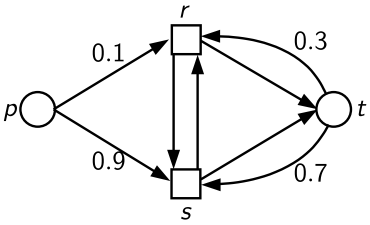
Controller Synthesis for MDPs with Qualitative Branching Time Objectives

Vojtěch Forejt

FI MU, Brno

Joint work with T. Brázdil and A. Kučera

Markov Decision Process (MDP)



- MDP is a tuple $\mathcal{M} = (V, E, (V_\square, V_\bigcirc), Prob)$
  - V is a finite set of vertices
  - $E \subseteq V \times V$ is a set of transitions
  - $(V_\square, V_\bigcirc)$ is a partition of V
  - $Prob : V_\bigcirc \times V \to [0, 1]$ is a probability assignment such that for every $s \in V_\bigcirc$: $\sum_{(s,t)\in E} Prob(s, t) = 1$

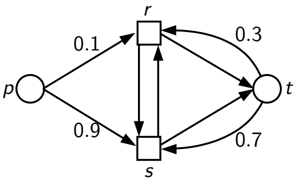
Strategies

Strategy σ is a function that to each $v_1 \dots v_n \in V^* V_\Box$ assigns a successor of v_n.

Strategies

Strategy σ is a function that to each $v_1 \dots v_n \in V^* V_\Box$ assigns a successor of v_n.

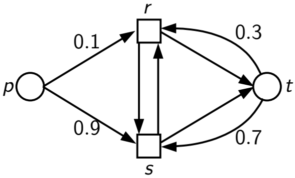
MDP:



Strategies

Strategy σ is a function that to each $v_1 \dots v_n \in V^* V_\Box$ assigns a successor of v_n.

MDP:



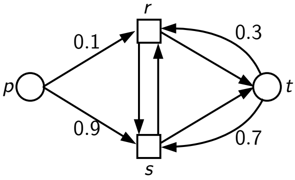
Strategy σ:

- $\sigma(\alpha r) = s$

Strategies

Strategy σ is a function that to each $v_1 \dots v_n \in V^* V_\Box$ assigns a successor of v_n.

MDP:



Strategy σ:

- $\sigma(\alpha r) = s$
- $\sigma(\alpha s) = r$

  for $\#_r(\alpha)$ odd;

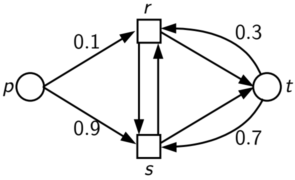
  $\sigma(\alpha s) = t$

  for $\#_r(\alpha)$ even;

Strategies

Strategy σ is a function that to each $v_1 \ldots v_n \in V^* V_\Box$ assigns a successor of v_n.

MDP:



Strategy σ:

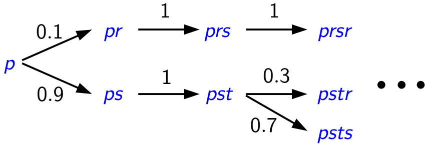
- $\sigma(\alpha r) = s$
- $\sigma(\alpha s) = r$

  for $\#_r(\alpha)$ odd;

  $\sigma(\alpha s) = t$

  for $\#_r(\alpha)$ even;

Play:



The Logic qPCTL

Syntax (simplified)

$$\varphi ::= a \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \mathcal{P}^{\bowtie\varrho}\mathbf{X}\varphi \mid \mathcal{P}^{\bowtie\varrho}\mathbf{F}\varphi \mid \mathcal{P}^{\bowtie\varrho}\mathbf{G}\varphi$$

- $\bowtie\varrho \in \{>0, =0, <1, =1\}$

The Logic qPCTL

Syntax (simplified)

$$\varphi ::= a \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \mathcal{P}^{\bowtie\varrho}\mathbf{X}\varphi \mid \mathcal{P}^{\bowtie\varrho}\mathbf{F}\varphi \mid \mathcal{P}^{\bowtie\varrho}\mathbf{G}\varphi$$

- $\bowtie\varrho \in \{>0, =0, <1, =1\}$
- For example $\mathcal{P}^{>0}\mathbf{G}(a \wedge \mathcal{P}^{=0}\mathbf{X}b) \wedge \mathcal{P}^{<1}\mathbf{F}a$

