Randomness for Free

Laurent Doyen

LSV, ENS Cachan & CNRS

joint work with Krishnendu Chatterjee,
Hugo Gimbert, Tom Henzinger

Stochastic games on graphs

- Games for synthesis
  - Reactive system synthesis = finding a winning strategy in a game
- Game played on a graph
  - Infinite number of rounds
  - Player's moves determine successor state
  - Outcome = infinite path in the graph

When is randomness more powerful ?

When is randomness for free ?

Stochastic games on graphs

- Games for synthesis
  - Reactive system synthesis = finding a winning strategy in a game
- Game played on a graph
  - Infinite number of rounds
  - Player's moves determine successor state
  - Outcome = infinite path in the graph

When is randomness more pow

When is randomness for free ?

… in game structures ?

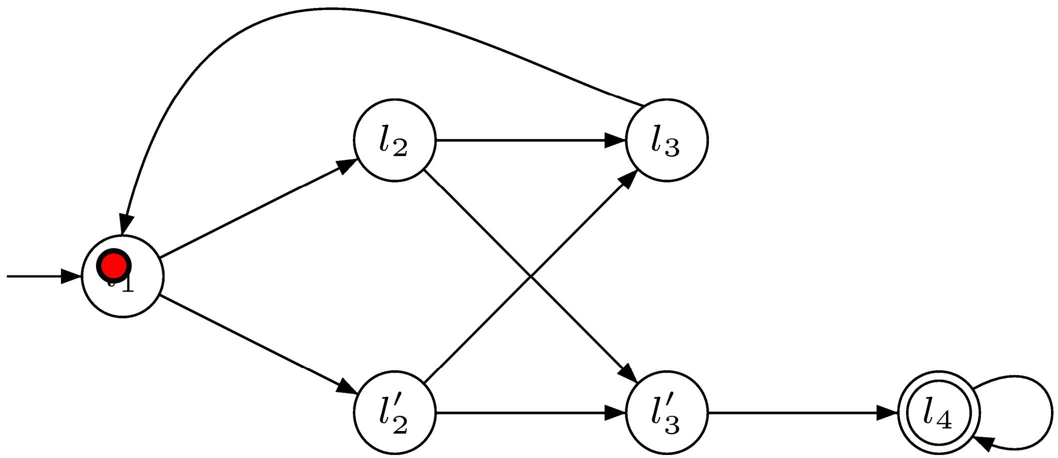
… in strategies ?

Stochastic games on graphs

- Classification according to Information & Interaction

Interaction: how players' moves are combined.

Information: what is visible to the players.



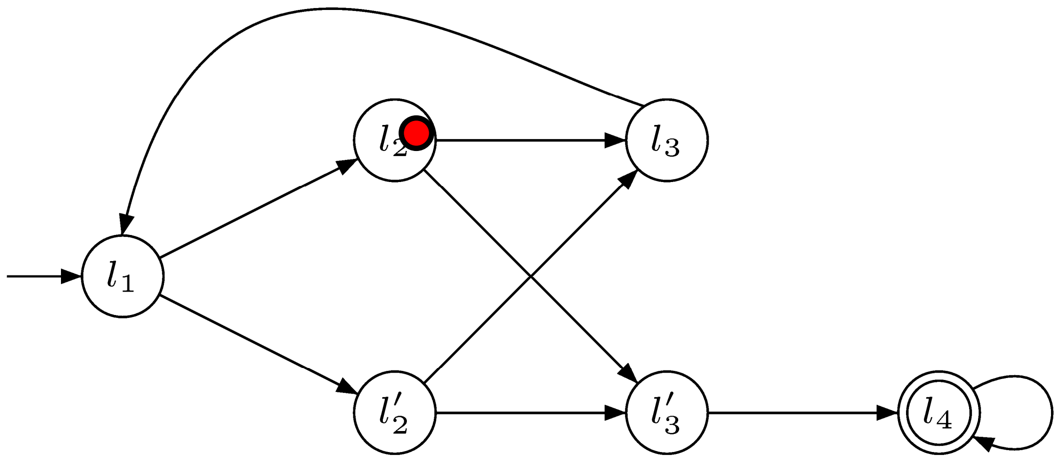
Round 1

Stochastic games on graphs

- Classification according to Information & Interaction

Interaction: how players' moves are combined.

Information: what is visible to the players.



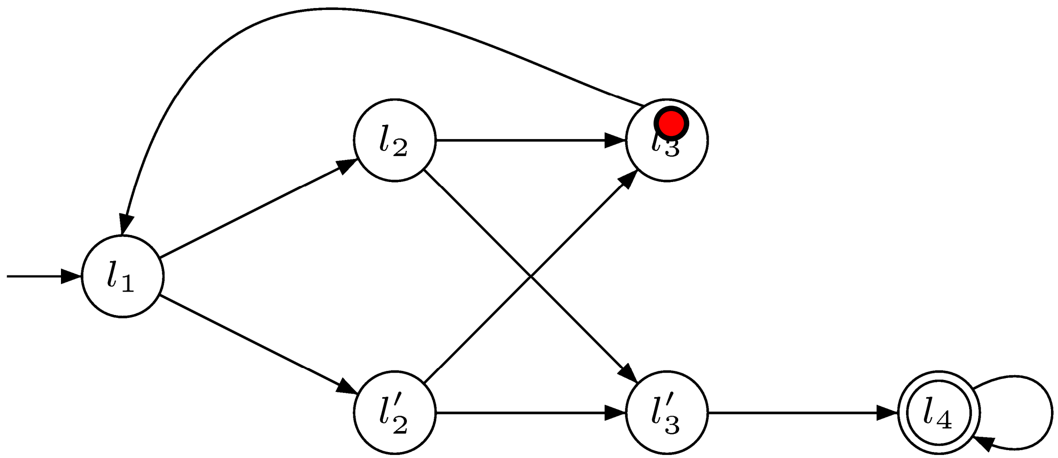
Round 2

Stochastic games on graphs

- Classification according to Information & Interaction

Interaction: how players' moves are combined.

Information: what is visible to the players.



Round 3

Mode of interaction

- Classification according to Information & Interaction

Interaction

General case: concurrent & stochastic

$$\delta : S \times A_1 \times A_2 \rightarrow \mathcal{D}(S)$$

Player 1's move

Player 2's move

Probability distribution on successor state

$2\,{}^1\!/_2$-player games

Players choose their moves simultaneously and independently

Mode of interaction

- Classification according to Information & Interaction

Interaction

General case: concurrent & stochastic

$$\delta : S \times A_1 \times A_2 \to \mathcal{D}(S)$$

Player 1's move

Player 2's move

Probability distribution on successor state

$1\,{}^1\!/_2$-player games (MDP) if A_1 or A_2 is a singleton.

Players choose their moves simultaneously and independently

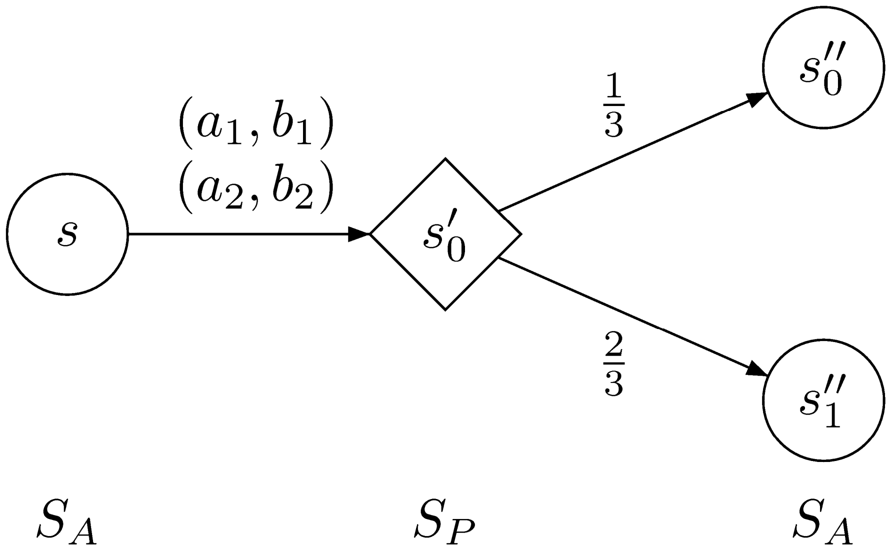
Mode of interaction

Interaction

General case: concurrent & stochastic

$$\delta : S \times A_1 \times A_2 \to \mathcal{D}(S) \quad \Longrightarrow \quad \delta : \begin{cases} S_A \times A_1 \times A_2 \to S_P \\ S_P \to \mathcal{D}(S_A) \end{cases}$$

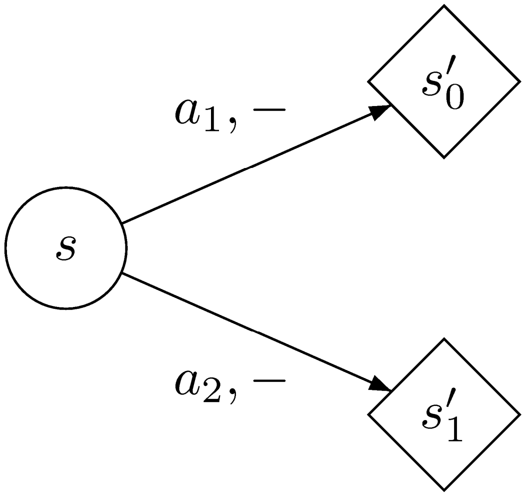
Separation of concurrency & probabilities



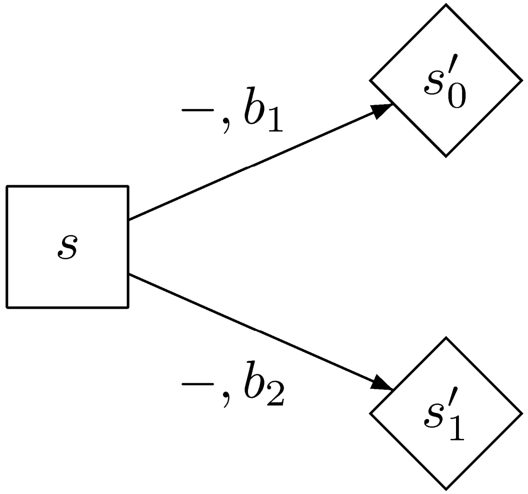
Mode of interaction

Interaction

Special case: turn-based



Player 1 state



Player 2 state

In each state, one player's move determines successor

Knowledge

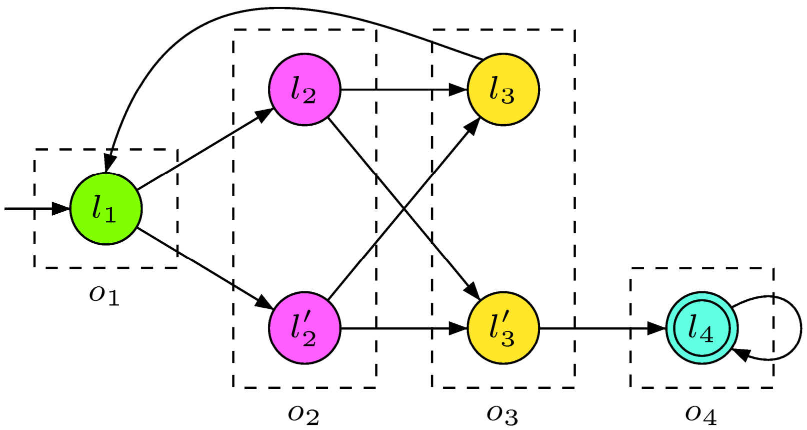
- Classification according to Information & Interaction

Information

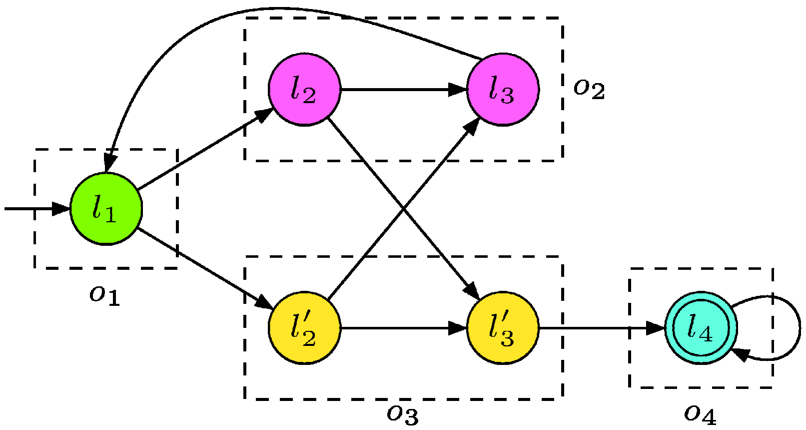
General case: partial observation

Two partitions $\mathcal{O}_1 \subseteq 2^S$ and $\mathcal{O}_2 \subseteq 2^S$

In state ℓ, player i sees $\mathsf{obs}_i(\ell)$ such that $\ell \in \mathsf{obs}_i(\ell)$



