Threshold Constraints with Guarantees for Parity Objectives in Markov Decision Processes

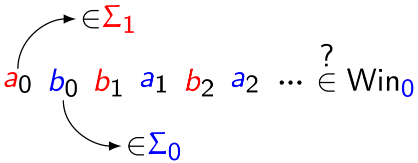
Raphaël Berthon [1,2]   Mickael Randour [2]   Jean-François Raskin [2]

[1]ENS Rennes, France

[2]Computer Science Department, ULB - Université libre de Bruxelles, Belgium

July 12, 2017

Context: Synthesis of reactive systems



Strategy synthesis problem

Given an effective representation of Win_0, decide if the system Player 0 has a winning strategy against the environment, Player 1 in this game, and synthesize this strategy.

Strategy: $\lambda_0 \colon (\Sigma_1.\Sigma_0)^*.\Sigma_1 \to \Sigma_0$

An easy game

Example

Let's take a game where two players choose letters in an alphabet $\Sigma = \{a, b\}$ and where $\text{Win}_0 = (ab \mid ba)^\omega$. The following play is winning:

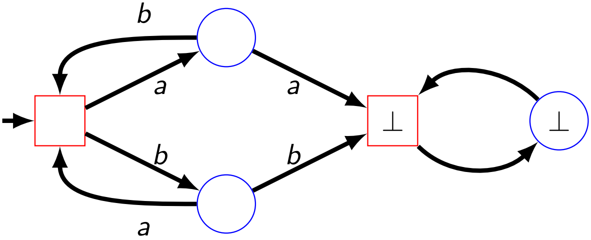
$ababbaabba\ldots$

How to represent a game?

$\exists\lambda_0.\ \forall\lambda_1.\ \texttt{Out}(\lambda_0,\lambda_1) \in \mathrm{Win}_0$

λ_0 must win against all strategies of Player 1

Previous example:

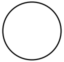


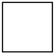
Theorem

Any game with a ω-regular winnning condition can be represented with a *Parity game*.

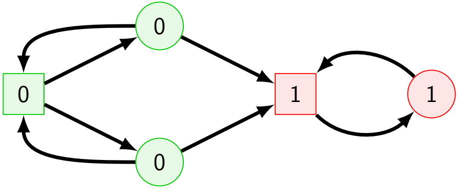
A parity game is:

- a game $\mathcal{M} = (G = (S, E), S_0, S_1)$
- a priority function $p \colon S \to \mathbb{N}$

 $= \mathcal{P}_0$

 $= \mathcal{P}_1$

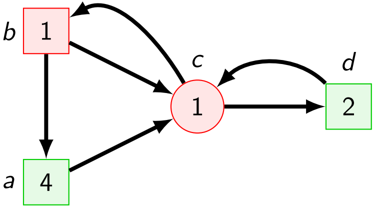
A run π is winning for $\mathcal{P}_0$ if the greatest priority seen infinitely often is even: $\max(\inf(\pi)) \in$ even.



Winning surely

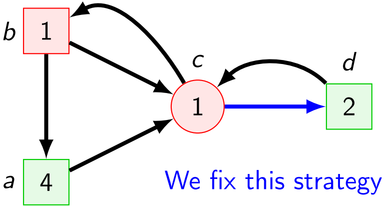
$\mathcal{P}_0$ is winning *surely* in a game if $\mathtt{Out}_{\lambda,s} \subseteq [\![p]\!]$: the system must win against all behaviors of the environment.

We denote the existence of such a strategy by $c \models_{\mathcal{M}} S(p)$.



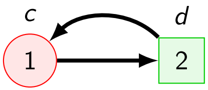
Winning surely

$\mathcal{P}_0$ is winning *surely* in a game if $\mathtt{Out}_{\lambda,s} \subseteq [\![p]\!]$: the system must win against all behaviors of the environment.
We denote the existence of such a strategy by $c \models_{\mathcal{M}} S(p)$.



Winning surely

$\mathcal{P}_0$ is winning *surely* in a game if $\mathtt{Out}_{\lambda,s} \subseteq [\![p]\!]$: the system must win against all behaviors of the environment.
We denote the existence of such a strategy by $c \models_{\mathcal{M}} S(p)$.