The Logic qPCTL

Syntax (simplified)

$$\varphi ::= a \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \mathcal{P}^{\bowtie\varrho}\mathbf{X}\varphi \mid \mathcal{P}^{\bowtie\varrho}\mathbf{F}\varphi \mid \mathcal{P}^{\bowtie\varrho}\mathbf{G}\varphi$$

- $\bowtie\varrho \in \{>0, =0, <1, =1\}$
- For example $\mathcal{P}^{>0}\mathbf{G}(a \wedge \mathcal{P}^{=0}\mathbf{X}b) \wedge \mathcal{P}^{<1}\mathbf{F}a$

Semantics For a vertex v,

- $v \models \mathcal{P}^{=1}\mathbf{F}\varphi$ iff
  - a vertex satisfying φ is reached with probability $=1$

The Logic qPCTL

Syntax (simplified)

$$\varphi ::= a \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \mathcal{P}^{\bowtie\varrho}\mathbf{X}\varphi \mid \mathcal{P}^{\bowtie\varrho}\mathbf{F}\varphi \mid \mathcal{P}^{\bowtie\varrho}\mathbf{G}\varphi$$

- $\bowtie\varrho \in \{>0, =0, <1, =1\}$
- For example $\mathcal{P}^{>0}\mathbf{G}(a \wedge \mathcal{P}^{=0}\mathbf{X}b) \wedge \mathcal{P}^{<1}\mathbf{F}a$

Semantics For a vertex v,

- $v \models \mathcal{P}^{=1}\mathbf{F}\varphi$ iff
  - a vertex satisfying φ is reached with probability $=1$
- $v \models \mathcal{P}^{>0}\mathbf{G}\varphi$ iff
  - φ will hold forever with probability >0

The Logic qPCTL

Syntax (simplified)

$$\varphi ::= a \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \mathcal{P}^{\bowtie\varrho}\mathbf{X}\varphi \mid \mathcal{P}^{\bowtie\varrho}\mathbf{F}\varphi \mid \mathcal{P}^{\bowtie\varrho}\mathbf{G}\varphi$$

- $\bowtie\varrho \in \{>0, =0, <1, =1\}$
- For example $\mathcal{P}^{>0}\mathbf{G}(a \wedge \mathcal{P}^{=0}\mathbf{X}b) \wedge \mathcal{P}^{<1}\mathbf{F}a$

Semantics For a vertex v,

- $v \models \mathcal{P}^{=1}\mathbf{F}\varphi$ iff
  - a vertex satisfying φ is reached with probability $=1$
- $v \models \mathcal{P}^{>0}\mathbf{G}\varphi$ iff
  - φ will hold forever with probability >0
- $v \models \mathcal{P}^{=1}\mathbf{X}\varphi$ iff
  - with probability $=1$, φ will hold in a next vertex

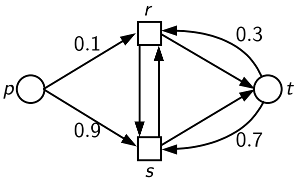
The Controller-Synthesis Problem

Given a MDP $\mathcal{M}$,
its vertex v,
and a qPCTL formula φ,
is there a strategy σ that ensures satisfying φ in v ?

The Controller-Synthesis Problem

Example I

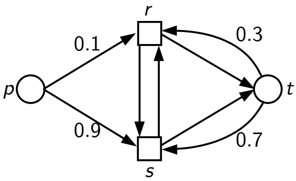
MDP:



The Controller-Synthesis Problem

Example I

MDP:

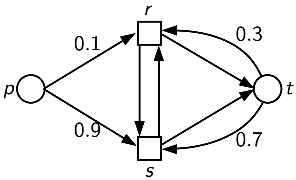


Vertex: r

The Controller-Synthesis Problem

Example I

MDP:



Vertex: r

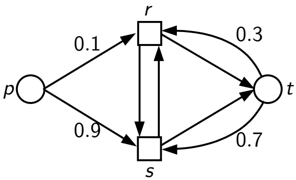
Formula φ:

$$(\mathcal{P}^{>0}\mathbf{X}s) \wedge (\mathcal{P}^{=1}\mathbf{X}t)$$

The Controller-Synthesis Problem

Example I

MDP:



The winning strategy does not exist.

