertex
- player ◊ – avoid it
Discrete-time stochastic game – example

- 2 players – □ and ◊
- in each vertex – a set of enabled actions
- for each action a probability distribution on the set of vertices

goals of the players
- player □ – reach the target vertex
- player ◊ – avoid it
Discrete-time stochastic game – example

- 2 players – □ and ◊
- in each vertex – a set of enabled actions
- for each action a probability distribution on the set of vertices

goals of the players
- player □ – reach the target vertex
- player ◊ – avoid it
Discrete-time stochastic game – example

- 2 players – □ and ◊
- in each vertex – a set of enabled actions
- for each action a probability distribution on the set of vertices

goals of the players
- player □ – reach the target vertex
- player ◊ – avoid it

Brázdíl, Forejt, Krčál, Křetínský, Kučera: Faculty of Informatics, Masaryk University
Continuous-Time Stochastic Games with Time-Bounded Reachability
Discrete-time stochastic game – example

- 2 players – □ and ◊
- in each vertex – a set of enabled actions
- for each action a probability distribution on the set of vertices

goals of the players
- player □ – reach the target vertex
- player ◊ – avoid it
Discrete-time stochastic game – example

- 2 players – □ and ◊
- in each vertex – a set of enabled actions
- for each action a probability distribution on the set of vertices

goals of the players
- player □ – reach the target vertex
- player ◊ – avoid it
Discrete-time stochastic game – example

- 2 players – □ and ◊
- in each vertex – a set of enabled actions
- for each action a probability distribution on the set of vertices

goals of the players

- player □ – reach the target vertex
- player ◊ – avoid it
Discrete-time stochastic game – example

- 2 players – ■ and ♦
- in each vertex – a set of enabled actions
- for each action a probability distribution on the set of vertices

goals of the players
- player ■ – reach the target vertex
- player ♦ – avoid it

Probability that the player ■ wins for strategies \( \sigma \) and \( \pi \) denoted by:
\[
P^{\sigma, \pi}_{\text{boss-has-no-time}}\left(\text{Reach}\{\text{student-prepares-slides}\}\right)
\]
Continuous-time stochastic game – example

now with time
the goal of the player □ – reach the target vertex in time 0.5

- performing an action takes time according to an exponential distribution
- each action has a rate - the parameter of the distribution
- rate ~ how many such events occur in one time unit
Continuous-time stochastic game – example

now with time
the goal of the player □ – reach the target vertex in time 0.5

- performing an action takes time according to an exponential distribution
- each action has a rate - the parameter of the distribution
- rate ∼ how many such events occur in one time unit
Continuous-time stochastic game – example

Now with time

The goal of the player | Reach the target vertex in time 0.5

- Performing an action takes time according to an exponential distribution
- Each action has a rate - the parameter of the distribution
- Rate $\sim$ How many such events occur in one time unit

Brázdíl, Forejt, Krčál, Křetínský, Kučera: Continuous-Time Stochastic Games with Time-Bounded Reachability
Continuous-time stochastic game – example

now with time

the goal of the player □ – reach the target vertex in time 0.5

- performing an action takes time according to an exponential distribution
- each action has a rate - the parameter of the distribution
- rate ~ how many such events occur in one time unit
Continuous-time stochastic game – example

now with time
the goal of the player □ – reach the target vertex in time 0.5

▶ performing an action takes time according to an exponential distribution
▶ each action has a rate - the parameter of the distribution
▶ rate ~ how many such events occur in one time unit
Continuous-time stochastic game – example

now with time
the goal of the player □ – reach the target vertex in time 0.5

► performing an action takes time according to an exponential distribution
► each action has a rate - the parameter of the distribution
► rate ~ how many such events occur in one time unit
Value of a game

\[ \inf_{\pi \in \Pi} P_v^{\sigma,\pi}(\text{Reach} \leq t(T)) \] for a fixed strategy \( \sigma \)
Value of a game

\[ \{ \inf_{\pi \in \Pi} P^\sigma_\pi(\text{Reach} \leq t(T)) \mid \sigma \in \Sigma \} \]

Theorem
Every CTG G has a value in every vertex.
Value of a game

\[
\{ \inf_{\pi \in \Pi} P^\sigma_\pi(\text{Reach} \leq t(T)) \mid \sigma \in \Sigma \}
\]

\[
\sup_{\sigma \in \Sigma} \inf_{\pi \in \Pi} P^\sigma_\pi(\text{Reach} \leq t(T))
\]
Value of a game

\[ \{ \inf_{\pi \in \Pi} P^{\sigma,\pi}_v(\text{Reach} \leq t(T)) \mid \sigma \in \Sigma \} \quad \{ \sup_{\sigma \in \Sigma} P^{\sigma,\pi}_v(\text{Reach} \leq t(T)) \mid \pi \in \Pi \} \]

\[ \sup_{\sigma \in \Sigma} \inf_{\pi \in \Pi} P^{\sigma,\pi}_v(\text{Reach} \leq t(T)) \quad \inf_{\pi \in \Pi} \sup_{\sigma \in \Sigma} P^{\sigma,\pi}_v(\text{Reach} \leq t(T)) \]
Value of a game

\[
\begin{align*}
\{ \inf_{\pi \in \Pi} \mathcal{P}_v^{\sigma,\pi}(\text{Reach} \leq t(T)) & \mid \sigma \in \Sigma \} \ ? \ \{ \sup_{\sigma \in \Sigma} \mathcal{P}_v^{\sigma,\pi}(\text{Reach} \leq t(T)) & \mid \pi \in \Pi \} \\
\sup_{\sigma \in \Sigma} \inf_{\pi \in \Pi} \mathcal{P}_v^{\sigma,\pi}(\text{Reach} \leq t(T)) & = \inf_{\pi \in \Pi} \sup_{\sigma \in \Sigma} \mathcal{P}_v^{\sigma,\pi}(\text{Reach} \leq t(T))
\end{align*}
\]