Player 1's view



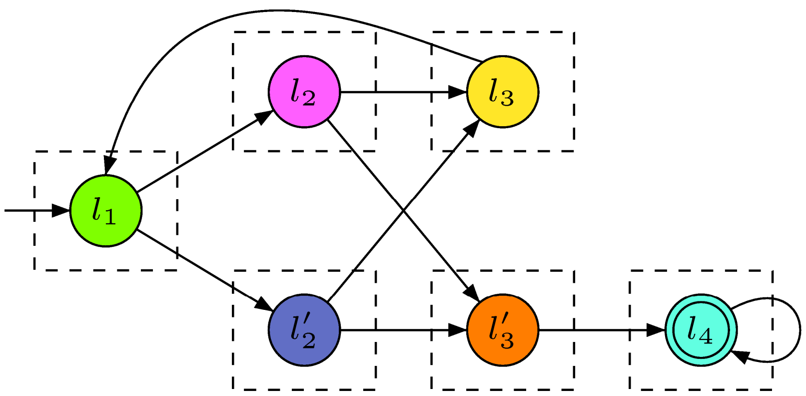
Player 2's view

Knowledge

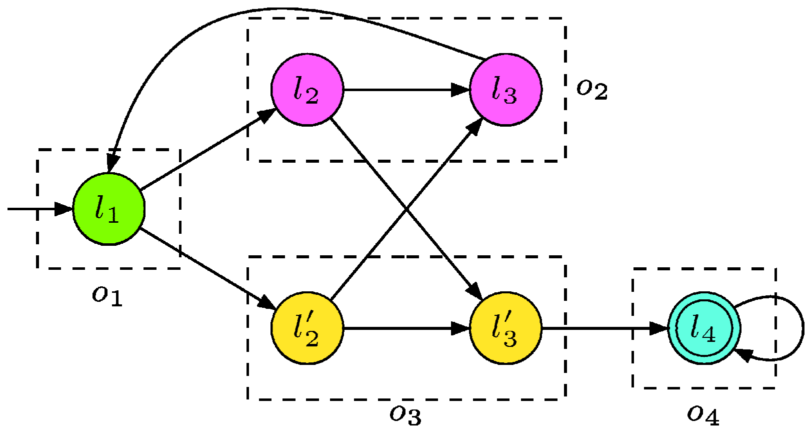
Information

Special case 1: one-sided complete observation

$$\mathcal{O}_1 = \{\{\ell\} \mid \ell \in S\} \quad \text{or} \quad \mathcal{O}_2 = \{\{\ell\} \mid \ell \in S\}$$



Player 1's view



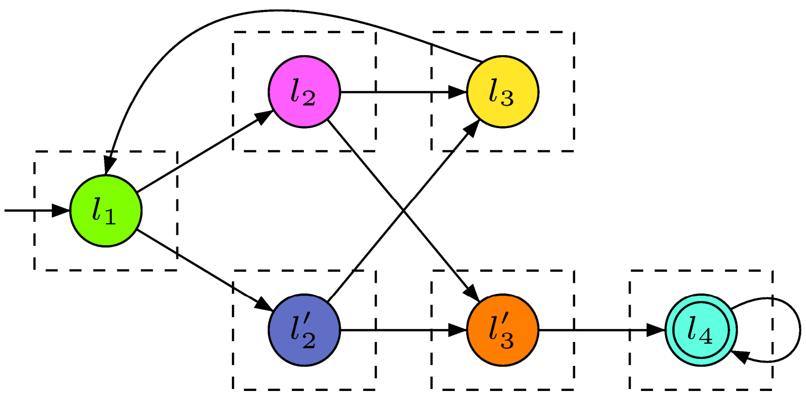
Player 2's view

Knowledge

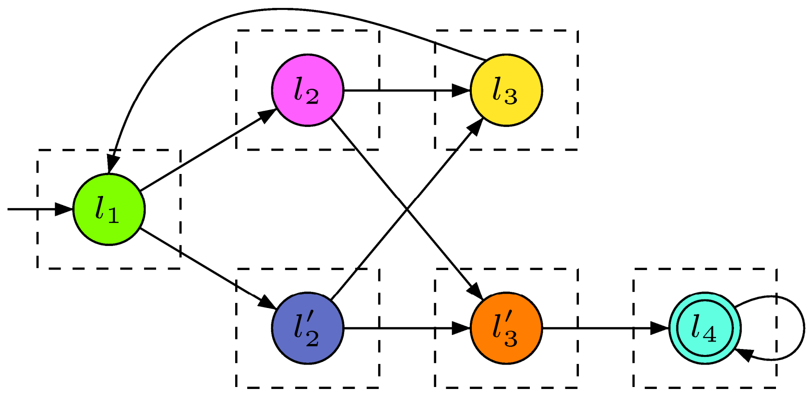
Information

Special case 2: complete observation

$\mathcal{O}_1 = \{\{\ell\} \mid \ell \in S\}$ and $\mathcal{O}_2 = \{\{\ell\} \mid \ell \in S\}$



Player 1's view



Player 2's view

Classification

- Classification according to Information & Interaction

Information	Interaction
partial observation	concurrent
one-sided complete obs.	turn-based
complete observation	

$2\,{}^{1}\!/\!_{2}$-player games

Classification

- Classification according to Information & Interaction

Information

partial observation

=

one-sided complete obs. POMDP

|

complete observation
MDP

Interaction

concurrent

=

turn-based

$1\,{}^1\!/_2$-player games

Markov Decision Process

Strategies

A strategy for Player i is a function $\sigma_i : S^+ \to \mathcal{D}(A_i)$ that maps histories to probability distribution over actions.

Strategies are observation-based:

$$\forall \rho, \rho' \in S^+ : \text{if } \mathsf{obs}_i(\rho) = \mathsf{obs}_i(\rho'), \texttt{ then } \sigma_i(\rho) = \sigma_i(\rho')$$

Strategies

A strategy for Player i is a function $\sigma_i : S^+ \to \mathcal{D}(A_i)$ that maps histories to probability distribution over actions.

Strategies are observation-based:

$$\forall \rho, \rho' \in S^+ : \text{if } \mathsf{obs}_i(\rho) = \mathsf{obs}_i(\rho'), \text{ then } \sigma_i(\rho) = \sigma_i(\rho')$$

Special case: pure strategies $\sigma_i : S^+ \to A_i$

Objectives

An objective is a measurable set of infinite sequences of states:

$$\varphi \subseteq S^\omega$$

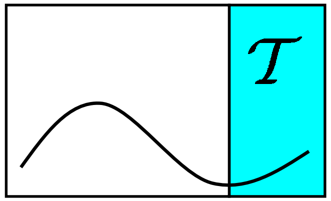
Objectives

An objective is a measurable set of infinite sequences of states:

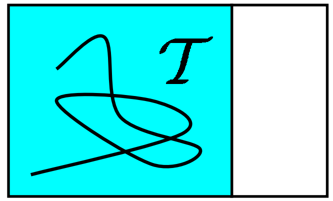
$$\varphi \subseteq S^{\omega}$$

Examples:

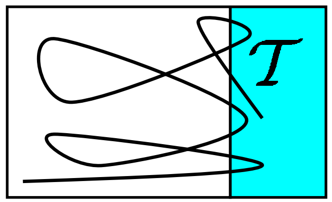
- Reachability, safety
- Büchi, coBüchi
- Parity
- Borel



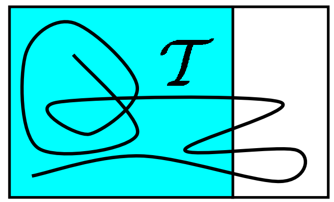
Reachability



Safety



Büchi



coBüchi

Value

Probability of finite prefix of a play:

$$P(s_0 \ldots s_n \mid s_0) = \prod_{i=1}^{n} p(s_i, s_{i+1})$$

$$p(s_i, s_{i+1}) = \sum_{a \in A_1, b \in A_2} \delta(s_i, \sigma_1(s_0 \ldots s_i)(a), \sigma_2(s_0 \ldots s_i)(b))(s_{i+1})$$

induces a unique probability measure on measurable sets of plays:

$$Pr_{s_0}^{\sigma_1,\sigma_2}(\cdot)$$

and a value function for Player 1:

$$\langle\langle 1 \rangle\rangle_{val}^{G}(\varphi)(s) = \sup_{\sigma_1 \in \Sigma_1} \inf_{\sigma_2 \in \Sigma_2} Pr_s^{\sigma_1,\sigma_2}(\varphi).$$

Value

Probability of finite prefix of a play:

$$P(s_0 \ldots s_n \mid s_0) = \prod_{i=1}^{n} p(s_i, s_{i+1})$$

$$p(s_i, s_{i+1}) = \sum_{a \in A_1, b \in A_2} \delta(s_i, \sigma_1(s_0 \ldots s_i)(a), \sigma_2(s_0 \ldots s_i)(b))(s_{i+1})$$

induces a unique probability measure on measurable set of plays:

Our reductions preserve values and existence of optimal strategies.

and a value f

$$\langle\langle 1 \rangle\rangle_{val}^{G}(\varphi)(s) = \sup_{\sigma_1 \in \Sigma_1} \inf_{\sigma_2 \in \Sigma_2} Pr_s^{\sigma_1, \sigma_2}(\varphi).$$

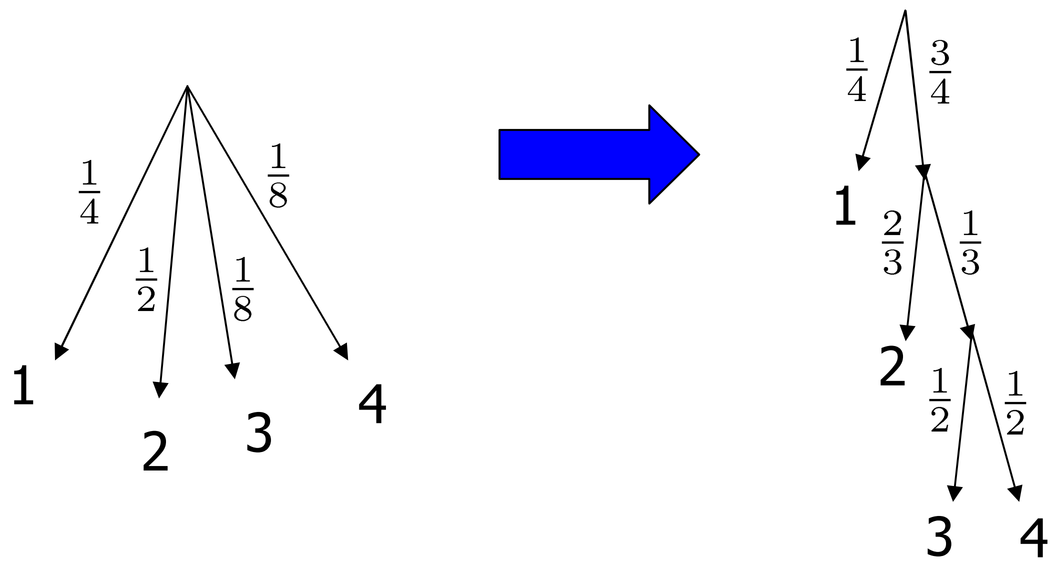
Outline

- Randomness in game structure
  - for free with complete-observation, concurrent
  - for free with one-sided, turn-based
- Randomness in strategies
  - for free in (PO)MDP
- Corollary: undecidability results