Vertex: r

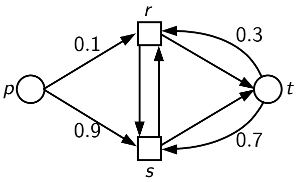
Formula φ:

$$(\mathcal{P}^{>0}\mathbf{X}s) \wedge (\mathcal{P}^{=1}\mathbf{X}t)$$

The Controller-Synthesis Problem

Example II

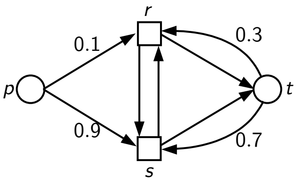
MDP:



The Controller-Synthesis Problem

Example II

MDP:

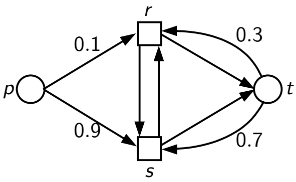


Vertex: p

The Controller-Synthesis Problem

Example II

MDP:



Vertex: p

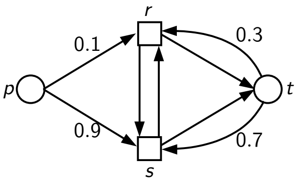
Formula φ:

$$\mathcal{P}^{>0}\mathbf{F}(r \wedge \mathcal{P}^{=1}\mathbf{G}\neg p)$$

The Controller-Synthesis Problem

Example II

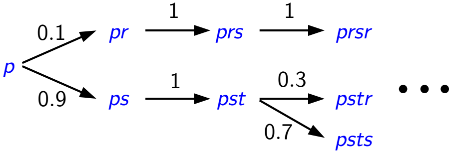
MDP:



Vertex: p

Formula φ:
$\mathcal{P}^{>0}\mathbf{F}(r \wedge \mathcal{P}^{=1}\mathbf{G}\neg p)$

- The winning strategy does exist!
  (Use the strategy from example)

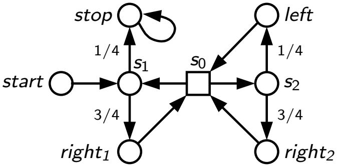


The Result

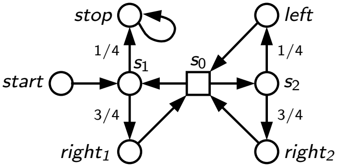
Theorem

The controller-synthesis problem for MDPs and qPCTL is EXPTIME-complete.

Simulating Random Walks

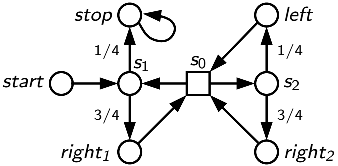


Simulating Random Walks



- Is there a strategy σ such that
  $\mathit{start} \models_\sigma \mathcal{P}^{>0}\mathbf{G}(\neg \mathit{stop} \wedge \mathcal{P}^{>0}\mathbf{F}\mathit{stop})$.

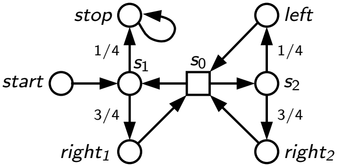
Simulating Random Walks



- Is there a strategy σ such that
  start $\models_\sigma \mathcal{P}^{>0}\mathbf{G}(\neg \mathit{stop} \wedge \mathcal{P}^{>0}\mathbf{F}\mathit{stop})$.

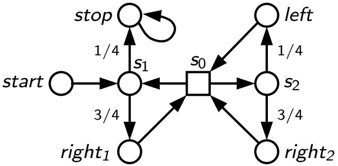
- Suppose there is a finite memory strategy σ. Then, each bottom SCC of $\mathcal{M}_\sigma$

Simulating Random Walks



- Is there a strategy σ such that *start* $\models_\sigma \mathcal{P}^{>0}\mathbf{G}(\neg stop \wedge \mathcal{P}^{>0}\mathbf{F} stop)$.
- Suppose there is a finite memory strategy σ. Then, each bottom SCC of $\mathcal{M}_\sigma$
  - either contains *stop* and violates $\mathbf{G}\neg stop$, or

Simulating Random Walks



- Is there a strategy σ such that
  start $\models_\sigma \mathcal{P}^{>0}\mathbf{G}(\neg \mathit{stop} \wedge \mathcal{P}^{>0}\mathbf{F}\mathit{stop})$.

- Suppose there is a finite memory strategy σ. Then, each bottom SCC of $\mathcal{M}_\sigma$
  - either contains *stop* and violates $\mathbf{G}\neg\mathit{stop}$, or
  - does not contain *stop* and *stop* is not reachable from this SCC.

