Control, probabilities and partial observation

Nathalie Bertrand
Inria Rennes Bretagne Atlantique

MOVEP 2016, Genova

1. Probabilistic automata
   - Presentation
   - Stochastic languages
   - Decision problems

2. Partially observable MDP
   - Presentation
   - POMDP analysis
   - Application to control for fault diagnosis

3. Conclusion

1. Probabilistic automata
   - Presentation
   - Stochastic languages
   - Decision problems

2. Partially observable MDP
   - Presentation
   - POMDP analysis
   - Application to control for fault diagnosis

3. Conclusion

An introductive example

Holiday planning

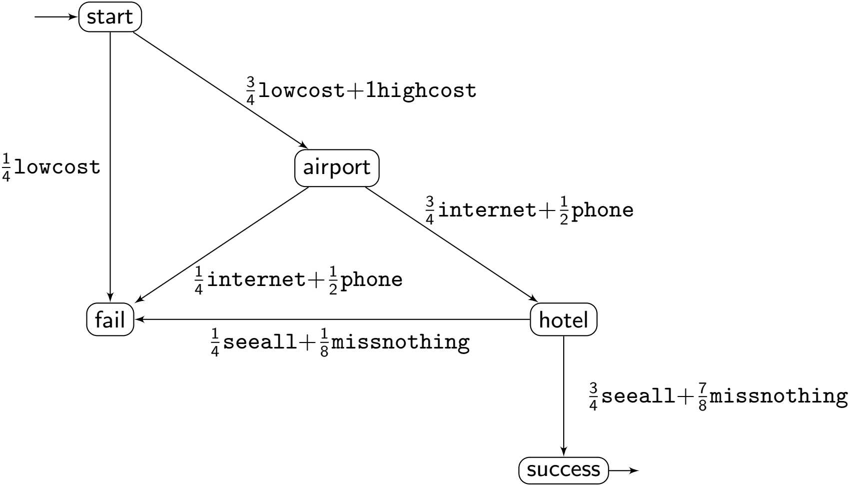
1. Choose and airline type `lowcost` or `highcost` ;
2. Book an accommodation on the `internet` or by `phone`;
3. Choose a tour `seeall` or `missnothing`.

Each action

- must be planned before holidays;
- may fail with some probability.

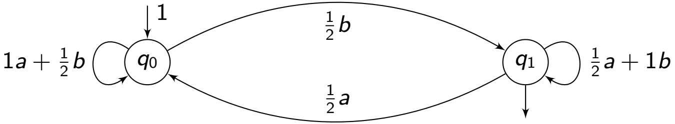
A possible plan: $\texttt{lowcost} \cdot \texttt{internet} \cdot \texttt{seeall}$

Example formalisation



The success probability of `lowcost` · `internet` · `seeall` is equal to $\frac{27}{64}$.

Probabilistic automata



Probabilistic automata

A PA $\mathcal{A} = (Q, A, \{\mathbf{P}_a\}_{a \in A}, \pi_0, F)$ is defined by:

- Q, a finite set of states; $Q = \{q_0, q_1\}$
- A, a finite alphabet of actions; $A = \{a, b\}$
- for every $a \in A$, a stochastic matrix $\mathbf{P}_a$ indexed by Q
  i.e. for every $q, q' \in Q$, $\mathbf{P}_a[q, q'] \geq 0$ and $\sum_{q' \in Q} \mathbf{P}_a[q, q'] = 1$;

$$\mathbf{P}_a = \begin{matrix} 1 & 0 \\ .5 & .5 \end{matrix} \qquad \mathbf{P}_b = \begin{matrix} .5 & .5 \\ 0 & 1 \end{matrix}$$

- π_0, the initial distribution over states; $\pi_0[q_0] = 1$
- $F \subseteq Q$, a subset of final states. $F = \{q_1\}$

Label $1a + \frac{1}{2}b$ on the loop at q_0 means $\mathbf{P}_a[q_0, q_0] = 1$ and $\mathbf{P}_b[q_0, q_0] = \frac{1}{2}$.