Preliminary

Rational probabilities → Probabilities $\frac{1}{2}$ only

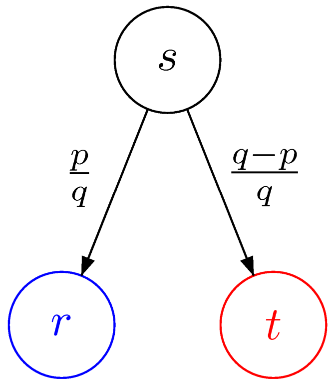
1. Make all states have two successors

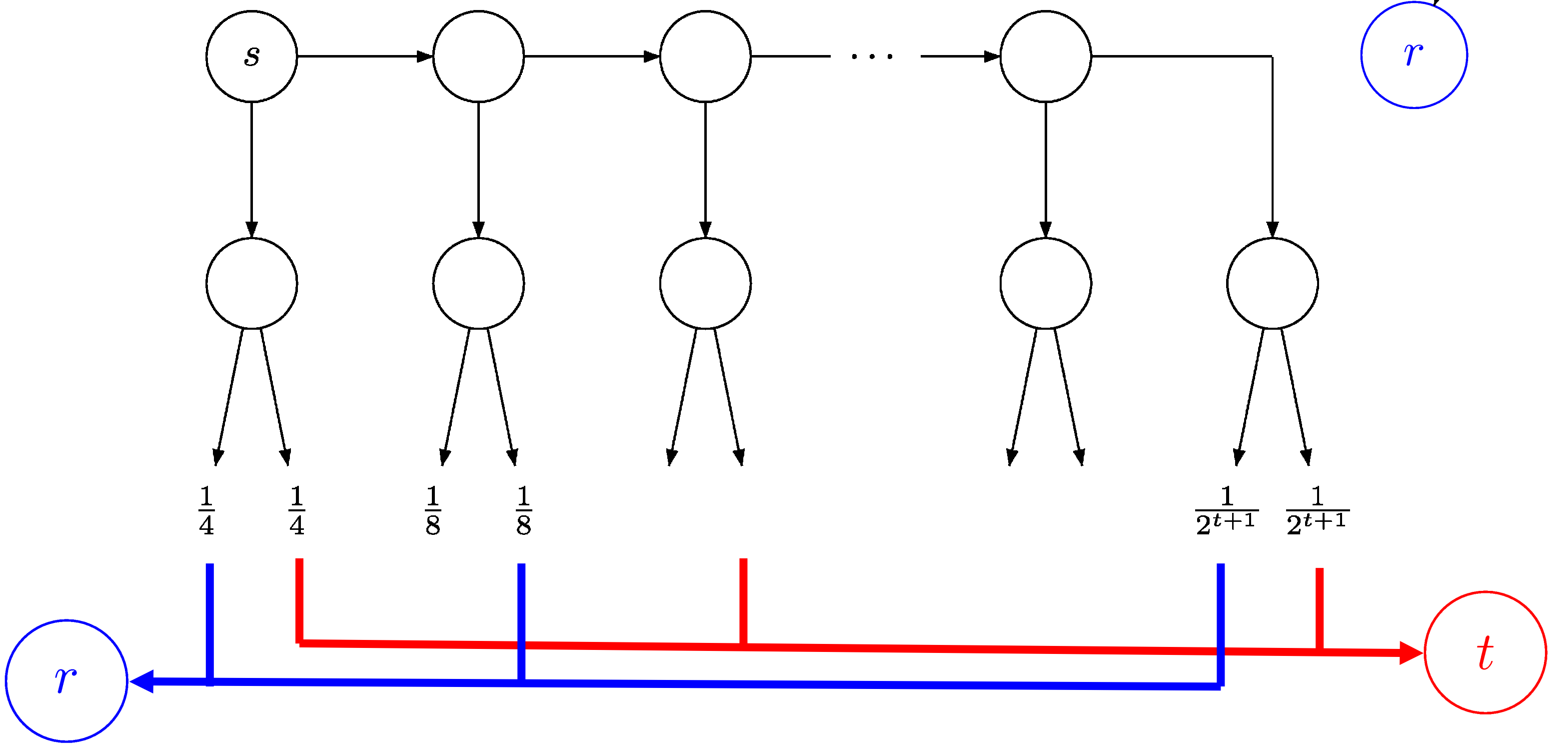


Preliminary

Rational probabilities ➡ Probabilities $\frac{1}{2}$ only

2. Use binary encoding of p and q-p to simulate [Zwick,Paterson 96] :

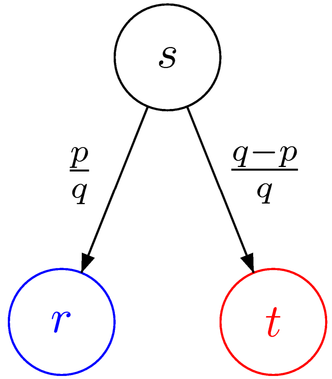


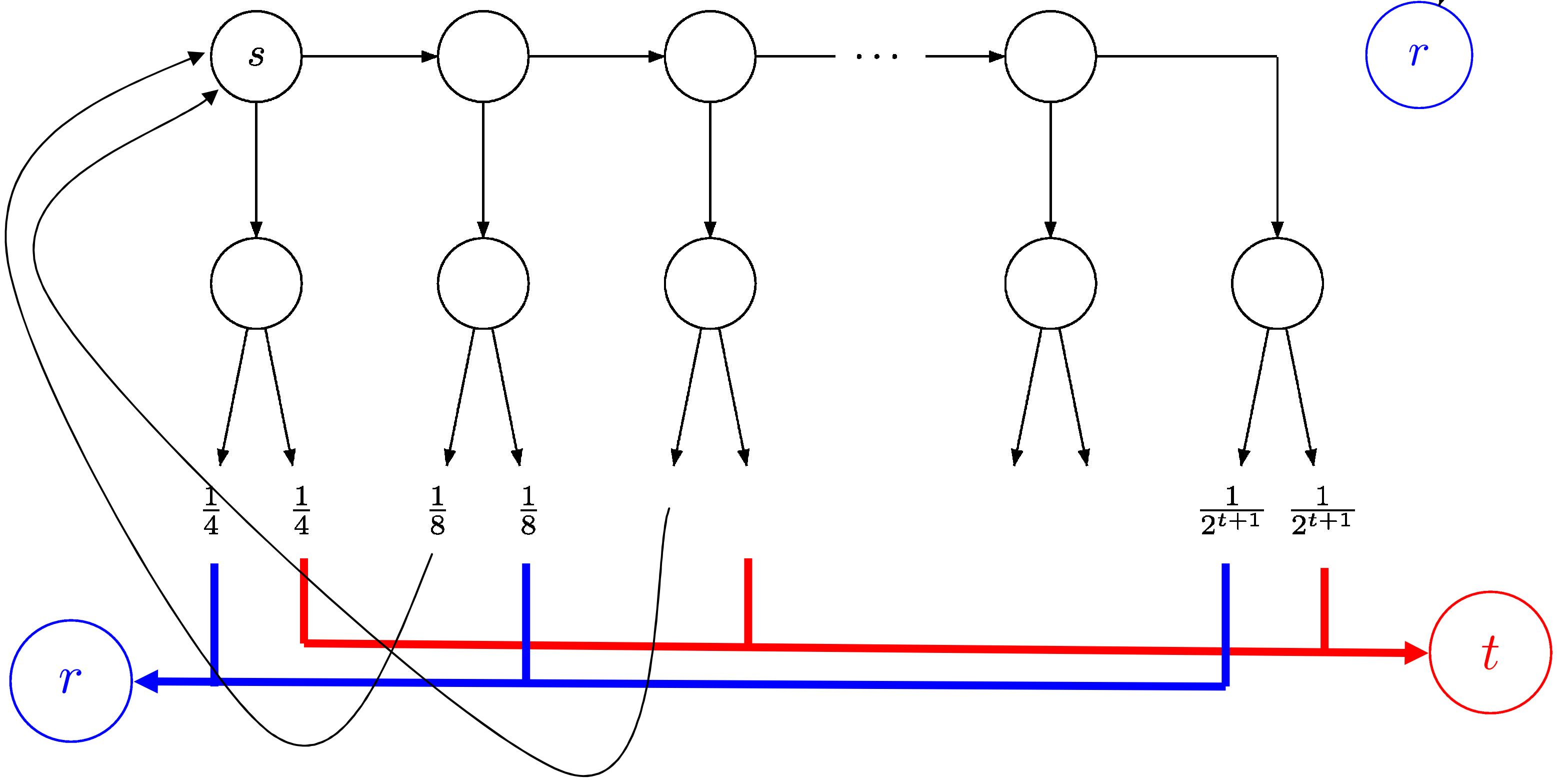


Preliminary

Rational probabilities ➡ Probabilities $\frac{1}{2}$ only

2. Use binary encoding of p and q-p to simulate [Zwick,Paterson 96] :

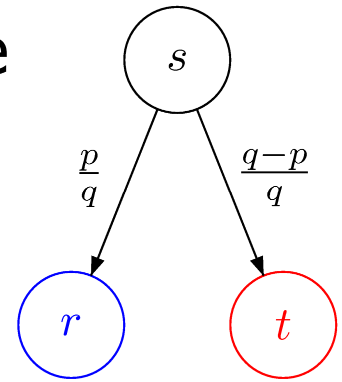


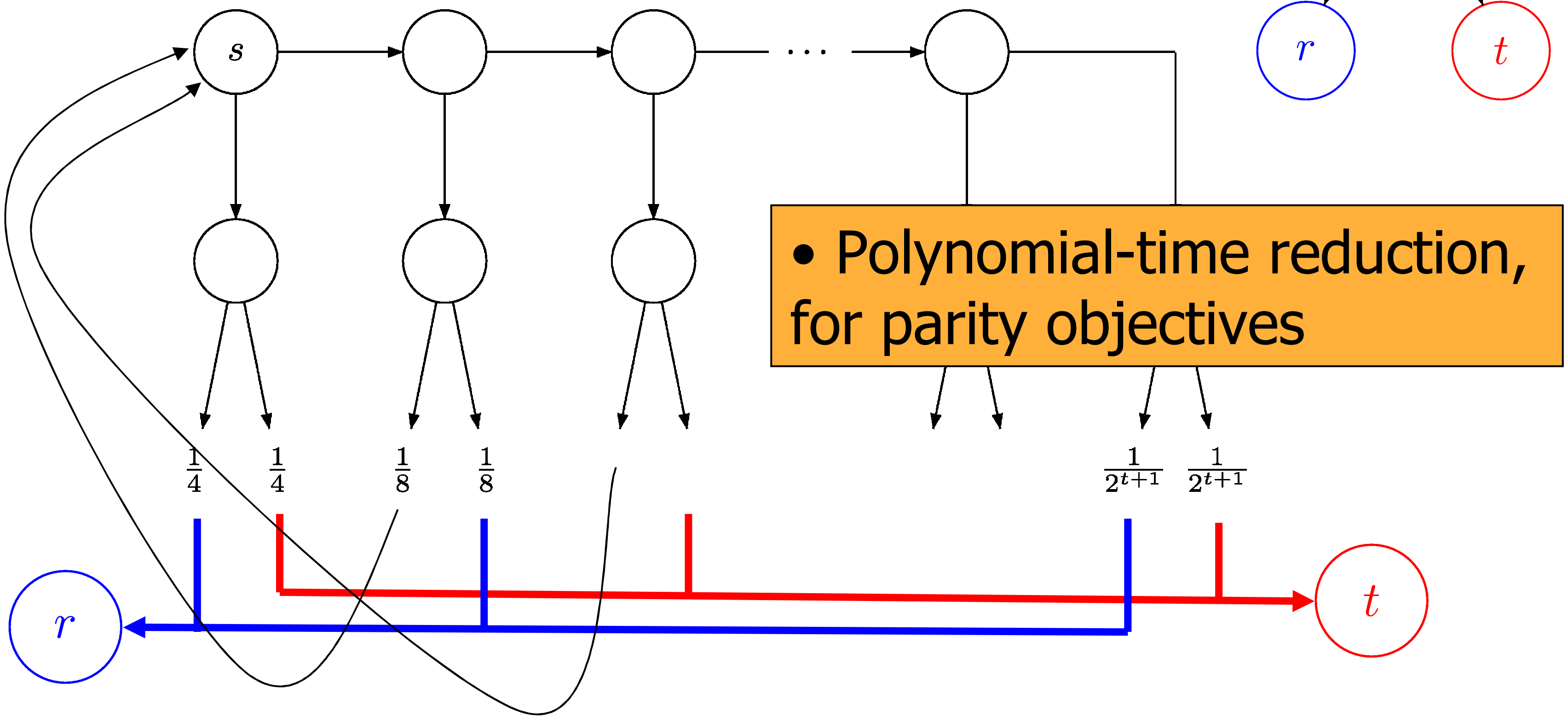


Preliminary

Rational probabilities → Probabilities $\frac{1}{2}$ only

2. Use binary encoding of p and q-p to simulate [Zwick,Paterson 96] :





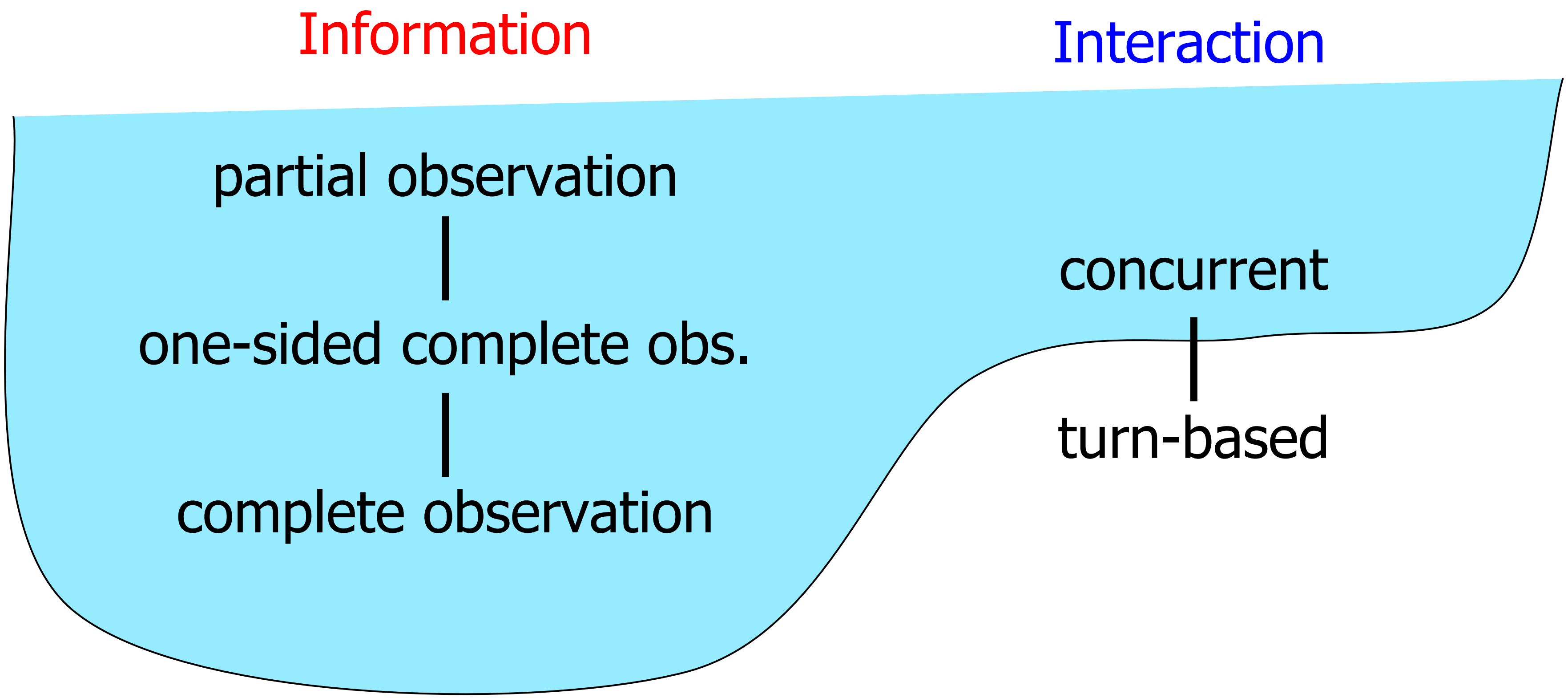
- Polynomial-time reduction, for parity objectives