Simulating Random Walks

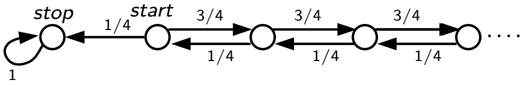
- In the following we have

$$\textit{start} \models_\sigma \mathcal{P}^{>0}\mathbf{G}(\neg \textit{stop} \wedge \mathcal{P}^{>0}\mathbf{F}\textit{stop})$$

Simulating Random Walks

- In the following we have

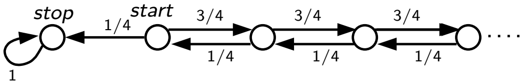
$$start \models_\sigma \mathcal{P}^{>0}\mathbf{G}(\neg stop \wedge \mathcal{P}^{>0}\mathbf{F}stop)$$



Simulating Random Walks

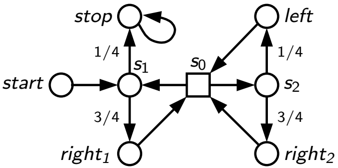
- In the following we have

$$\mathit{start} \models_\sigma \mathcal{P}^{>0}\mathbf{G}(\neg \mathit{stop} \land \mathcal{P}^{>0}\mathbf{F}\mathit{stop})$$

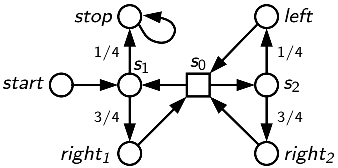


- Here, *stop* is reached from *start* only with probability $\frac{1}{3}$, although *stop* will always remain reachable.

Simulating Random Walks

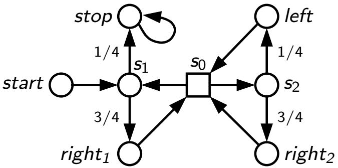


Simulating Random Walks



- Is there a strategy σ such that
  $start \models_\sigma \mathcal{P}^{>0}\mathbf{G}(\neg stop \wedge \mathcal{P}^{>0}\mathbf{F}stop)$

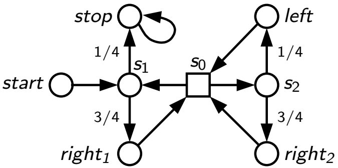
Simulating Random Walks



- Is there a strategy σ such that
  $start \models_\sigma \mathcal{P}^{>0}\mathbf{G}(\neg stop \wedge \mathcal{P}^{>0}\mathbf{F}stop)$

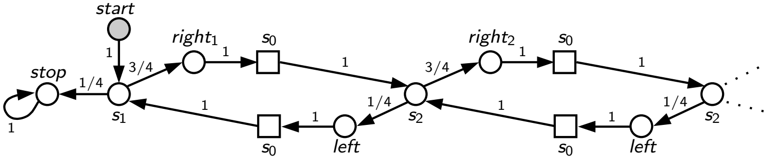
$$\sigma(ws_0) = \begin{cases} s_2 & \text{if } \#_{right}(w) > \#_{left}(w), \\ s_1 & \text{otherwise.} \end{cases}$$

Simulating Random Walks



- Is there a strategy σ such that $start \models_\sigma \mathcal{P}^{>0}\mathbf{G}(\neg stop \wedge \mathcal{P}^{>0}\mathbf{F} stop)$

$$\sigma(ws_0) = \begin{cases} s_2 & \text{if } \#_{right}(w) > \#_{left}(w), \\ s_1 & \text{otherwise.} \end{cases}$$



Sketch of the Solution

- Given a MDP $\mathcal{M}$, a vertex v and φ, we create a MDP $\mathcal{M}'$ and a set of simple linear time "subgoals".

Sketch of the Solution

- Given a MDP $\mathcal{M}$, a vertex v and φ, we create a MDP $\mathcal{M}'$ and a set of simple linear time "subgoals".
- There is a winning strategy for $\mathcal{M}$, v and φ if there is a strategy for subgoals in $\mathcal{M}'$.

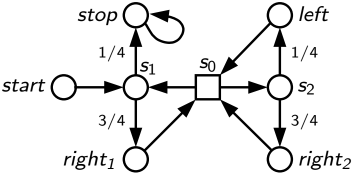
Sketch of the Solution

- Given a MDP $\mathcal{M}$, a vertex v and φ, we create a MDP $\mathcal{M}'$ and a set of simple linear time "subgoals".
- There is a winning strategy for $\mathcal{M}$, v and φ if there is a strategy for subgoals in $\mathcal{M}'$.
- Find finite memory strategies π, $\pi' \ldots$ for each vertex of $\mathcal{M}'$ and each subgoal.

