On the strategy synthesis problem in MDPs: probabilistic CTL and rolling windows

Séminaire IRISA

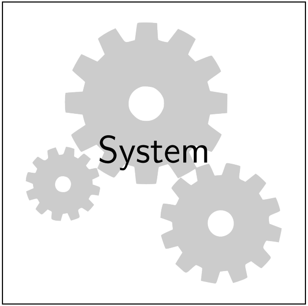
Damien Busatto-Gaston, Benjamin Bordais, Debraj Chakraborty
Shibashis Guha and Jean-François Raskin

Université Libre de Bruxelles

June 9, 2022

Introduction: Vérification

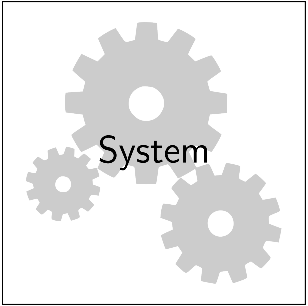
Ensuring system safety:



$\models$ Specification

Introduction: Vérification

Ensuring system safety:

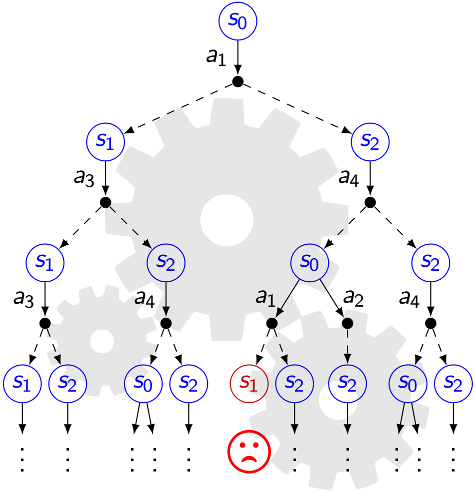


$\models$ Specification

Using *formal methods*:

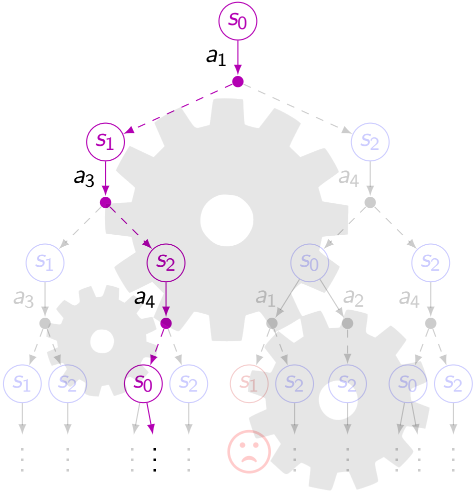
- model: finite automata, transition systems
- Specification: property in temporal logic

Test, model-checking, synthesis



$\models$ Specification

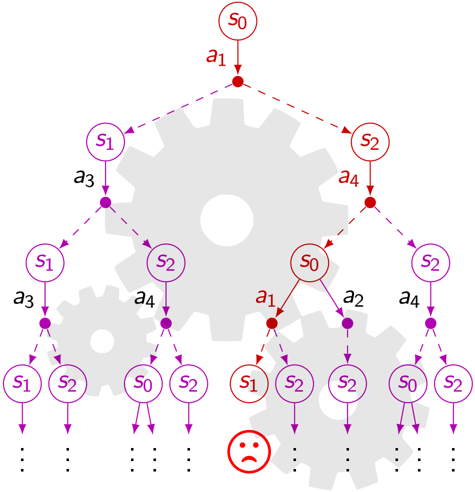
Test, model-checking, synthesis



$\models$ Specification

- Test

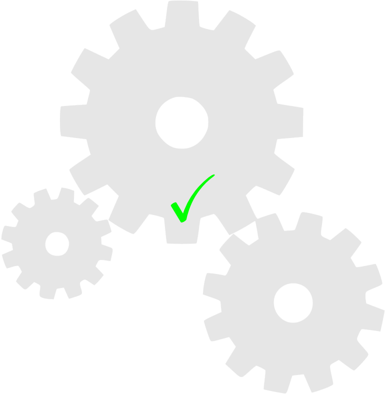
Test, model-checking, synthesis



$\models$ Specification

- Model-checking: $\boxed{\text{System}}$ $\models$ Specification

Test, model-checking, synthesis

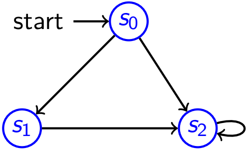


$\Leftarrow$ Specification

- Model-checking: $\boxed{\text{System}}$ $\models$ Specification
- Synthesis: Specification $\rightarrow$ $\boxed{\text{System}}$