Outline

- Randomness in game structure
  - for free with complete-observation, concurrent
  - for free with one-sided, turn-based
- Randomness in strategies
  - for free in (PO)MDP
- Corollary: undecidability results

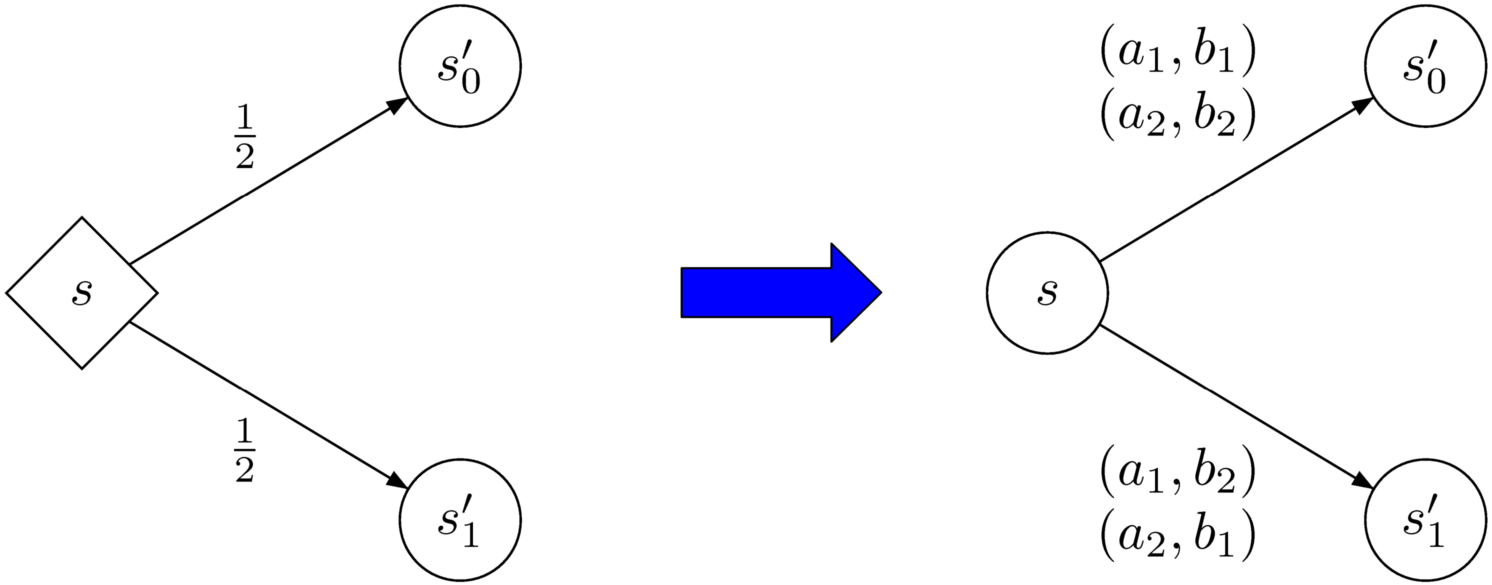
Classification

- Classification according to Information & Interaction



Reduction

Simulate probabilistic state with concurrent state



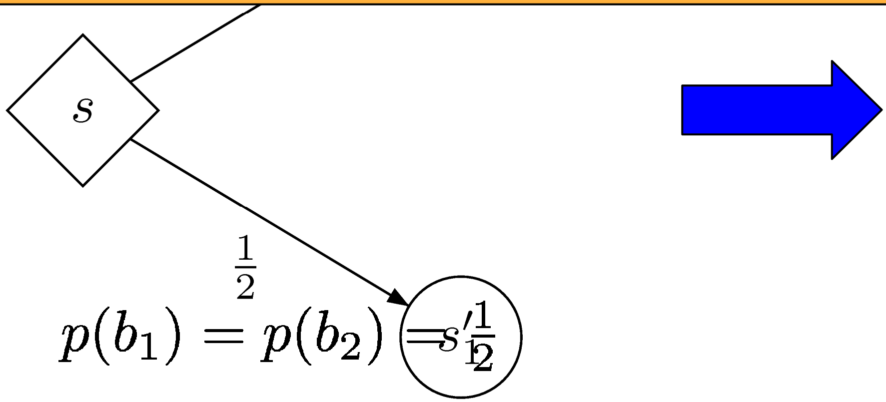
	a_1	a_2
b_1	s'_0	s'_1
b_2	s'_1	s'_0

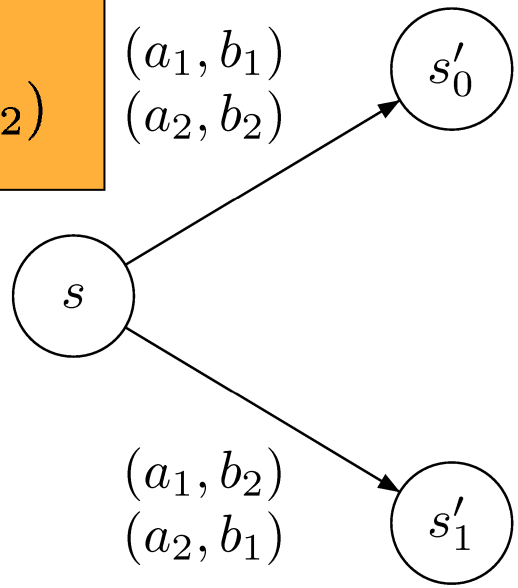
Reduction

Simulate probabilistic state with concurrent state

Probability to go from s to s'_0:

$$p(s, s'_0) = p(a_1) \cdot p(b_1) + p(a_2) \cdot p(b_2)$$





	a_1	a_2
b_1	s'_0	s'_1
b_2	s'_1	s'_0

Reduction

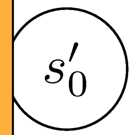
Simulate probabilistic state with concurrent state

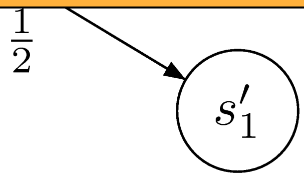
Probability to go from s to s'_0:

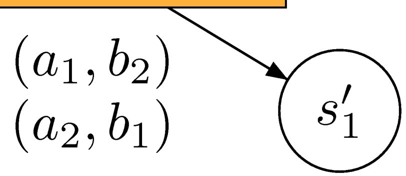
$$p(s, s'_0) = p(a_1) \cdot p(b_1) + p(a_2) \cdot p(b_2)$$

If $p(a_1) = p(a_2) = \frac{1}{2}$, then

$$p(s, s'_0) = \frac{1}{2} \cdot p(b_1) + \frac{1}{2} \cdot p(b_2) = \frac{1}{2} \cdot (p(b_1) + p(b_2)) = \frac{1}{2}$$







	a_1	a_2
b_1	s'_0	s'_1
b_2	s'_1	s'_0

Reduction

Simulate probabilistic state with concurrent state

Probability to go from s to s′$_0$:

$$p(s, s'_0) = p(a_1) \cdot p(b_1) + p(a_2) \cdot p(b_2)$$

If $p(a_1) = p(a_2) = \frac{1}{2}$, then

$$p(s, s'_0) = \frac{1}{2} \cdot p(b_1) + \frac{1}{2} \cdot p(b_2) = \frac{1}{2} \cdot (p(b_1) + p(b_2)) = \frac{1}{2}$$

If $p(b_1) = p(b_2) = \frac{1}{2}$, then

$$p(s, s'_0) = p(a_1) \cdot \frac{1}{2} + p(a_2) \cdot \frac{1}{2} = (p(a_1) + p(a_2)) \cdot \frac{1}{2} = \frac{1}{2}$$

b_1	s'_0	s'_1
b_2	s'_1	s'_0

Reduction

Simulate probabilistic state with concurrent state

Probability to g

$p(s, s'_0) =$

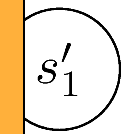
Each player can unilaterally decide to simulate the original game.

If $p(a_1) = p(a_2) = \frac{1}{2}$, then

$$p(s, s'_0) = \frac{1}{2} \cdot p(b_1) + \frac{1}{2} \cdot p(b_2) = \frac{1}{2} \cdot (p(b_1) + p(b_2)) = \frac{1}{2}$$

If $p(b_1) = p(b_2) = \frac{1}{2}$, then

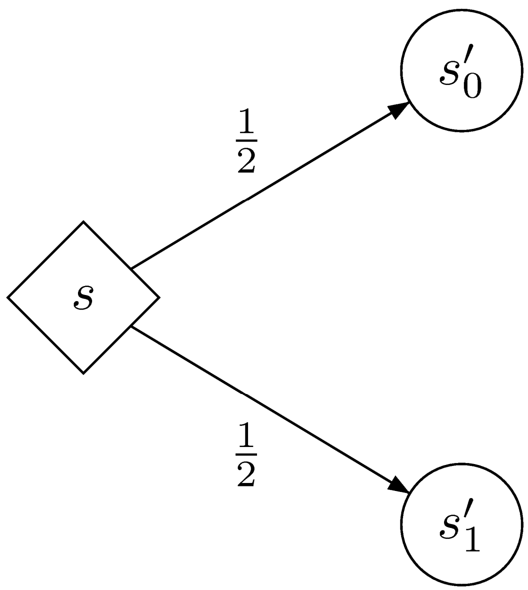
$$p(s, s'_0) = p(a_1) \cdot \frac{1}{2} + p(a_2) \cdot \frac{1}{2} = (p(a_1) + p(a_2)) \cdot \frac{1}{2} = \frac{1}{2}$$



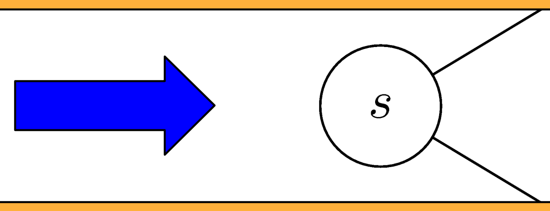
b_1	s'_0	s'_1
b_2	s'_1	s'_0

Reduction

Simulate probabilistic state with concurrent state



Player 1 can obtain at least the value v(s) by playing all actions uniformly at random: v'(s) ≥ v(s)



Player 2 can force the value to be at most v(s) by playing all actions uniformly at random: v'(s) ≤ v(s)

	a_1	a_2
b_1	s'_0	s'_1
b_2	s'_1	s'_0

Reduction

For {complete,one-sided,partial} observation, given a game with rational probabilities $\langle G, \varphi \rangle$

we can construct a concurrent game with deterministic transition function $\langle \bar{G}, \bar{\varphi} \rangle$

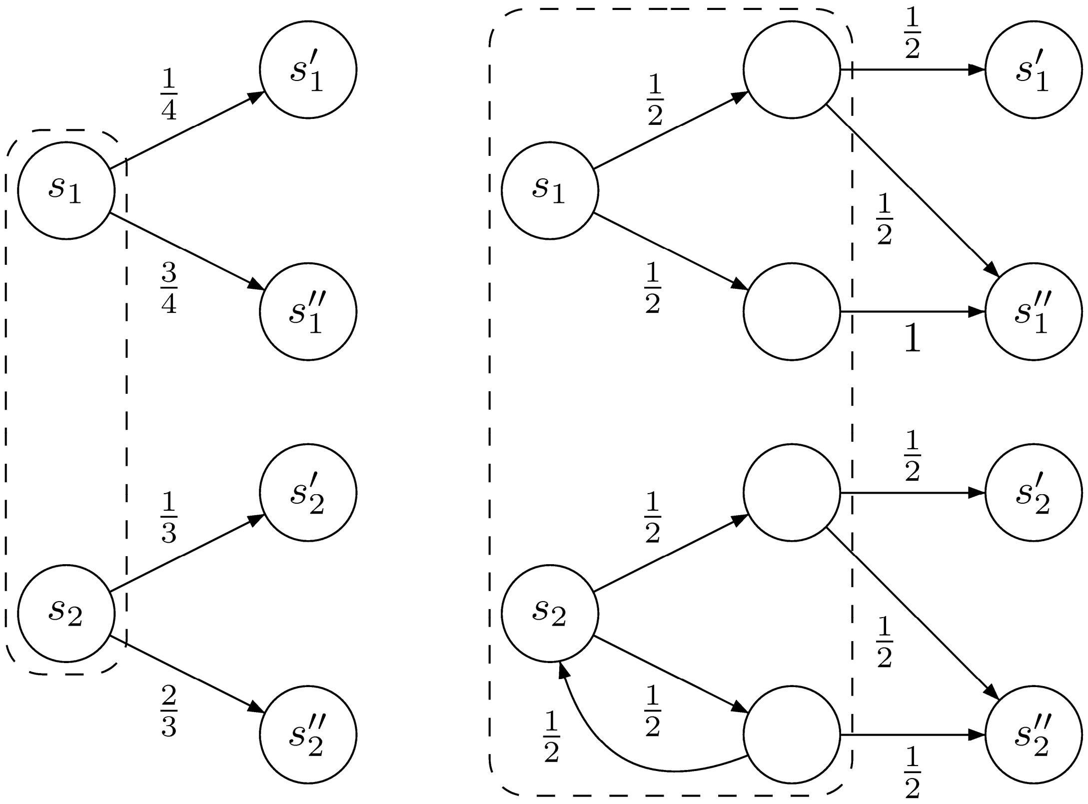
such that:

$$\langle\langle 1 \rangle\rangle_{val}^{G}(\varphi)(s) = \langle\langle 1 \rangle\rangle_{val}^{\bar{G}}(\bar{\varphi})(s).$$

and existence of optimal observation-based strategies is preserved.