Control in PA

Strategies are words
what is the probability to reach a final state after word w?

Acceptance probability

The *acceptance probability* of $w = a_1 \dots a_n$ by $\mathcal{A}$ is:

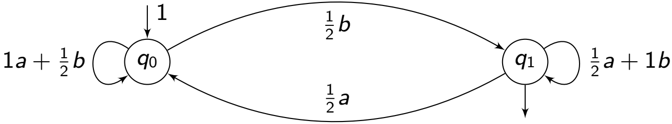
$$\mathbf{Pr}_{\mathcal{A}}(w) = \sum_{q \in Q} \pi_{\mathbf{0}}[q] \sum_{q' \in F} \left(\prod_{i=1}^{n} \mathbf{P}_{a_i} \right) [q, q']$$

For short

$$\mathbf{Pr}_{\mathcal{A}}(w) = \pi_0 \mathbf{P}_w \mathbf{1}_F^T$$

where $\mathbf{P}_w = \prod_{i=1}^{n} \mathbf{P}_{a_i}$ and $\mathbf{1}_F$ is the indicating vector of subset F.

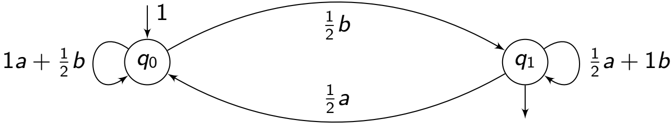
Illustration



Inductive computation of $\mathbf{Pr}_{\mathcal{A}}(abba)$ from $\mathbf{Pr}_{\mathcal{A}}(\varepsilon) = 0$.

- $\mathbf{Pr}_{\mathcal{A}}(a) = \frac{1}{2}\mathbf{Pr}_{\mathcal{A}}(\varepsilon) = 0$
- $\mathbf{Pr}_{\mathcal{A}}(ab) = \mathbf{Pr}_{\mathcal{A}}(a) + \frac{1}{2}(1 - \mathbf{Pr}_{\mathcal{A}}(a)) = \frac{1}{2}$
- $\mathbf{Pr}_{\mathcal{A}}(abb) = \mathbf{Pr}_{\mathcal{A}}(ab) + \frac{1}{2}(1 - \mathbf{Pr}_{\mathcal{A}}(ab)) = \frac{3}{4}$
- $\mathbf{Pr}_{\mathcal{A}}(abba) = \frac{1}{2}\mathbf{Pr}_{\mathcal{A}}(abb) = \frac{3}{8}$

Illustration



Inductive computation of $\mathbf{Pr}_{\mathcal{A}}(abba)$ from $\mathbf{Pr}_{\mathcal{A}}(\varepsilon) = 0$.

- $\mathbf{Pr}_{\mathcal{A}}(a) = \frac{1}{2}\mathbf{Pr}_{\mathcal{A}}(\varepsilon) = 0$
- $\mathbf{Pr}_{\mathcal{A}}(ab) = \mathbf{Pr}_{\mathcal{A}}(a) + \frac{1}{2}(1 - \mathbf{Pr}_{\mathcal{A}}(a)) = \frac{1}{2}$
- $\mathbf{Pr}_{\mathcal{A}}(abb) = \mathbf{Pr}_{\mathcal{A}}(ab) + \frac{1}{2}(1 - \mathbf{Pr}_{\mathcal{A}}(ab)) = \frac{3}{4}$
- $\mathbf{Pr}_{\mathcal{A}}(abba) = \frac{1}{2}\mathbf{Pr}_{\mathcal{A}}(abb) = \frac{3}{8}$

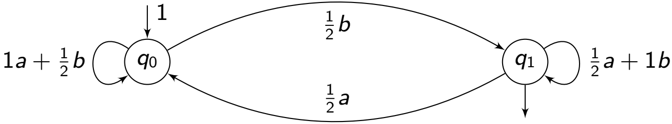
In general:

$$\mathbf{Pr}_{\mathcal{A}}(wa) = \frac{1}{2}\mathbf{Pr}_{\mathcal{A}}(w) \quad \text{and} \quad \mathbf{Pr}_{\mathcal{A}}(wb) = \frac{1}{2}(1 + \mathbf{Pr}_{\mathcal{A}}(w))$$