Introduction (Transition Systems)

- Formal verification: prevent erroneous behaviour of systems
- model: transition system M



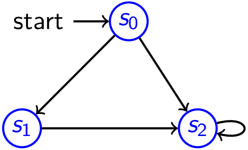
- specification: formula ϕ in some logic *e.g.* CTL

All paths reach s_2 and there is a path that avoids s_1

$$\mathsf{A}\,\mathsf{F}\, s_2 \wedge \mathsf{E}\,\mathsf{G}\,\neg s_1$$

Introduction (Transition Systems)

- Formal verification: prevent erroneous behaviour of systems
- model: transition system M



- specification: formula ϕ in some logic *e.g.* CTL

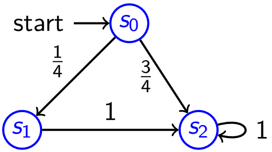
All paths reach s_2 and there is a path that avoids s_1

$$\mathsf{A}\,\mathsf{F}\,s_2 \wedge \mathsf{E}\,\mathsf{G}\,\neg s_1$$

- early 80s, *satisfiability* problem: given ϕ, does there exists M s.t. $M \models \phi$? $\rightsquigarrow$ EXPTIME-complete [Emerson and Halpern]
- Late 80s, *model-checking*: given M and ϕ, does $M \models \phi$? $\rightsquigarrow$ PTIME [Clarke and Emerson]

Introduction (Markov Chains)

- What about stochastic systems?
- model: Markov chain (MC) M



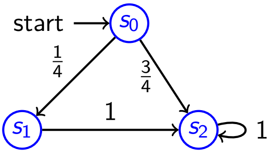
- specification: formula ϕ in some probabilistic logic *e.g.* PCTL

  Instead of A and E, compare probabilities to thresholds:

  $$\mathbb{P}\left[\mathsf{G}\,\neg s_1\right] \geq \frac{1}{2}$$

Introduction (Markov Chains)

- What about stochastic systems?
- model: Markov chain (MC) M



- specification: formula ϕ in some probabilistic logic *e.g.* PCTL

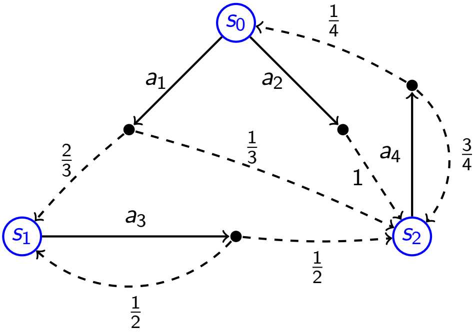
Instead of A and E, compare probabilities to thresholds:

$$\mathbb{P}\left[\mathsf{G}\,\neg s_1\right] \geq \frac{1}{2}$$

- *model-checking*: given M and ϕ, does $M \models \phi$? $\rightsquigarrow$ PTIME
- *satisfiability* problem: given ϕ, does there exists M s.t. $M \models \phi$? $\rightsquigarrow$ open problem

Introduction (Markov Decision Process)

- Reactive synthesis?
- model: Markov Decision Process (MDP) $\mathcal{M}$



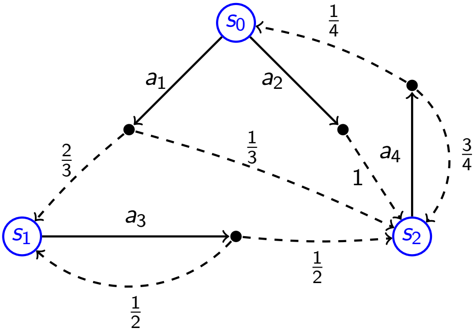
- specification: formula ϕ in some probabilistic logic *e.g.* PCTL

Instead of A and E, compare probabilities to thresholds:

$$\mathbb{P}\left[\mathsf{G}\,\neg s_1\right] \geq \frac{1}{2}$$

Introduction (Markov Decision Process)