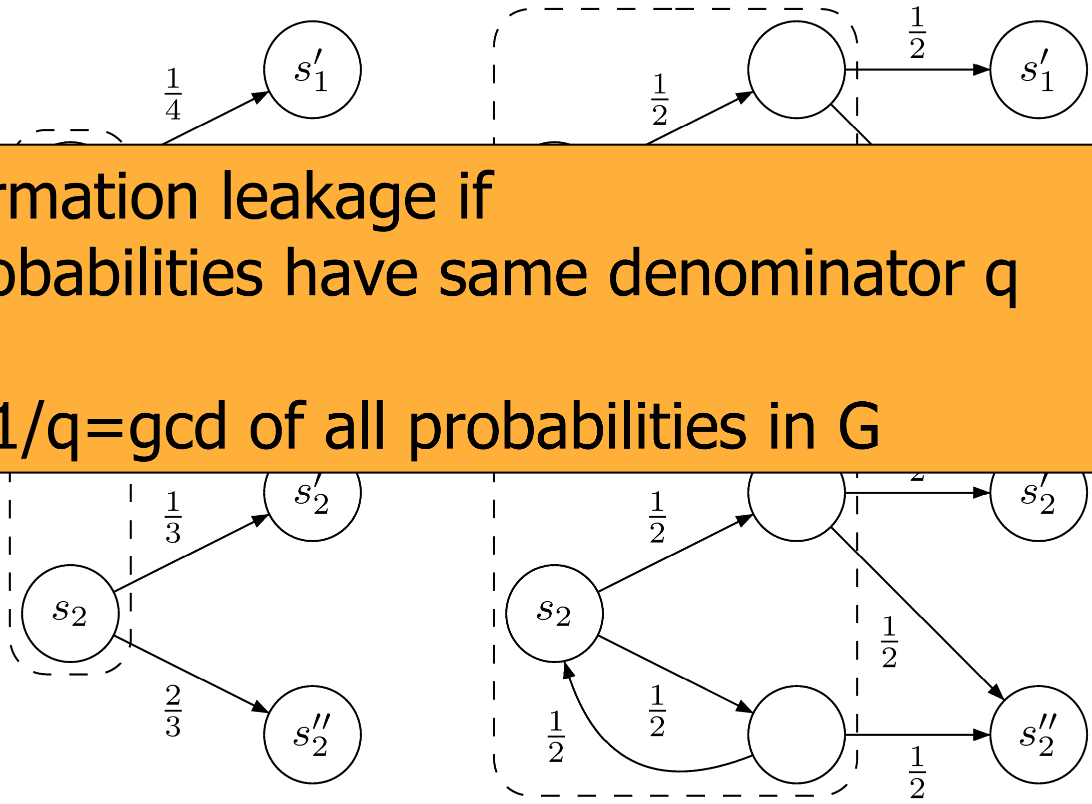
The reduction is in polynomial time for complete-observation games.

Partial information



Information leak from the back edges…

Partial information

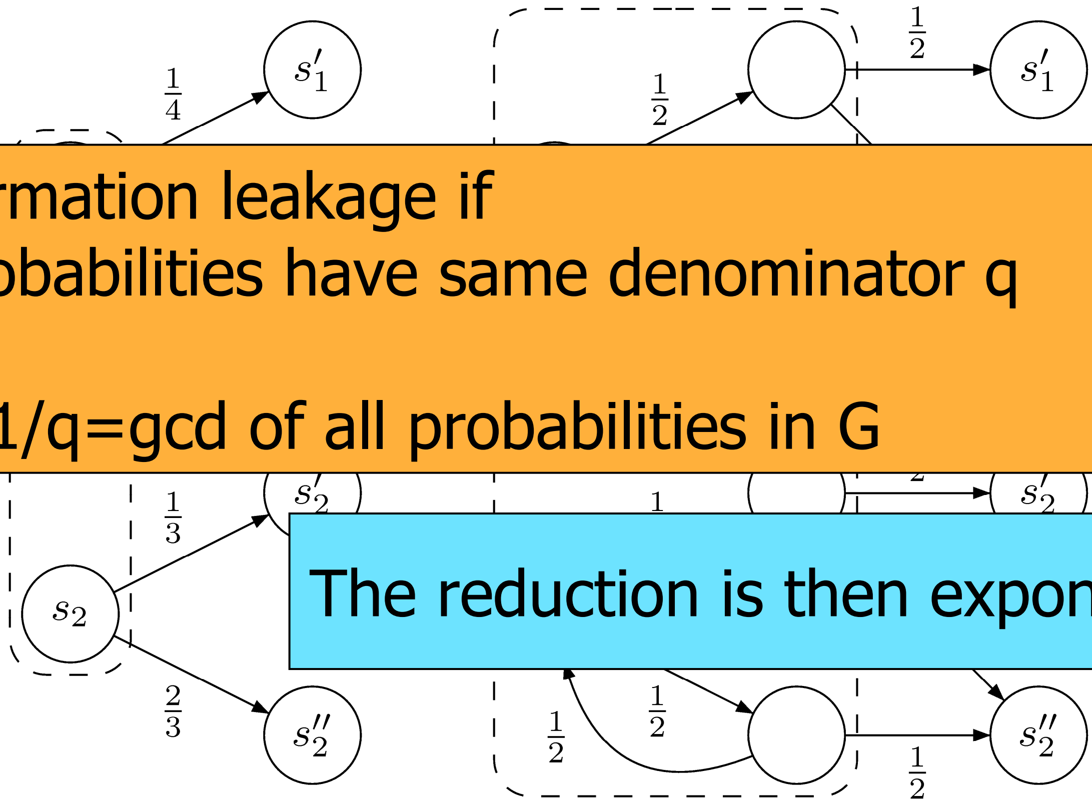


No information leakage if
all probabilities have same denominator q

➔ use 1/q=gcd of all probabilities in G

Information leak from the back edges…

Partial information



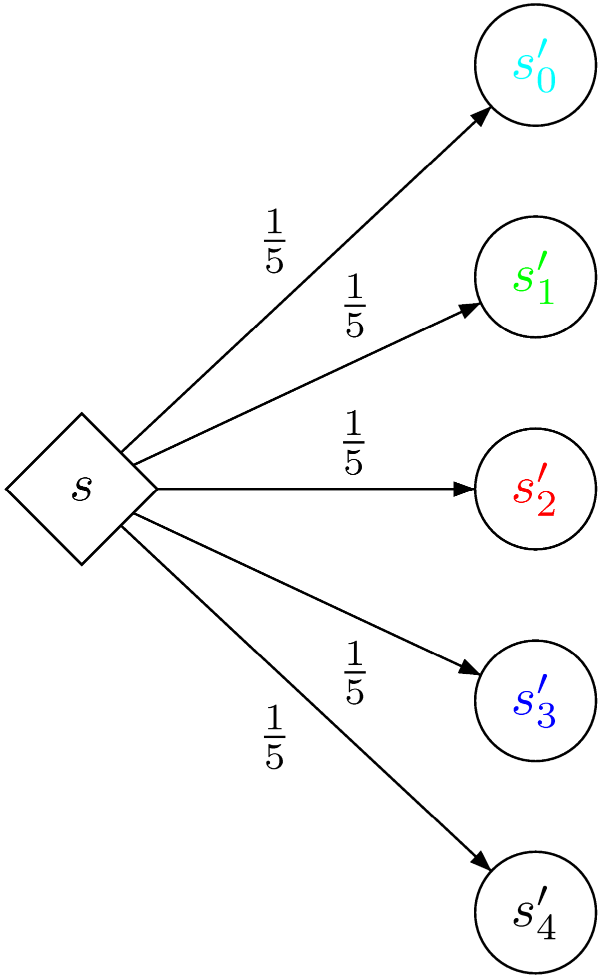
No information leakage if
all probabilities have same denominator q

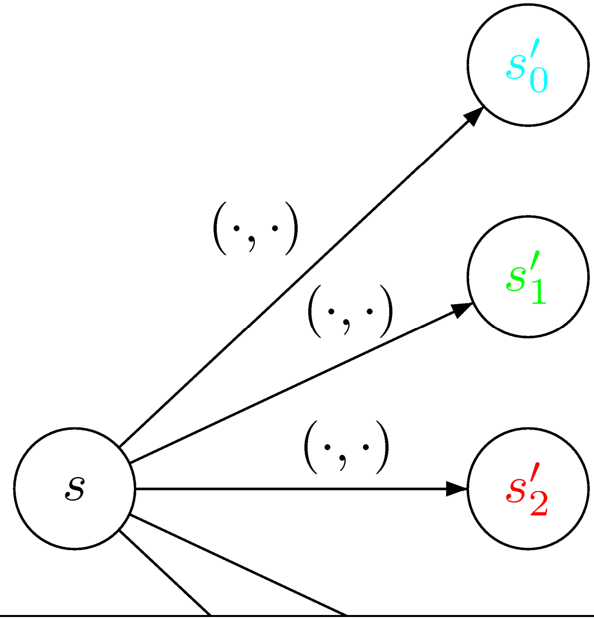
➔ use 1/q=gcd of all probabilities in G

The reduction is then exponential.

Information leak from the back edges…

Example





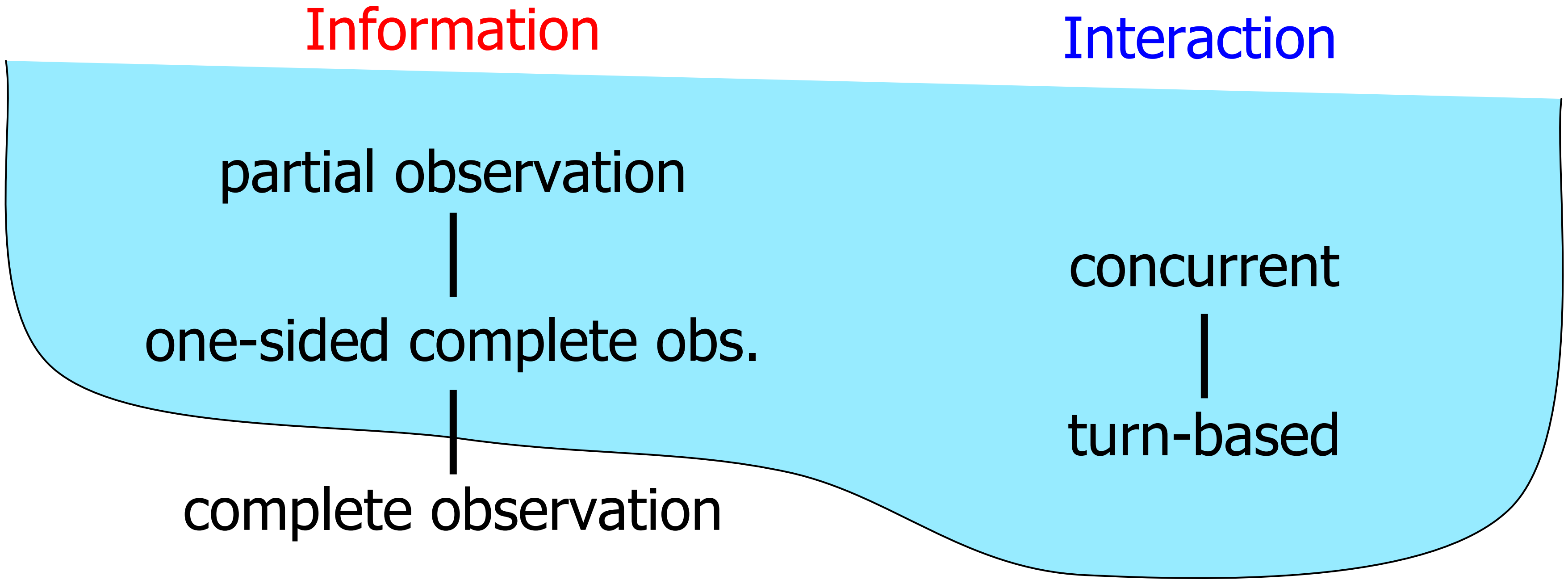
s	a_1	a_2	a_3	a_4	a_5
b_1	s'_0	s'_1	s'_2	s'_3	s'_4
b_2	s'_1	s'_2	s'_3	s'_4	s'_0
b_3	s'_2	s'_3	s'_4	s'_0	s'_1
b_4	s'_3	s'_4	s'_0	s'_1	s'_2
b_5	s'_4	s'_0	s'_1	s'_2	s'_3