Thus giving an explicit acceptance probability:

$$\mathbf{Pr}_{\mathcal{A}}(a_1 \dots a_n) = \sum_{i=1}^{n} 2^{i-n-1} \cdot \mathbf{1}_{a_i=b}$$

Illustration



Inductive computation of $\mathbf{Pr}_{\mathcal{A}}(abba)$ from $\mathbf{Pr}_{\mathcal{A}}(\varepsilon) = 0$.

- $\mathbf{Pr}_{\mathcal{A}}(a) = \frac{1}{2}\mathbf{Pr}_{\mathcal{A}}(\varepsilon) = 0$
- $\mathbf{Pr}_{\mathcal{A}}(ab) = \mathbf{Pr}_{\mathcal{A}}(a) + \frac{1}{2}(1 - \mathbf{Pr}_{\mathcal{A}}(a)) = \frac{1}{2}$
- $\mathbf{Pr}_{\mathcal{A}}(abb) = \mathbf{Pr}_{\mathcal{A}}(ab) + \frac{1}{2}(1 - \mathbf{Pr}_{\mathcal{A}}(ab)) = \frac{3}{4}$
- $\mathbf{Pr}_{\mathcal{A}}(abba) = \frac{1}{2}\mathbf{Pr}_{\mathcal{A}}(abb) = \frac{3}{8}$

In general:

$$\mathbf{Pr}_{\mathcal{A}}(wa) = \frac{1}{2}\mathbf{Pr}_{\mathcal{A}}(w) \quad \text{and} \quad \mathbf{Pr}_{\mathcal{A}}(wb) = \frac{1}{2}(1 + \mathbf{Pr}_{\mathcal{A}}(w))$$

Thus giving an explicit acceptance probability:

$$\mathbf{Pr}_{\mathcal{A}}(a_1 \dots a_n) = \sum_{i=1}^{n} 2^{i-n-1} \cdot \mathbf{1}_{a_i = b}$$

Which word maximizes the acceptance probability?

1. Probabilistic automata
   - Presentation
   - Stochastic languages
   - Decision problems

2. Partially observable MDP
   - Presentation
   - POMDP analysis
   - Application to control for fault diagnosis

3. Conclusion

Languages defined by PA

Selection of good strategies

Stochastic languages

For $\mathcal{A}$ a PA, $\theta \in [0,1]$ a *threshold* and $\bowtie \in \{<, \leq, >, \geq, =, \neq\}$ an operator, the *stochastic language* $L_{\bowtie\theta}(\mathcal{A})$ is defined by

$$L_{\bowtie\theta}(\mathcal{A}) = \{w \in A^* \mid \mathbf{Pr}_{\mathcal{A}}(w) \bowtie \theta\}$$

We further define subclasses of stochastic languages.

Rational languages

- A PA is *rational* if its probabilities are in $\mathbb{Q}$.
- A stochastic language is *rational* if it is specified by a rational PA and a rational threshold.

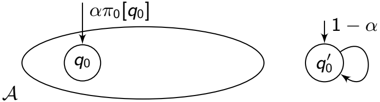
Removing syntactic sugar
Getting rid of useless thresholds and operators

Unique threshold

For every PA $\mathcal{A}$, threshold θ and comparison operator $\bowtie$, there exists $\mathcal{A}'$ s.t.

$$L_{\bowtie\frac{1}{2}}(\mathcal{A}') = L_{\bowtie\theta}(\mathcal{A})$$

Proof



- Case $\theta > \frac{1}{2}$
  set $q_0' \notin F$ and $\alpha = \frac{1}{2\theta}$;
- Case $\theta < \frac{1}{2}$
  set $q_0' \in F$ and $\alpha = \frac{1}{2(1-\theta)}$.