- Reactive synthesis?
- model: Markov Decision Process (MDP) $\mathcal{M}$



- specification: formula ϕ in some probabilistic logic *e.g.* PCTL

  Instead of A and E, compare probabilities to thresholds:

  $$\mathbb{P}[\mathsf{G}\,\neg s_1] \geq \frac{1}{2}$$

- *strategy synthesis*: given $\mathcal{M}$ and ϕ, does there exists a strategy σ s.t. $\mathcal{M} \models_\sigma \phi$? $\rightsquigarrow$ Σ_1^1-hard (highly undecidable)

Introduction

- MDP | Controller strategy | Stochastic environment $\rightsquigarrow$ System (MC)
- Strategy synthesis for a given formula in a fragment of PCTL
  $\rightsquigarrow$ may be decidable
- What kind of strategy is allowed? $\rightsquigarrow$ memoryless/history-dependent, deterministic/randomised

Introduction

- MDP | Controller strategy | Stochastic environment $\rightsquigarrow$ System (MC)
- Strategy synthesis for a given formula in a fragment of PCTL
  $\rightsquigarrow$ may be decidable
- What kind of strategy is allowed? $\rightsquigarrow$ memoryless/history-dependent, deterministic/randomised

- Window formulae: express a "local" property
- Global window formulae: enforce globally a "local" property

Introduction

- MDP | Controller strategy | Stochastic environment $\rightsquigarrow$ System (MC)
- Strategy synthesis for a given formula in a fragment of PCTL $\rightsquigarrow$ may be decidable
- What kind of strategy is allowed? $\rightsquigarrow$ memoryless/history-dependent, deterministic/randomised

- Window formulae: express a "local" property
- Global window formulae: enforce globally a "local" property

- decision procedure for window properties
- global window remains undecidable
- synthesis of deterministic strategies becomes decidable!
- corolaries for PCTL satisfiability

Probabilistic CTL

Definition (PCTL syntax)

A formula of PCTL is generated by the nonterminal Φ in the following grammar:

$$\Phi := p \mid \neg\Phi \mid \Phi_1 \wedge \Phi_2 \mid \Phi_1 \vee \Phi_2 \mid \mathbb{P}\,[\varphi] \bowtie c$$
$$\varphi := \mathsf{X}\Phi \mid \Phi_1 \mathsf{U}^{\ell} \Phi_2 \mid \Phi_1 \mathsf{U} \Phi_2$$

where p ranges over the atomic propositions in AP, ℓ ranges over $\mathbb{N}$, $c \in \mathbb{Q}$ and $\bowtie \in \{\leq, <, \geq, >\}$.

- F, G, W can be defined from U
- U^{ℓ}: ϕ_2 reached within the first ℓ steps
- ℓ encoded in unary

Probabilistic CTL with linear inequalities

Definition (L-PCTL in normal form, syntax)

A formula of L-PCTL is generated by the nonterminal Φ in the following grammar:

$$\Phi := p \mid \neg p \mid \Phi_1 \wedge \Phi_2 \mid \Phi_1 \vee \Phi_2 \mid \sum_{i=1}^{n} c_i \mathbb{P}\left[\varphi_i\right] \succcurlyeq c_0$$

$$\varphi := \mathsf{X}^\ell \Phi \mid \Phi_1 \mathsf{U}^\ell \Phi_2 \mid \Phi_1 \mathsf{W}^\ell \Phi_2 \mid \Phi_1 \mathsf{U} \Phi_2 \mid \Phi_1 \mathsf{W} \Phi_2$$

where p ranges over the atomic propositions in AP, ℓ ranges over $\mathbb{N}$, and $n \in \mathbb{N}_{>0}$, $(c_0, \cdots, c_n) \in \mathbb{Z}^n$, $\succcurlyeq \in \{\geq, >\}$ define linear inequalities.

- F, G, F^ℓ and G^ℓ as syntactic sugar
- ℓ encoded in unary

Examples of PCTL formulae

- $\mathbb{P}\left[\mathsf{F}^2 a\right] \geq 0.5$
  The probability of reaching a within 2 steps is at least 0.5.