Outline

- Randomness in game structure
  - for free with complete-observation, concurrent
  - for free with one-sided, turn-based
- Randomness in strategies
  - for free in (PO)MDP
- Corollary: undecidability results

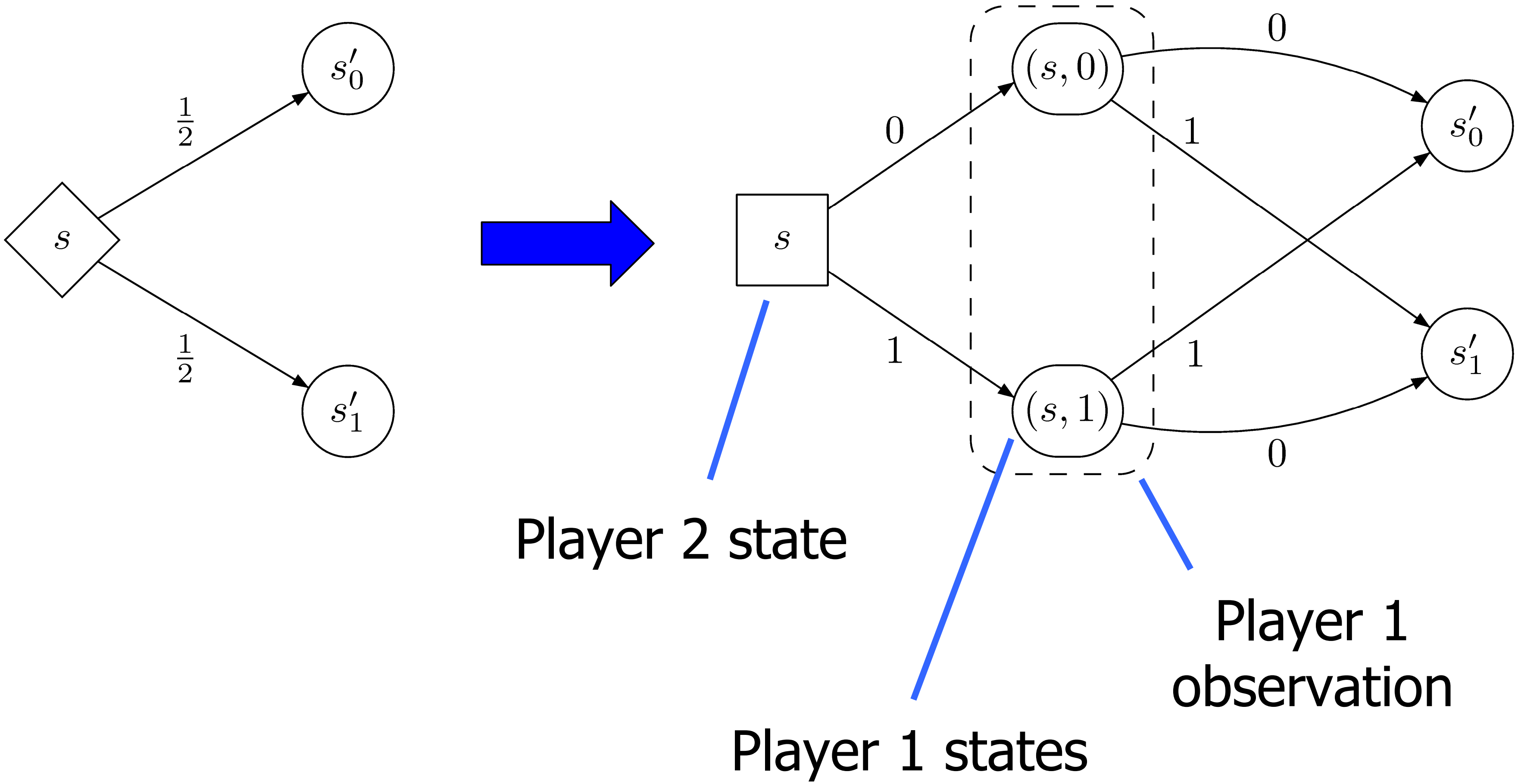
Overview

- Classification according to Information & Interaction



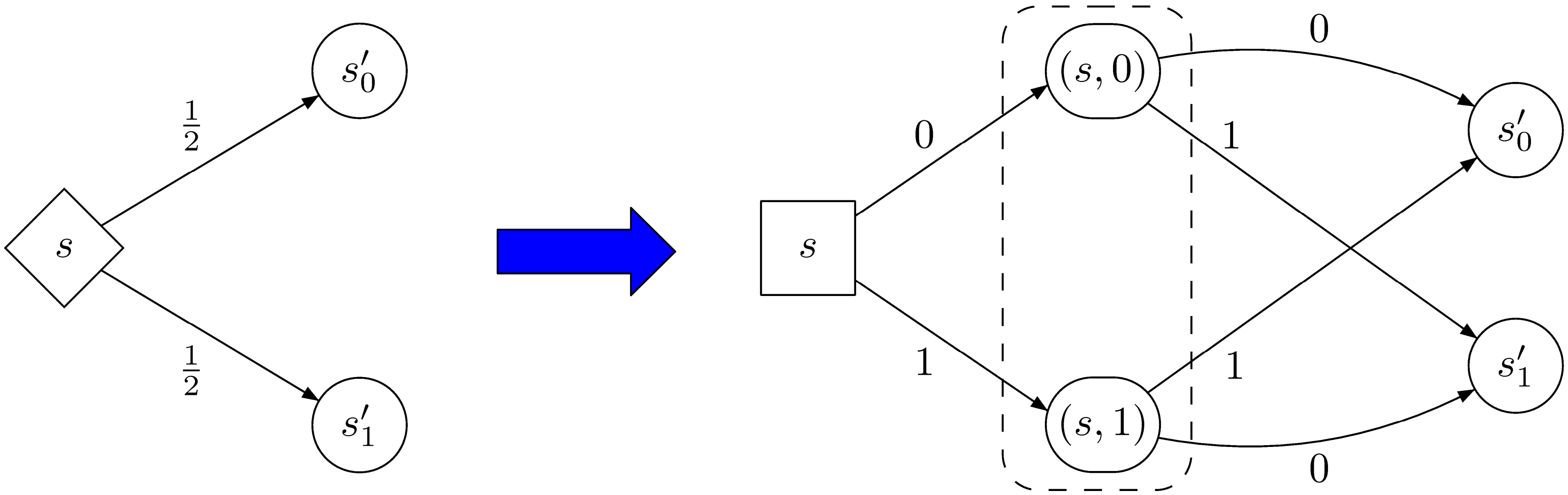
Reduction

Simulate probabilistic state with imperfect information turn-based states



Reduction

Simulate probabilistic state with imperfect information turn-based states



Each player can unilaterally decide to simulate the probabilistic state by playing uniformly at random:

Player 2 chooses states (s,0),(s,1) unifiormly at random
Player 1 chooses actions 0,1 unifiormly at random

Reduction

Games with rational probabilities can be reduced to turn-based-games with deterministic transitions and (at least) one-sided complete observation.

Values and existence of optimal strategies are preserved.

Randomness for free

In transition function (this talk):

	$2\,{}^1\!/_2$-player			$1\,{}^1\!/_2$-player	
	complete	one-sided	partial	MDP	POMDP
turn-based		free	free		
concurrent	free	free	free	(NA)	(NA)

When randomness is not for free

- Complete-information turn-based ($2\frac{1}{2}$) games

  - in deterministic games, value is either 0 or 1 [Martin98]
  - MDPs with reachability objective can have values in [0,1]

  ➡ Randomness is not for free.

- $1\frac{1}{2}$-player games (MDP & POMDP)

  - in deterministic partial info $1\frac{1}{2}$-player games, value is either 0 or 1 [see later]
  - MDPs have value in [0,1]

  ➡ Randomness is not for free.

Randomness for free

In transition function (this talk):

	$2\,{}^1\!/_2$-player			$1\,{}^1\!/_2$-player	
	complete	one-sided	partial	MDP	POMDP
turn-based	not	free	free	not	not
concurrent	free	free	free	(NA)	(NA)

Randomness for free

In transition function (this talk):

	$2\,^1\!/_2$-player			$1\,^1\!/_2$-player	
	complete	one-sided	partial	MDP	POMDP
turn-based	not	free	free	not	not
concurrent	free	free	free	(NA)	(NA)

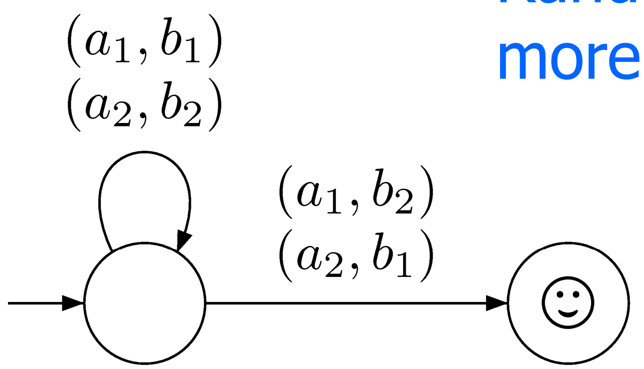
In strategies (Everett'57, Martin'98, CDHR'07):

	$2\,^1\!/_2$-player			$1\,^1\!/_2$-player	
	complete	one-sided	partial	MDP	POMDP
turn-based	$\epsilon > 0$	not	not	$\epsilon \geq 0$	?
concurrent	not	not	not	(NA)	(NA)

Randomness in strategies

Example

- concurrent, complete observation
- reachability



Randomized strategies are more powerful

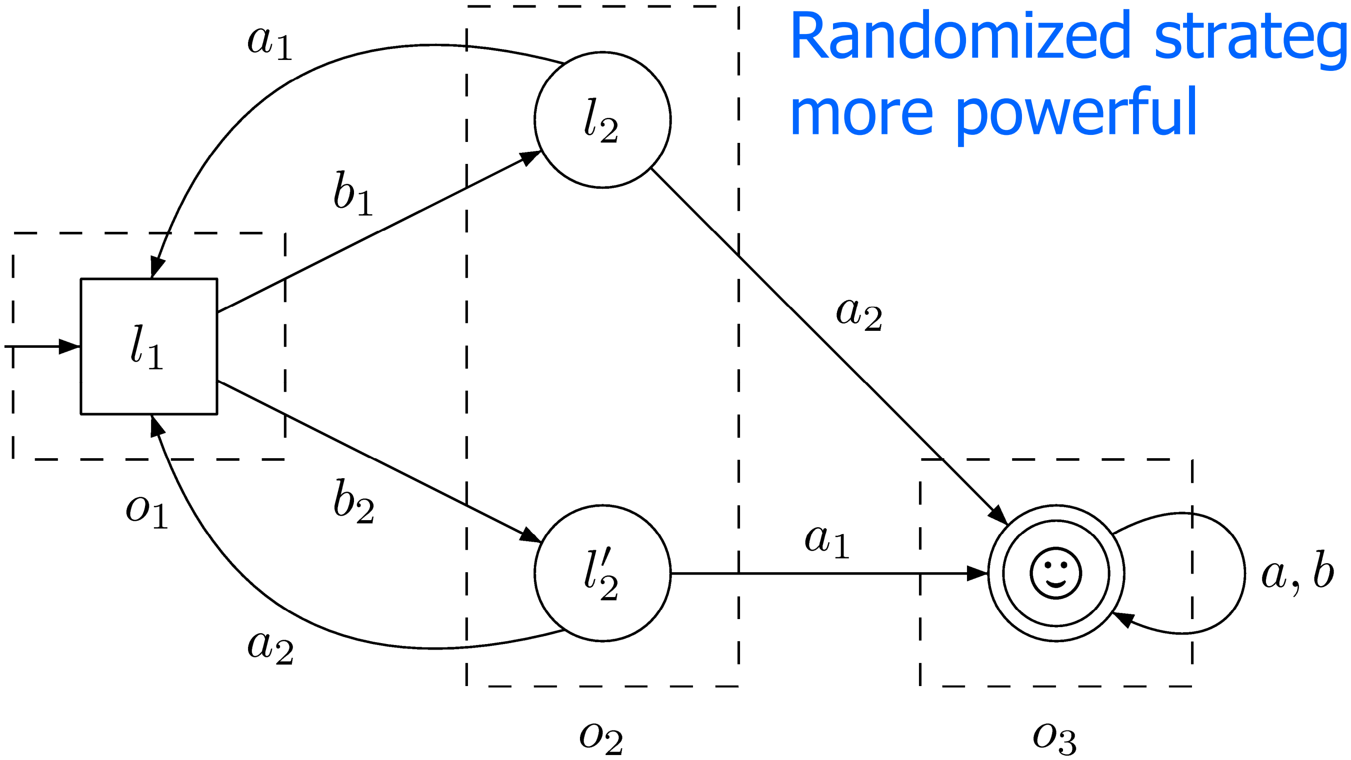
Reminder: randomized strategy for Player i: $\sigma_i : S^+ \to \mathcal{D}(A_i)$