Removing syntactic sugar

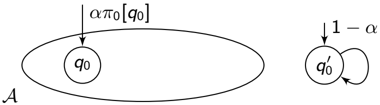
Getting rid of useless thresholds and operators

Unique threshold

For every PA $\mathcal{A}$, threshold θ and comparison operator $\bowtie$, there exists $\mathcal{A}'$ s.t.

$$L_{\bowtie\frac{1}{2}}(\mathcal{A}') = L_{\bowtie\theta}(\mathcal{A})$$

Proof



- Case $\theta > \frac{1}{2}$
  set $q_0' \notin F$ and $\alpha = \frac{1}{2\theta}$;
- Case $\theta < \frac{1}{2}$
  set $q_0' \in F$ and $\alpha = \frac{1}{2(1-\theta)}$.

Restricting operators

Comparison operators $\geq$ and $>$ suffice.

Proof idea

- $\leq$ and $<$ removed by complementation of final states;
- $\mathcal{A}'$ runs two copies of $\mathcal{A}$ in parallel, and $F' = F \times (Q \setminus F)$ then:
  - $\mathbf{Pr}_{\mathcal{A}'}(w) = \mathbf{Pr}_{\mathcal{A}}(w)(1 - \mathbf{Pr}_{\mathcal{A}}(w))$
  - $L_{\geq\frac{1}{4}}(\mathcal{A}') = L_{=\frac{1}{2}}(\mathcal{A})$

Regular vs stochastic languages

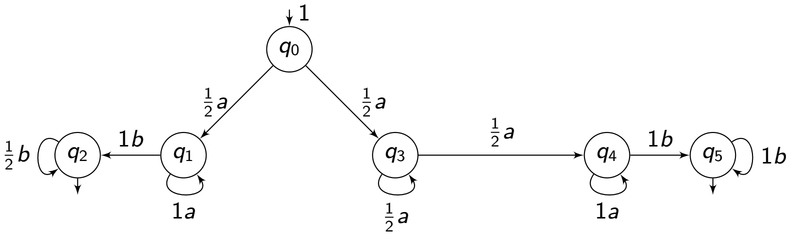
Regular vs stochastic

Regular languages are rational stochastic.

Proof

A DFA is a PA with transition probabilities in $\{0, 1\}$.

A counting PA



absorbing sink state is omitted

Accepted words are of the form $w = a^m b^n$ with $m > 0$, $n > 0$.
Accepting runs on w are:

- the run $q_0 q_1^m q_2^n$, with probability $\frac{1}{2^n}$;
- the family of runs $q_0 q_3^r q_4^s q_5^n$ with $r, s > 0$ and $r + s = m$, with total probability $\frac{1}{2} - \frac{1}{2^m}$.

Altogether $\mathbf{Pr}_{\mathcal{A}}(w) = \frac{1}{2} + \frac{1}{2^n} - \frac{1}{2^m}$.

$$\mathcal{L}_{=\frac{1}{2}}(\mathcal{A}) = \{a^n b^n \mid n > 0\}$$

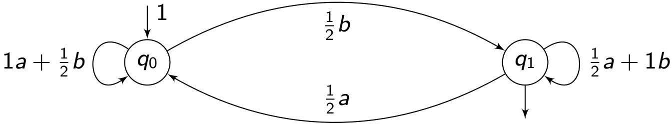
Stochastic vs context-free languages

Stochastic vs context-free languages

Context-free languages and stochastic languages are incomparable.

- $L = \{a^{n_1} b a^{n_2} b \dots a^{n_k} b a^* \mid \exists i > 1 \ n_i = n_1\}$
  is a context-free language that is not stochastic.
- $L = \{a^n b^n c^n \mid n > 0\}$
  is a rational stochastic language that is not contex-free.
  - $\{a^n b^n \mid n > 0\} = \{a^n b^n c^+ \mid n > 0\} \cap \{a^+ b^n c^n \mid n > 0\}$
  - family $\{\mathcal{L}_{=\theta}(\mathcal{A}) \mid \mathcal{A} \text{ PA }\}$ is closed under intersection