Examples of PCTL formulae

- $\mathbb{P}\left[\mathsf{F}^2 a\right] \geq 0.5$
  The probability of reaching a within 2 steps is at least 0.5.
- $\mathbb{P}\left[\mathsf{G}\,\mathbb{P}\left[a\,\mathsf{U}^5\,b\right] \geq 0.95\right] = 1$
  With probability 1, I enforce globally $\mathbb{P}\left[a\,\mathsf{U}^5\,b\right] \geq 0.95$.

Examples of PCTL formulae

- $\mathbb{P}\left[\mathsf{F}^2 a\right] \geq 0.5$
  The probability of reaching a within 2 steps is at least 0.5.
- $\mathbb{P}\left[\mathsf{G}\,\mathbb{P}\left[a\,\mathsf{U}^5\, b\right] \geq 0.95\right] = 1$
  With probability 1, I enforce globally $\mathbb{P}\left[a\,\mathsf{U}^5\, b\right] \geq 0.95$.
- $\mathbb{P}\left[\mathsf{F}^{10} a\right] \geq 2\mathbb{P}\left[\mathsf{X}\, b\right]$
  The probability of reaching a in ten steps is at least twice as big as the probability that the next state has label b.

Examples of PCTL formulae

- $\mathbb{P}\left[\mathsf{F}^2 a\right] \geq 0.5$
  The probability of reaching a within 2 steps is at least 0.5.
- $\mathbb{P}\left[\mathsf{G}\,\mathbb{P}\left[a\,\mathsf{U}^5\,b\right] \geq 0.95\right] = 1$
  With probability 1, I enforce globally $\mathbb{P}\left[a\,\mathsf{U}^5\,b\right] \geq 0.95$.
- $\mathbb{P}\left[\mathsf{F}^{10} a\right] \geq 2\mathbb{P}\left[\mathsf{X}\,b\right]$
  The probability of reaching a in ten steps is at least twice as big as the probability that the next state has label b.
- $\mathbb{P}\left[\mathsf{G}\,\mathbb{P}\left[\mathsf{F}\,a\right] = 1\right] = 1$
  The probability of visiting a infinitely often is 1.

Rolling windows

Definition

An L-PCTL formula Φ (in normal form) is a *window formula* if the horizon label ℓ of every path operator in Φ is finite, so that the unbounded U and W are not used.

Rolling windows

Definition

An L-PCTL formula Φ (in normal form) is a *window formula* if the horizon label ℓ of every path operator in Φ is finite, so that the unbounded U and W are not used.

Definition

A *global window* formula is a formula of the shape $\mathsf{A}\,\mathsf{G}\,\Phi$, with Φ a window L-PCTL formula. It is satisfied by a state s of M if every infinite path in $\mathrm{Paths}_M(s)$ satisfies the path formula $\mathsf{G}\,\Phi$, or equivalently if every state reachable from s satisfies Φ.

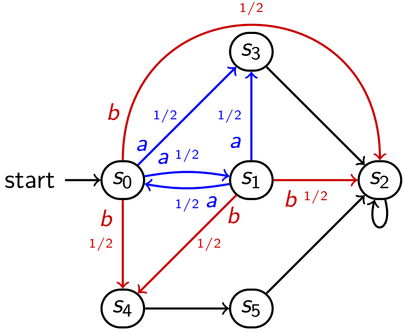
Lemma

The global window formula $\mathsf{A}\,\mathsf{G}\,\Phi$ *is satisfied on a state* s *of* M *if and only if* s *satisfies the L-PCTL formula* $\mathbb{P}\,[\mathsf{G}\,\Phi] = 1$.

Examples of PCTL formulae

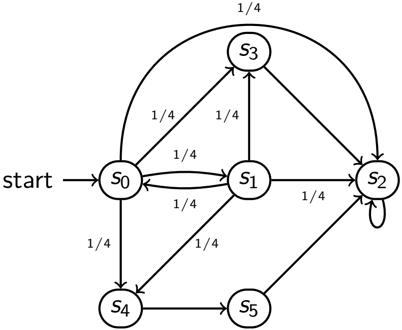
- $\mathbb{P}\left[\mathsf{F}^2 a\right] \geq 0.5$
  $\rightsquigarrow$ window formula
- $\mathbb{P}\left[\mathsf{G}\,\mathbb{P}\left[a\,\mathsf{U}^5\, b\right] \geq 0.95\right] = 1$
  $\rightsquigarrow$ global window formula
- $\mathbb{P}\left[\mathsf{F}^{10} a\right] \geq 2\mathbb{P}\left[\mathsf{X}\, b\right]$
  $\rightsquigarrow$ window formula
- $\mathbb{P}\left[\mathsf{G}\,\mathbb{P}\left[\mathsf{F}\, a\right] = 1\right] = 1$
  $\rightsquigarrow$ not in our fragments