Winning with probability greater than a threshold

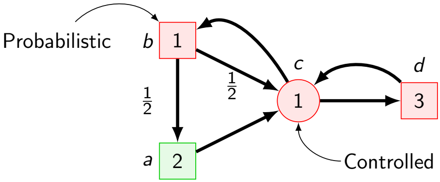
What if $\mathcal{P}_1$ is probabilistic?

$\rightarrow$ Stochastic model of the expected behavior of the environment: Markov Decision Process (MDP)

$\rightarrow$ Not any more surely winning strategy, but strategy winning with probability $\geqslant k$, $k \in \mathbb{Q} \cap [0, 1]$

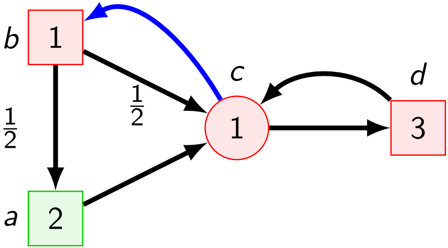
$\rightarrow$ Denoted $s \models_{\mathcal{M}} \mathrm{P}_{\geqslant k}(p)$

Here, $c \models_{\mathcal{M}} \mathrm{P}_{\geqslant 1}(p)$, denoted "almost surely": $c \models_{\mathcal{M}} AS(p)$



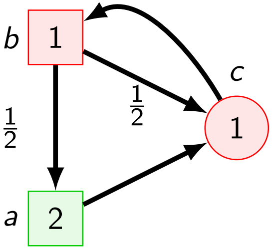
An example with threshold 1

$c \models_{\mathcal{M}} AS(p)$
We fix this strategy:



An example with threshold 1

$c \models_{\mathcal{M}} AS(p)$
In the following Markov chain, a is visited infinitely often with probablity 1.

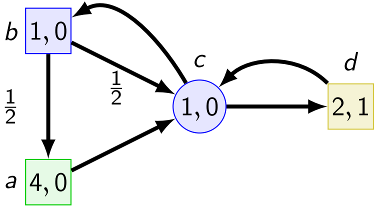


What if we have two parity objectives, p_1 and p_2, an initial state s, and want a strategy λ ensuring:

$$1 : \mathtt{Out}_{\lambda,s} \subseteq [\![p_1]\!]$$

$$2 : \mathbb{P}_{\lambda,s}(p_2) \geqslant 1$$

Here $c \models_{\mathcal{M}} S(p_1) \wedge AS(p_2)$

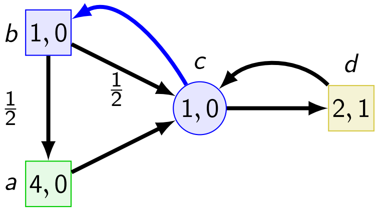


Infinite-memory winning strategy

$c \models_{\mathcal{M}} S(p_1) \wedge AS(p_2)$

In this example, there is a strategy ensuring both constraints. It needs infinite memory. We proceed by rounds made of an increasing number of steps.

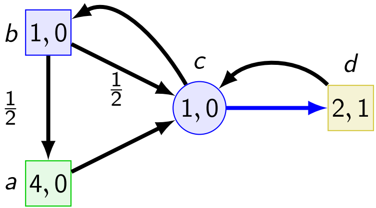
Round i



Play $c \to b$, i times

Infinite-memory winning strategy

$c \models_{\mathcal{M}} S(p_1) \wedge AS(p_2)$
In this example, there is a strategy ensuring both constraints. It needs infinite memory. We proceed by rounds made of an increasing number of steps.
Round i



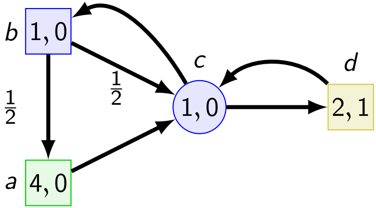
If a has not been reached this round, play $c \to d$ one time.

Strategy

$c \models_{\mathcal{M}} S(p_1) \wedge AS(p_2)$

Summary at round i:

- Play i times $c \to b$.
- If the current maximum of the round is odd, play one time $c \to d$.
- begin the next round $i+1$



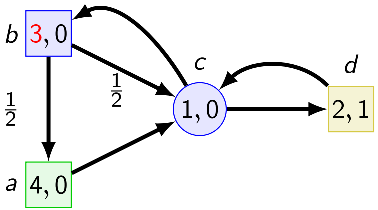
$c \to d$ is taken infinitely often with probability 0.