pure strategy for Player i: $\sigma_i : S^+ \to A_i$

Randomness in strategies

Example

- turn-based, one-sided complete observation
- reachability

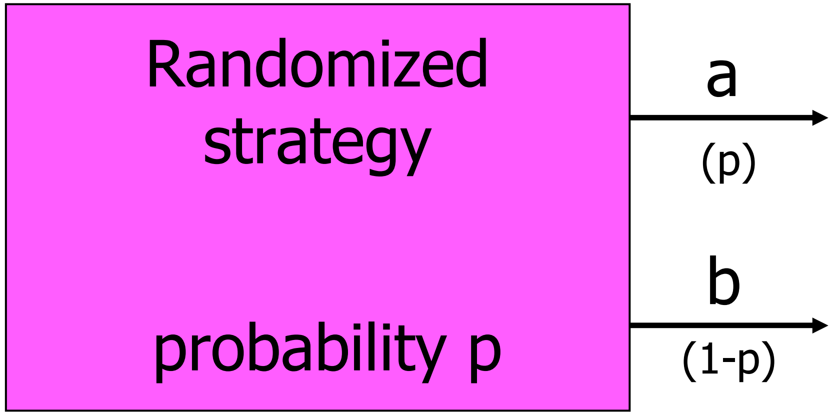


Randomized strategies are more powerful

Outline

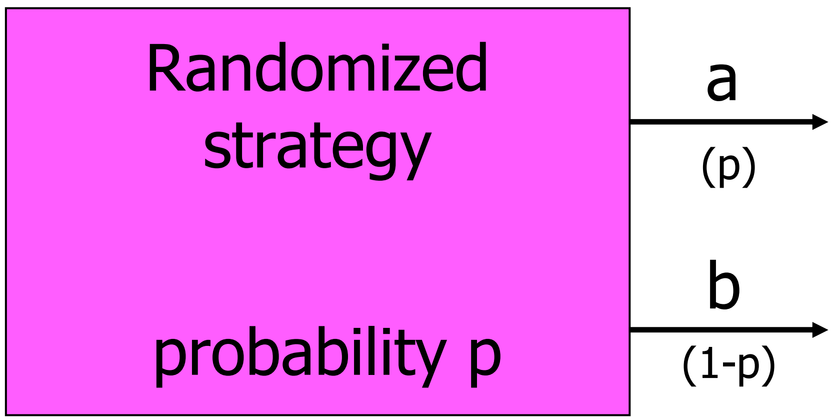
- Randomness in game structure
  - for free with complete-observation, concurrent
  - for free with one-sided, turn-based
- Randomness in strategies
  - for free in (PO)MDP
- Corollary: undecidability results

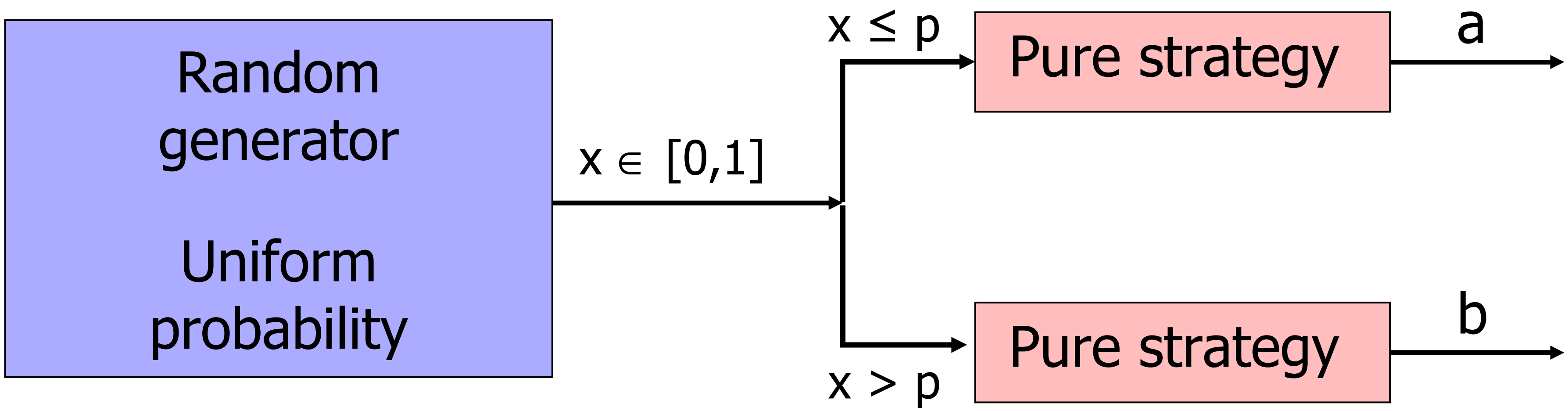
Randomness in strategies




```
RandStrategy(p)
{
   x = rand(0…1)
   if x≤p then out(a)
   else out(b)
}
```

Randomness in strategies





Randomness for free in strategies

$\langle G, \varphi \rangle$ 1½-player game (POMDP), s_0 initial state.

For every randomized observation-based strategy σ, there exists a pure observation-based strategy σ_P such that:

$$\Pr^{\sigma}(\varphi) \leq \Pr^{\sigma_P}(\varphi)$$

Proof. (assume alphabet of size 2, and fan-out = 2)

Given σ, we show that the value $\Pr^{\sigma}(\varphi)$ of σ can be obtained as the average of the value of pure strategies σ_x:

$$\Pr^{\sigma}(\varphi) = \int_{\mathcal{D}} \Pr^{\sigma_x}(\varphi) \; d\nu$$

Randomness for free in strategies

Strategies σ_x are obtained by «de-randomization» of σ

Assume $\sigma(s_1)(a) = p$ and $\sigma(s_1)(b) = 1-p$

σ is equivalent to playing σ_0 with frequency p
and σ_1 with frequency 1-p where:

Randomness for free in strategies

Strategies σ_x are obtained by «de-randomization» of σ

Assume $\sigma(s_1)(a) = p$ and $\sigma(s_1)(b) = 1-p$

σ is equivalent to playing σ_0 with frequency p and σ_1 with frequency 1-p where:

$$\sigma_0(\rho) \begin{cases} \text{plays } a & \text{if } \rho = s \\ \text{plays like } \sigma(\rho) & \text{otherwise} \end{cases}$$

$$\sigma_1(\rho) \begin{cases} \text{plays } b & \text{if } \rho = s \\ \text{plays like } \sigma(\rho) & \text{otherwise} \end{cases}$$

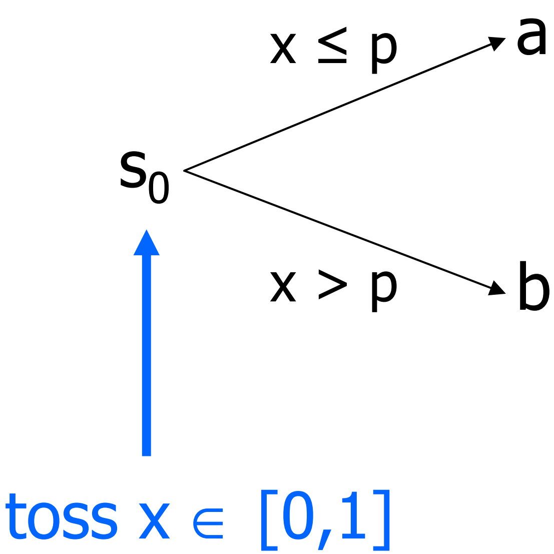
$$\mathrm{Pr}^{\sigma}(\varphi) = p \cdot \mathrm{Pr}^{\sigma_0}(\varphi) + (1-p) \cdot \mathrm{Pr}^{\sigma_1}(\varphi)$$

Randomness for free in strategies

Equivalently, toss a coin $x \in [0,1]$,
play σ_0 if $x \leq p$, and play σ_1 if $x > p$.

[$p = \sigma(s_1)(a)$]

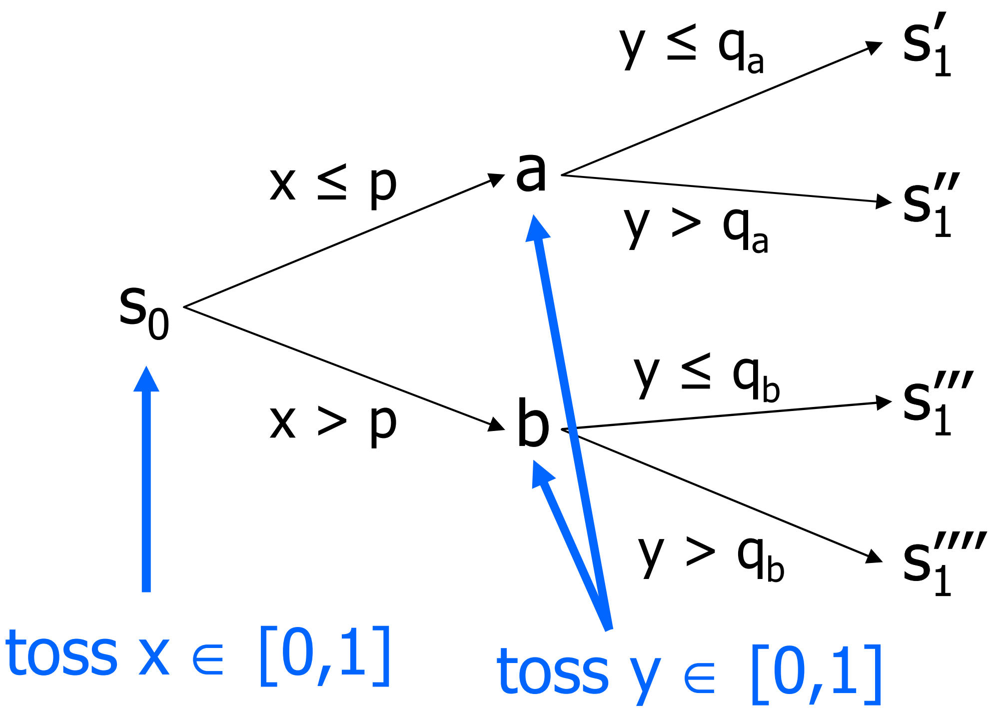
Playing σ in G can be viewed as a sequence of coin tosses:



Randomness for free in strategies

Equivalently, toss a coin $x \in [0,1]$,
play σ_0 if $x \leq p$, and play σ_1 if $x > p$. [$p = \sigma(s_1)(a)$]

Playing σ in G can be viewed as a sequence of coin tosses:



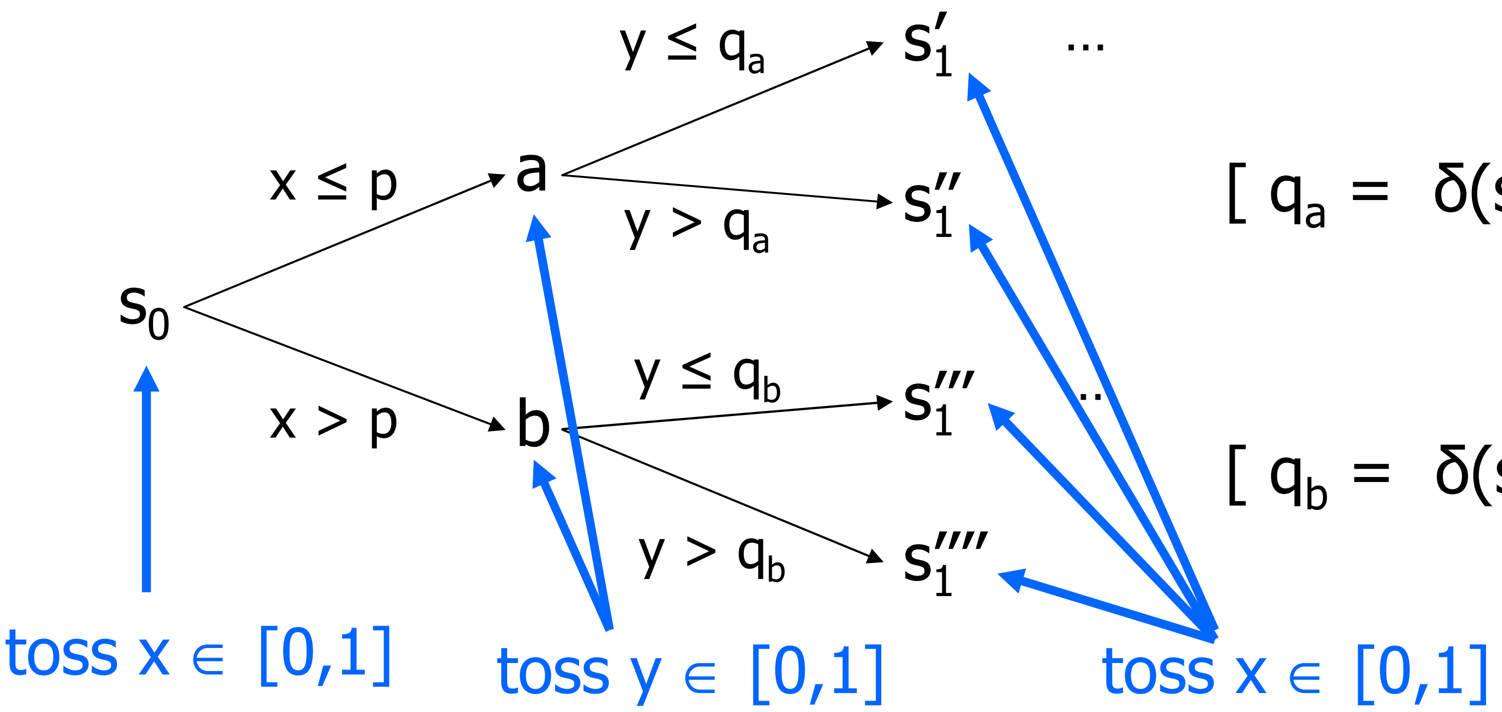
[$q_a = \delta(s_1,a)(s_1')$]

[$q_b = \delta(s_1,b)(s_1''')$]

Randomness for free in strategies

Equivalently, toss a coin $x \in [0,1]$, play σ_0 if $x \leq p$, and play σ_1 if $x > p$. $[\ p = \sigma(s_1)(a)\ ]$

Playing σ in G can be viewed as a sequence of coin tosses:



$[\ q_a = \delta(s_1,a)(s_1')\ ]$

$[\ q_b = \delta(s_1,b)(s_1''')\ ]$

Randomness for free in strategies

Given an infinite sequence $x=(x_n)_{n\geq 0} \in [0,1]^\omega$, define for all s_0, s_1, …, s_n:

$$\sigma_x(s_0, s_1, \ldots, s_n) = \begin{cases} a & \text{if } x_n \leq \sigma(s_0, s_1, \ldots, s_n)(a) \\ b & \text{otherwise.} \end{cases}$$

σ_x is a pure and observation-based strategy !