Strategy synthesis



- window $\varphi = \mathbb{P}\left[\mathsf{F}^2 s_2\right] \geq \frac{9}{16}$
- global window $\mathsf{A}\,\mathsf{G}\,\varphi$
- $\rightsquigarrow$ play a and b with probability $\frac{1}{2}$

Strategy synthesis



- window $\varphi = \mathbb{P}\left[\mathsf{F}^2 s_2\right] \geq \frac{9}{16}$
- global window $\mathsf{A}\,\mathsf{G}\,\varphi$

- $\rightsquigarrow$ play a and b with probability $\frac{1}{2}$
- proba of reaching s_2 in 2 steps: $\frac{1}{4} + \frac{1}{16} + \frac{1}{4} = \frac{9}{16}$
- no other strategy satisfies $\mathsf{A}\,\mathsf{G}\,\varphi$

Complexity of strategy synthesis

Table: the synthesis problem for L-PCTL and MDPs

strategies:	Memoryless	History-dependent
Deterministic	NP-complete	Σ_1^1-complete
Randomized	in EXPTIME SQRT-SUM-hard	Σ_1^1-hard

[1] if the formula is flat and non-strict

Complexity of strategy synthesis

Table: the synthesis problem for L-PCTL and MDPs

strategies:	Memoryless	History-dependent
Deterministic	NP-complete	Σ_1^1-complete
Randomized	in EXPTIME SQRT-SUM-hard	Σ_1^1-hard

Theorem

Strategy synthesis for a window formula is decidable in EXPSPACE. For memoryless (resp. deterministic) startegies, PSPACE instead.

[1] if the formula is flat and non-strict

Complexity of strategy synthesis

Table: the synthesis problem for global window L-PCTL and MDPs

strategies:	Memoryless	History-dependent
Deterministic	NP-complete NP-complete	Σ_1^1-complete in 2EXPTIME EXPTIME-hard
Randomized	in EXPTIME SQRT-SUM-hard in PSPACE SQRT-SUM-hard	Σ_1^1-hard Σ_1^1-hard coRE-complete[1]

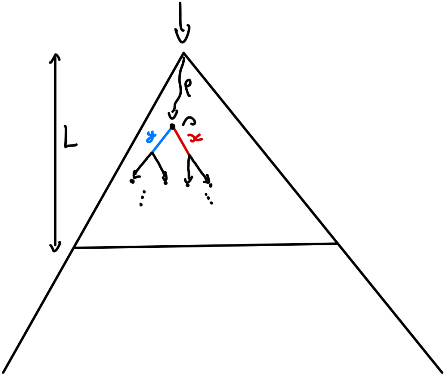
Theorem

Strategy synthesis for a window formula is decidable in EXPSPACE. For memoryless (resp. deterministic) startegies, PSPACE instead.

[1] if the formula is flat and non-strict

Window formulae

- unfold the MDP
- assign a variable $x, y, \ldots$ to every action in this tree
- truncate at an horizon L: the window length of the formula
- a strategy is now a point in $\mathbb{R}^{\{x,y,\ldots\}}$



Theory of the reals

Definition

The *first-order theory of the reals* (FO-$\mathbb{R}$) is the set of all well-formed sentences of first-order logic that involve universal and existential quantifiers and logical combinations of equalities and inequalities of real polynomials. $\exists$-$\mathbb{R}$ is the existential fragment.

Proposition

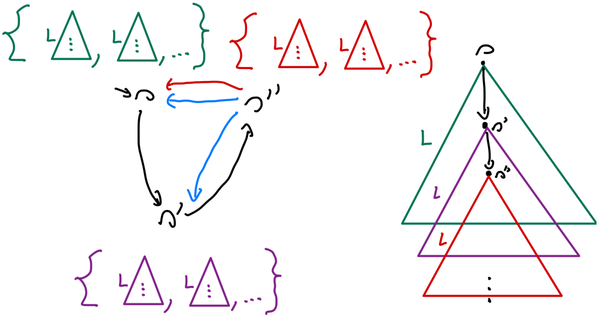
Let s be a state of $\mathcal{M}$ and Φ be a window L-PCTL formula. The set of strategies σ such that $s \models_\sigma \Phi$ can be represented in $\exists$-$\mathbb{R}$ as a formula of exponential size.