Winning each condition is not sufficient

Remark

Having a strategy λ_1 for $c \models_{\mathcal{M}} S(p_1)$ and a strategy λ_2 for $c \models_{\mathcal{M}} \mathrm{P}_{\geqslant 1}(p_2)$ is not sufficient to have a strategy for $c \models_{\mathcal{M}} S(p_1) \wedge \mathrm{P}_{\geqslant 1}(p_2)$.

The following example has no winning strategy for $S(p_1) \wedge AS(p_2)$.

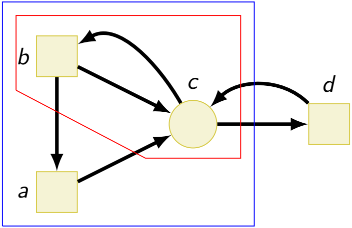


End-component

Definition

A subgraph C is an end-component if:

- C is strongly connected
- $\mathrm{Post}_{\square}(C) \subseteq C$



Theorem

For all strategy λ, $\mathbb{P}(\inf(\lambda) = \mathrm{EC}) = 1$.

Safe reachability

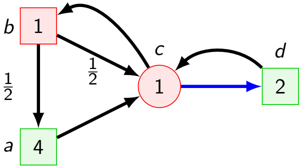
Definition

A set of states T can be reached safely from a state s with respect to a parity condition p if $s \models S(p) \wedge AS(\Diamond T)$.

This problem can be decided in NP$\cap$ co-NP [Almagor et al., 2016].

On this example, a can be reached safely from c with respect to p.

We alternate between the two possible actions:



Safe reachability

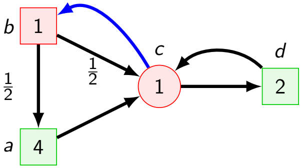
Definition

A set of states T can be reached safely from a state s with respect to a parity condition p if $s \models S(p) \wedge AS(\Diamond T)$.

This problem can be decided in NP$\cap$ co-NP [Almagor et al., 2016].

On this example, a can be reached safely from c with respect to p.

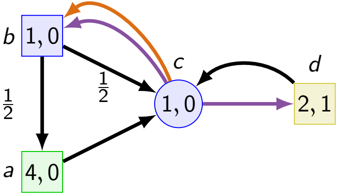
We alternate between the two possible actions:



Condition for sure and almost-sure

An end-component C is *ultra-good* (UGEC) if we have:

- from all state, a strategy λ_1 reaching safely the maximum of p_1 with respect to p_1
- from all state, a strategy λ_2 having probability 1 of satisfying both p_1 and p_2



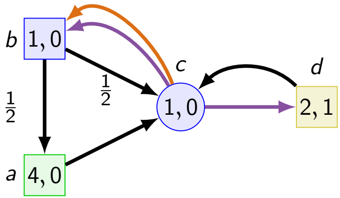
Lemma

The following holds: $\forall s \in UGEC : \; s \models_{\mathcal{M}} S(p_1) \wedge AS(p_2)$

Strategy for sure and almost-sure

Strategy at round i:

- Play i times λ_2.
- If the current maximum of the round is odd, play λ_1 until reaching the maximum of p_1.
- begin the next round $i+1$



Lemma

The following holds: $\forall s \in UGEC : \; s \models_{\mathcal{M}} S(p_1) \wedge AS(p_2)$

General result

Theorem

Given an MDP $\mathcal{M}$, a state $s_0 \in S$, and two priority functions p_1, p_2, it can be decided in NP ∩ coNP *if*

$$s_0 \models \mathbb{S}(p_1) \wedge \mathbb{P}_{\sim k}(p_2)$$

for $\sim \in \{>, \geqslant\}$ and $k \in \mathbb{Q} \cap [0, 1]$.
If the answer is Yes*, then there exists an infinite-memory witness strategy, and infinite memory is in general necessary. This decision problem is at least as hard as solving parity games.*

→ Proof of this result relies on the notion of UGEC and safe reachability.

Conclusion

- Synthesis for ω-regular objectives with sure condition and probabilistic one
- NP$\cap$co-NP algorithm. Uses parity games as a black box.
- Infinite memory is required, but the finite-memory cases can be easily obtained
- Can be extended to more than two objectives for sure and almost-sure cases
- No solution yet for more than one threshold case

Almagor, S., Kupferman, O., and Velner, Y. (2016).
Minimizing expected cost under hard boolean constraints, with applications to quantitative synthesis.
CoRR, abs/1604.07064.