Sketch of the Solution

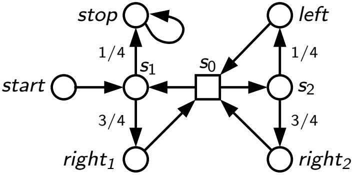
- Given a MDP $\mathcal{M}$, a vertex v and φ, we create a MDP $\mathcal{M}'$ and a set of simple linear time "subgoals".
- There is a winning strategy for $\mathcal{M}$, v and φ if there is a strategy for subgoals in $\mathcal{M}'$.
- Find finite memory strategies π, $\pi' \ldots$ for each vertex of $\mathcal{M}'$ and each subgoal.
- The winning strategy σ is a composition of these strategies. A counter simulating random walk is used to determine which π to use.

Solution – Example



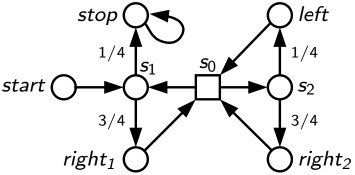
- For $\mathcal{P}^{>0}\mathbf{G}(\neg stop \wedge \mathcal{P}^{>0}\mathbf{F} stop)$ the subgoals are $\mathbf{G}\neg stop$ and $\mathbf{F} stop$.

Solution – Example



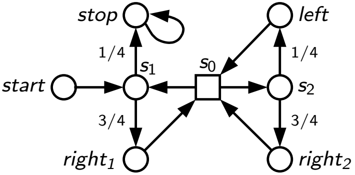
- For $\mathcal{P}^{>0}\mathbf{G}(\neg stop \wedge \mathcal{P}^{>0}\mathbf{F} stop)$ the subgoals are $\mathbf{G}\neg stop$ and $\mathbf{F} stop$.
- For $\mathbf{G}\neg stop$, a winning strategy π_1 is "go to s_2" and for $\mathbf{F} stop$ a winning strategy π_2 is "go to s_1".

Solution – Example

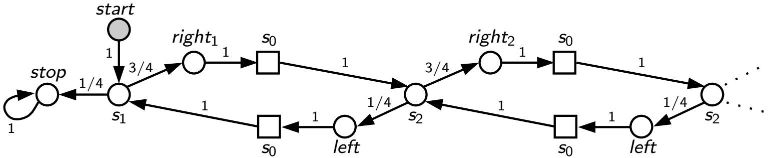


- For $\mathcal{P}^{>0}\mathbf{G}(\neg stop \wedge \mathcal{P}^{>0}\mathbf{F} stop)$ the subgoals are $\mathbf{G}\neg stop$ and $\mathbf{F} stop$.
- For $\mathbf{G}\neg stop$, a winning strategy π_1 is "go to s_2" and for $\mathbf{F} stop$ a winning strategy π_2 is "go to s_1".
- Reset counter to 0 at *start*. Increment it in *right*s and decrement in *left*. When the counter is > 0, use π_1, otherwise use π_2.

Solution – Example



- For $\mathcal{P}^{>0}\mathbf{G}(\neg stop \wedge \mathcal{P}^{>0}\mathbf{F} stop)$ the subgoals are $\mathbf{G}\neg stop$ and $\mathbf{F} stop$.
- For $\mathbf{G}\neg stop$, a winning strategy π_1 is "go to s_2" and for $\mathbf{F} stop$ a winning strategy π_2 is "go to s_1".
- Reset counter to 0 at *start*. Increment it in *right*s and decrement in *left*. When the counter is > 0, use π_1, otherwise use π_2.



Notes and Future Work

Notes

- The problem is decidable even for randomized strategies, qualitative PCTL* and qualitative PECTL*.

Notes and Future Work

Notes

- The problem is decidable even for randomized strategies, qualitative PCTL* and qualitative PECTL*.
- The problem is undecidable for quantitative PCTL [BBFK06]

Notes and Future Work

Notes

- The problem is decidable even for randomized strategies, qualitative PCTL* and qualitative PECTL*.
- The problem is undecidable for quantitative PCTL [BBFK06]

Future work

- Can we extend the result to games of 2 and $\frac{1}{2}$ players?

Thank you for your attention