- intuition: use local consistency equations, that link the probability of a path formula on a given state to the probabilities of sub-formulae on the next states
- $\mathbb{P}_s[\mathrm{X}^2\, a] = \Sigma_{s \xrightarrow{a,p} s'} p \cdot \sigma(s, a) \cdot \mathbb{P}_{s'}[\mathrm{X}\, a]$
- a variable for every state and sub-formula

Global window formula $A\,G\,\Phi$

Fixed-point characterisation:

- For every state, we now have an $\exists$-$\mathbb{R}$ formula describing the set of (local) strategies that satisfy Φ
- does there exists a (global) strategy so that the sub-strategies obtained at each reachable state satisfy Φ?



Global window formula $\mathsf{A}\,\mathsf{G}\,\Phi$

Fixed-point characterisation:

- For every state, we now have an $\exists$-$\mathbb{R}$ formula describing the set of (local) strategies that satisfy Φ
- does there exists a (global) strategy so that the sub-strategies obtained at each reachable state satisfy Φ?
- intuition: On s, define an operator f that removes local strategies σ that are incompatible with the next state s' (no continuation of σ that satisfy Φ on s')
- take the greatest fixed point of f: a set Π_s of strategies for each s.

Proposition

Let s be a state, and let Φ be an L-PCTL window formula. Then, there exists a strategy σ so that $s \models_\sigma \mathsf{A}\,\mathsf{G}\,\Phi$ if and only if $\Pi_s \neq \emptyset$.

Decidability results

For deterministic (history-dependent) strategies:

- the strategies in the fixed point computation are described by their first L steps
- there are finitely many that are deterministic
- $\rightsquigarrow$ the fixed point is reached after finitely many steps
- pick strategies successively that stay in the fixed point
- $\rightsquigarrow$ finite memory strategy!

Theorem

The synthesis problem for global window L-PCTL formulae is in 2EXPTIME when restricted to deterministic strategies. Moreover, it is EXPTIME-hard.

Decidability results

For deterministic (history-dependent) strategies:

- the strategies in the fixed point computation are described by their first L steps
- there are finitely many that are deterministic
- $\rightsquigarrow$ the fixed point is reached after finitely many steps
- pick strategies successively that stay in the fixed point
- $\rightsquigarrow$ finite memory strategy!

Theorem

The synthesis problem for global window L-PCTL formulae is in 2EXPTIME when restricted to deterministic strategies. Moreover, it is EXPTIME-hard.

- lower bound: simulate an alternating TM of polynomial space
- this denotes from the general problem that is highly undecidable.

Decidability results

For memoryless (randomised) strategies:

- the previous $\exists$-$\mathbb{R}$ formulae can be greatly simplified by merging variables from the same state
- we can express the set of strategies of the fixed point directly, as a formula in $\exists$-$\mathbb{R}$

Theorem

The synthesis problem for global window L-PCTL formulae is in PSPACE when restricted to memoryless strategies.

Decidability results

For memoryless (randomised) strategies:

- the previous $\exists$-$\mathbb{R}$ formulae can be greatly simplified by merging variables from the same state
- we can express the set of strategies of the fixed point directly, as a formula in $\exists$-$\mathbb{R}$

Theorem

The synthesis problem for global window L-PCTL formulae is in PSPACE when restricted to memoryless strategies.

- lower bound: square root sum problem
- given $a_1 \dots a_n$ and b, does $\Sigma_i \sqrt{a_i} \leq b$?
- open problem of computational geometry, somewhere between NP and PSPACE

General case

If we do not restrict strategies, the fixed point characterisation does not yield decidability

- the fixed point may never be reached
- it cannot be expressed directly in the theory of the reals

- However, for a restricted class of formulae we can get semi-decidability:
- the set of points described by our formulae in the theory of the reals are compact sets
- if the fixed point is empty, it must be reached in finitely many steps $\rightsquigarrow$ co-recursively enumerable

Theorem

The synthesis problem for flat, non-strict global window L-PCTL formulae is coRE-complete.