The game has a value in vertex \( v \), denoted by \( \text{val}(v) \), if

\[ \sup_{\sigma \in \Sigma} \inf_{\pi \in \Pi} \mathcal{P}_v^{\sigma,\pi}(\text{Reach} \leq t(T)) = \inf_{\pi \in \Pi} \sup_{\sigma \in \Sigma} \mathcal{P}_v^{\sigma,\pi}(\text{Reach} \leq t(T)) \]
Value of a game

The game has a value in vertex $v$, denoted by $val(v)$, if

$$\sup_{\sigma \in \Sigma} \inf_{\pi \in \Pi} P_{\sigma, \pi}(Reach \leq t(T)) = \inf_{\pi \in \Pi} \sup_{\sigma \in \Sigma} P_{\sigma, \pi}(Reach \leq t(T))$$

**Theorem**

*Every CTG $G$ has a value in every vertex.*
Optimal strategies

Every CTG $G$ has a value!
Optimal strategies

Every CTG $G$ has a value!

- Hence, there are $\epsilon$-optimal strategies!

For every $\epsilon > 0$ exists a strategy $\sigma_\epsilon$

\[
\inf_{\pi \in \Pi} \mathcal{P}_{\sigma_\epsilon, \pi} (\text{Reach} \leq t(T)) \geq \text{val}(v) - \epsilon
\]
Optimal strategies

Every CTG $G$ has a value!

- Hence, there are $\epsilon$-optimal strategies!

- What about optimal strategies?

for every $\epsilon > 0$ exists a strategy $\sigma_\epsilon$

\[
\inf_{\pi \in \Pi} P_{v, \pi}^{\sigma_\epsilon} (\text{Reach} \leq t(T)) \geq \text{val}(v) - \epsilon
\]

\[
\text{inf}_{\pi \in \Pi} P_{v, \pi}^{\sigma_\epsilon} (\text{Reach} \leq t(T)) \geq \text{val}(v) - \epsilon
\]

does there exist a fixed strategy $\sigma^*$ such that:

\[
\text{inf}_{\pi \in \Pi} P_{v, \pi}^{\sigma^*} (\text{Reach} \leq t(T)) = \text{val}(v)
\]
Optimal strategies in CTGs

Optimal strategies do not exist in general

\[ \text{val}(\text{start}) = 0 \]

but for all strategies \( \pi \):

\[ P \bar{\sigma}, \pi \text{start} (\text{Reach} \leq 1 (\{\text{goal}\})) > 0 \]

Theorem

In every finitely-branching CTG \( G \), the player ♦ has an optimal strategy.

Theorem

In every finitely-branching CTG \( G \) with bounded rates, the player □ has an optimal strategy.
Optimal strategies in CTGs

Optimal strategies do not exist in general

\[ \text{val}(\text{start}) = 0 \]
Optimal strategies in CTGs

Optimal strategies do not exist in general

\[ val(start) = 0 \text{ but for all strategies } \pi : P_{start}(\text{Reach}^{\leq 1}(\{\text{goal}\})) > 0 \]
Optimal strategies in CTGs

Optimal strategies do not exist in general

\[ \text{val}(\text{start}) = 0 \quad \text{but for all strategies } \pi : \mathcal{P}_{\text{start}}^{\bar{\sigma},\pi}(\text{Reach}^{\leq 1}(\{\text{goal}\})) > 0 \]

Theorem

In every finitely-branching CTG \( G \), the player ♦ has an optimal strategy.

Theorem

In every finitely-branching CTG \( G \) with bounded rates, the player □ has an optimal strategy.
Finite description of an optimal strategy in finite uniform CTGs

- Strategies – functions with infinite domain (i.e. infinite set of histories)
Finite description of an optimal strategy in finite uniform CTGs

- Strategies – functions with infinite domain (i.e. infinite set of histories)
- Is there a finite description of some optimal strategy?
Finite description of an optimal strategy in finite uniform CTGs

- Strategies – functions with infinite domain (i.e. infinite set of histories)
- Is there a finite description of some optimal strategy?

Greedy strategies

A *greedy strategy* – reach the target in as few steps as possible.
Finite description of an optimal strategy in finite uniform CTGs

- Strategies – functions with infinite domain (i.e. infinite set of histories)
- Is there a finite description of some optimal strategy?

Greedy strategies
A greedy strategy – reach the target in as few steps as possible.
Finite description of an optimal strategy in finite uniform CTGs

- Strategies – functions with infinite domain (i.e. infinite set of histories)
- Is there a finite description of some optimal strategy?

Greedy strategies
A greedy strategy – reach the target in as few steps as possible.

Optimal strategy with finite description
- Up to the $k$-th step behaves like an optimal history-dependent strategy
- After the $k$-th step behaves greedily
Algorithmical results

**Theorem**

*For a finite CTG $G$ and $\epsilon > 0$, there are $\epsilon$-optimal strategies for both players computable in time*

$$|V| \cdot |A| \cdot b p^2 \cdot \frac{1}{\epsilon} O(1) \cdot \left( \left( \max R \right) \cdot t + \ln \frac{1}{\epsilon} \right)^O(|R|).$$

**Theorem**

*For a finite uniform CTG $G$, the finite descriptions of optimal strategies for both players are effectively computable.*
Thank you for your attention!