σ_x plays like σ, assuming that the result of the coin tosses is the sequence x.

The value $\mathrm{Pr}^\sigma(\varphi)$ of σ is the « average of the outcome » of the strategies σ_x.

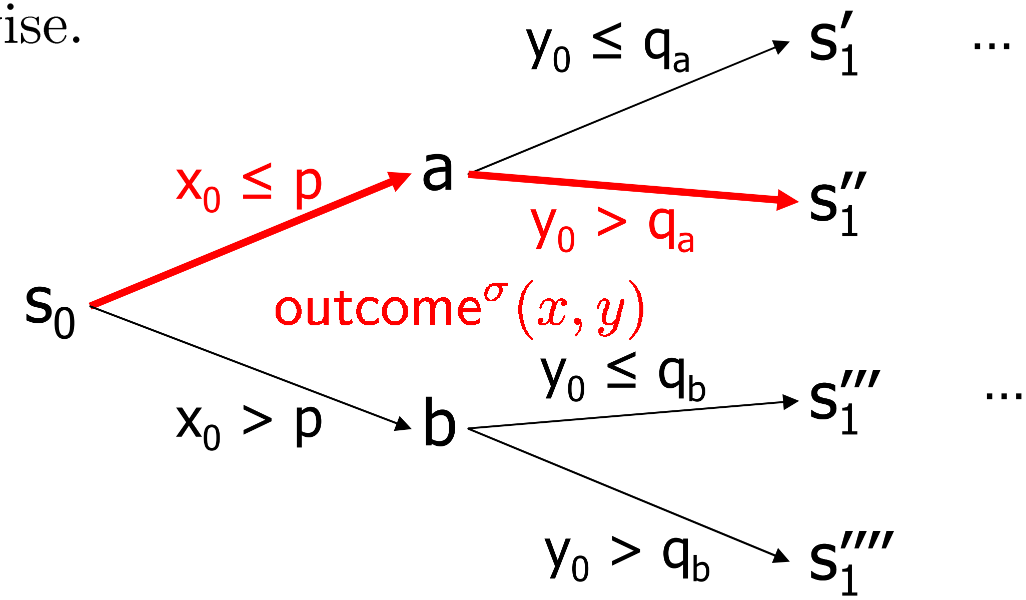
Randomness for free in strategies

Assume $x=(x_n)_{n\geq 0}$ and $y=(y_n)_{n\geq 0}$ are fixed.

Let $\mathsf{outcome}^{\sigma}(x,y) = s_0 a_1 s_1 a_2 s_2 \ldots$ where

$$a_{n+1} = \begin{cases} a & \text{if } x_n \leq \sigma(s_0 s_1 \cdots s_n)(a), \\ b & \text{otherwise.} \end{cases}$$

$$s_{n+1} = \begin{cases} L(s_n, a_{n+1}) & \text{if } y_n \leq \delta(s_n, a_{n+1})(L(s_n, a_{n+1})), \\ R(s_n, a_{n+1}) & \text{otherwise.} \end{cases}$$



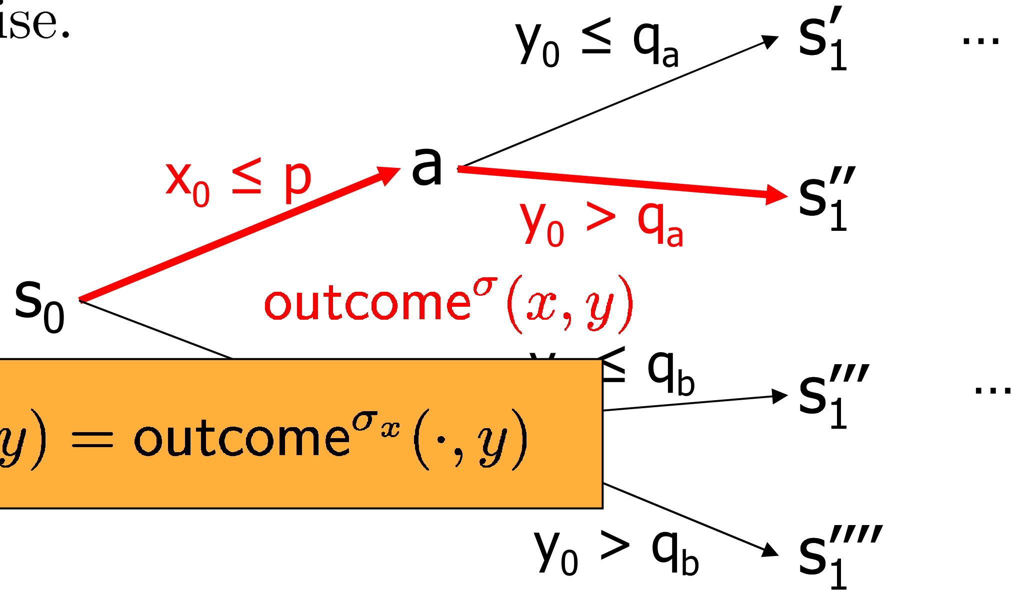
Randomness for free in strategies

Assume $x=(x_n)_{n\geq 0}$ and $y=(y_n)_{n\geq 0}$ are fixed.

Let $\text{outcome}^{\sigma}(x,y) = s_0 a_1 s_1 a_2 s_2 \ldots$ where

$$a_{n+1} = \begin{cases} a & \text{if } x_n \leq \sigma(s_0 s_1 \cdots s_n)(a), \\ b & \text{otherwise.} \end{cases}$$

$$s_{n+1} = \begin{cases} L(s_n, a_{n+1}) & \text{if } y_n \leq \delta(s_n, a_{n+1})(L(s_n, a_{n+1})), \\ R(s_n, a_{n+1}) & \text{otherwise.} \end{cases}$$



Theorem: $\text{outcome}^{\sigma}(x,y) = \text{outcome}^{\sigma_x}(\cdot,y)$

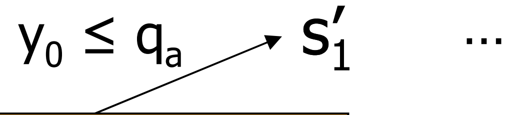
Randomness for free in strategies

Assume $x=(x_n)_{n\geq 0}$ and $y=(y_n)_{n\geq 0}$ are fixed.

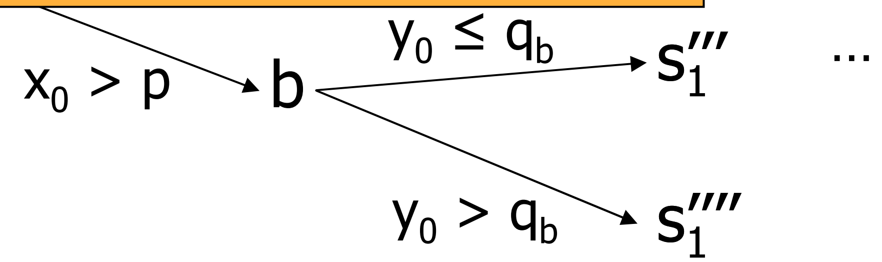
Let $\mathtt{outcome}^{\sigma}(x,y) = s_0 a_1 s_1 a_2 s_2 \ldots$ where

$$a_{n+1} = \begin{cases} a & \text{if } x_n \leq \sigma(s_0 s_1 \cdots s_n)(a), \\ b & \text{otherwise.} \end{cases}$$

$$s_{n+1} = \begin{cases} L(s_n, a_{n+1}) & \text{if } y_n \leq \delta(s_n, a_{n+1})(L(s_n, a_{n+1})), \\ R(s_n, a_{n+1}) & \text{otherwise.} \end{cases}$$



Playing σ in G is equivalent to choosing (x_n) and (y_n) uniformly at random, and then producing $\mathtt{outcome}^{\sigma}(x,y)$



Randomness for free in strategies

Assume $x=(x_n)_{n\geq 0}$ and $y=(y_n)_{n\geq 0}$ are fixed.

Let $\text{outcome}^{\sigma}(x,y) = s_0 a_1 s_1 a_2 s_2 \ldots$ where

$$a_{n+1} = \begin{cases} a & \text{if } x_n \leq \sigma(s_0 s_1 \cdots s_n)(a), \\ b & \text{otherwise.} \end{cases}$$

$$s_{n+1} = \begin{cases} L(s_n, a_{n+1}) & \text{if } y_n \leq \delta(s_n, a_{n+1})(L(s_n, a_{n+1})), \\ R(s_n, a_{n+1}) & \text{otherwise.} \end{cases}$$

$y_0 \leq q_a$ → s_1' …

Playing σ in G is equivalent to choosing (x_n) and (y_n) uniformly at ra

The random sequences (x_n) and (y_n) can be generated separately, and independently !

…

$y_0 > q_b$ → s_1''''

Proof

The value $\mathrm{Pr}^{\sigma}(\varphi)$ of σ is the « average of the outcome » of the strategies σ_x.

$$\mathrm{Pr}^{\sigma}(\varphi) = \int_{p \in (SA)^{\omega}} \mathbf{1}_{\varphi}(p)\ d\mu^{\sigma}(p)$$

$$= \int_{(x,y) \in [0,1]^{\omega} \times [0,1]^{\omega}} \mathbf{1}_{\varphi}(\mathsf{outcome}^{\sigma}(x,y))\ d(\nu \times \nu)(x,y)$$

$$= \int_{x \in [0,1]^{\omega}} \left(\int_{y \in [0,1]^{\omega}} \mathbf{1}_{\varphi}(\mathsf{outcome}^{\sigma}(x,y))\ d\nu(y) \right)\ d\nu(x)$$

$$= \int_{x \in [0,1]^{\omega}} \left(\int_{y \in [0,1]^{\omega}} \mathbf{1}_{\varphi}(\mathsf{outcome}^{\sigma_x}(\cdot,y))\ d\nu(y) \right)\ d\nu(x)$$

$$= \int_{x \in [0,1]^{\omega}} \mathrm{Pr}^{\sigma_x}(\varphi)\ d\nu(x)$$

Proof

The value $\Pr^\sigma(\varphi)$ of σ is the « average of the outcome » of the strategies σ_x.

$$\Pr^\sigma(\varphi) = \int_{p \in (SA)^\omega} \mathbf{1}_\varphi(p)\ d\mu^\sigma(p)$$

$$\Pr^\sigma(\varphi) \leq \Pr^{\sigma_x}(\varphi) \text{ for some } \sigma_x$$

$$\ldots^\sigma(x,y))\ d(\nu \times \nu)(x,y)$$

$$= \int_{x \in \ldots} \left(\int \ldots \mathbf{1}_\varphi(\text{outcome}^\sigma(x,y))\ d\nu(y) \right)\ d\nu(x)$$

Pure and randomized strategies are equally powerful in POMDPs !

$$= \int_{x \in \ldots} \ldots$$

$$= \int_{x \in [0,1]^\omega} \Pr^{\sigma_x}(\varphi)\ d\nu(x)$$

Randomness for free

In transition function:

	$2\,^1\!/_2$-player			$1\,^1\!/_2$-player	
	complete	one-sided	partial	MDP	POMDP
turn-based	not	free	free	not	not
concurrent	free	free	free	(NA)	(NA)

In strategies:

	$2\,^1\!/_2$-player			$1\,^1\!/_2$-player	
	complete	one-sided	partial	MDP	POMDP
turn-based	$\epsilon > 0$	not	not	$\epsilon \geq 0$	$\epsilon \geq 0$
concurrent	not	not	not	(NA)	(NA)

Corollary

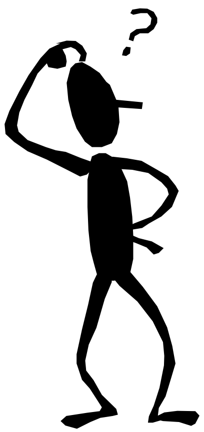
Using [BaierBertrandGrößer'08]:

Almost-sure coBüchi (and positive Büchi) with randomized (or pure) strategies is undecidable for POMDPs.

Using randomness for free in transition functions :

Almost-sure coBüchi (and positive Büchi) is undecidable for deterministic turn-based one-sided complete observation games.

Thank you !



Questions ?