General case

If we do not restrict strategies, the fixed point characterisation does not yield decidability

- the fixed point may never be reached
- it cannot be expressed directly in the theory of the reals

- However, for a restricted class of formulae we can get semi-decidability:
- the set of points described by our formulae in the theory of the reals are compact sets
- if the fixed point is empty, it must be reached in finitely many steps $\rightsquigarrow$ co-recursively enumerable

Theorem

The synthesis problem for flat, non-strict global window L-PCTL formulae is coRE-complete.

- lower bound: halting problem for two-counter Minsky machines
- this denotes from the general problem that is highly undecidable.

Undecidability

Theorem

The synthesis problem for global window L-PCTL formulae is Σ_1^1-hard.

- reduction from repeated reachability on two-counter Minsky machines
- we construct an MDP and a formula so that the only winning strategy accurately follows the execution of the Minsky machine
- the values of the counters are stored in the probabilities: $\frac{1}{2^{c_1} 3^{c_2}}$

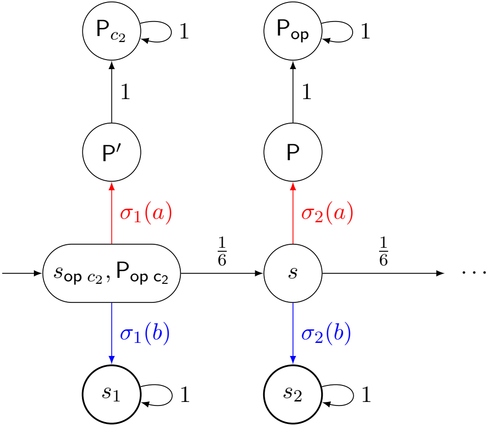
Undecidability

Theorem

The synthesis problem for global window L-PCTL formulae is Σ_1^1-hard.

- reduction from repeated reachability on two-counter Minsky machines
- we construct an MDP and a formula so that the only winning strategy accurately follows the execution of the Minsky machine
- the values of the counters are stored in the probabilities: $\frac{1}{2^{c_1} 3^{c_2}}$
- this denotes from the previous undecidability proof of PCTL synthesis that used deterministic strategies and used many unbounded F operators.

Need for linear inequalities in our reduction



The gadget Gd_{c_2++} for the operation $c_2 := c_2 + 1$.

$$\mathbb{P}(\mathsf{F}^2\mathsf{P}_{c_2}) = 6 \cdot 2 \cdot \mathbb{P}(\mathsf{F}^3\mathsf{P}_{\mathsf{op}})$$

PCTL satisfiability

Definition

PCTL *satisfiability* problem: given a formula ϕ, does there exists a Markov chain M s.t. $M \models \phi$?

- long standing open problem
- the Markov chain may need to be infinite or to have "weird" probabilities
- finite satisfiability also an open problem
- fixed granularity MC: restrict the probabilities to rationals $\frac{a}{N}$ of bounded denominators?

PCTL satisfiability

Definition

PCTL *satisfiability* problem: given a formula ϕ, does there exists a Markov chain M s.t. $M \models \phi$?

- long standing open problem
- the Markov chain may need to be infinite or to have "weird" probabilities
- finite satisfiability also an open problem
- fixed granularity MC: restrict the probabilities to rationals $\frac{a}{N}$ of bounded denominators?

Theorem

The satisfiability problem for global window L-PCTL formulae is decidable when restricted to Markov chains of fixed granularity. In this case, the formula has a model if and only if it has a finite model.

Conclusion

Contributions

- Notion of local property?
  - window PCTL formulae
- Rolling window constraint on executions?
  - global window PCTL
- How does it compare to full PCTL?
  - near the decidability threshold
  - a single unbounded until is enough for undecidability
  - deterministic strategy synthesis decidable
- Satisfiability?
  - yields a decidable fragment of PCTL for simple Markov chains

Conclusion

Contributions

- Notion of local property?
  - window PCTL formulae
- Rolling window constraint on executions?
  - global window PCTL
- How does it compare to full PCTL?
  - near the decidability threshold
  - a single unbounded until is enough for undecidability
  - deterministic strategy synthesis decidable
- Satisfiability?
  - yields a decidable fragment of PCTL for simple Markov chains

Related topics

- Window Mean-payoff
- Prompt LTL

Thank you