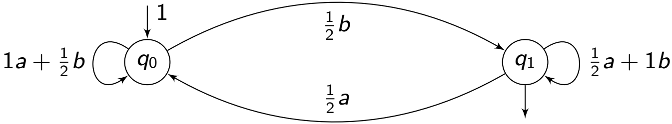
Stochastic vs contextual languages



For $w = w_1 \ldots w_n$, $\mathbf{Pr}_{\mathcal{A}}(w) = 0.\varphi(w_1)\ldots\varphi(w_n)$ with $\varphi(a) = 0$ and $\varphi(b) = 1$.

$$\mathcal{L}_{>\theta}(\mathcal{A}) = \{r \in [0,1] \mid \mathrm{bin}(r) > \theta\}$$

Stochastic vs contextual languages



For $w = w_1 \dots w_n$, $\mathbf{Pr}_{\mathcal{A}}(w) = 0.\varphi(w_1)\dots\varphi(w_n)$ with $\varphi(a) = 0$ and $\varphi(b) = 1$.

$$\mathcal{L}_{>\theta}(\mathcal{A}) = \{r \in [0,1] \mid \mathrm{bin}(r) > \theta\}$$

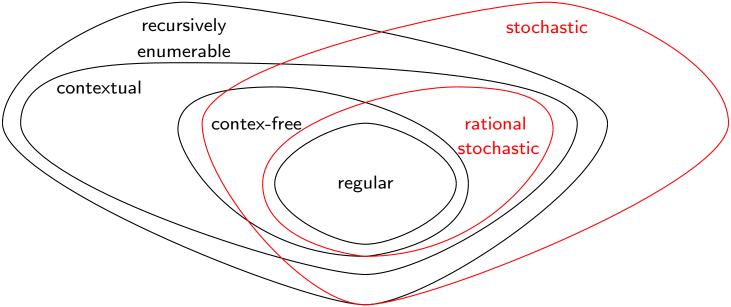
$$\theta < \theta' \Rightarrow \mathcal{L}_{>\theta'}(\mathcal{A}) \subsetneq \mathcal{L}_{>\theta}(\mathcal{A})$$

Cardinality of stochastic languages

There are uncountably many stochastic languages.

Consequence: "Most" stochastic languages are not recursively enumerable.
Not valid for rational stochastic languages!

Comparison with Chomsky's hierarchy



1. Probabilistic automata
   - Presentation
   - Stochastic languages
   - Decision problems

2. Partially observable MDP
   - Presentation
   - POMDP analysis
   - Application to control for fault diagnosis

3. Conclusion

Two decision problems

Quantitative language equivalence

Input: $\mathcal{A}$ and $\mathcal{A}'$ PA

Output: yes iff $\forall w \in A^* \ \mathbf{Pr}_{\mathcal{A}}(w) = \mathbf{Pr}_{\mathcal{A}'}(w)$

Boolean language equivalence

Input: $\mathcal{A}$ and $\mathcal{A}'$ PA, θ, θ' thresholds, $\bowtie, \bowtie'$ comparison operators

Output: yes iff $L_{\bowtie\theta}(\mathcal{A}) = L_{\bowtie'\theta'}(\mathcal{A}')$

Two decision problems

Quantitative language equivalence

Input: $\mathcal{A}$ and $\mathcal{A}'$ PA

Output: yes iff $\forall w \in A^*$ $\mathbf{Pr}_{\mathcal{A}}(w) = \mathbf{Pr}_{\mathcal{A}'}(w)$

Boolean language equivalence

Input: $\mathcal{A}$ and $\mathcal{A}'$ PA, θ, θ' thresholds, $\bowtie, \bowtie'$ comparison operators

Output: yes iff $L_{\bowtie\theta}(\mathcal{A}) = L_{\bowtie'\theta'}(\mathcal{A}')$

Note: for deterministic automata

- the two problems coincide
- decidable in PTIME by a product construction
- a witness of non-equivalence has size at most $|Q||Q'|$.

Quantitative language equivalence

Quantitative language equivalence

Quantitative language equivalence is decidable in PTIME.

Quantitative language equivalence

Quantitative language equivalence

Quantitative language equivalence is decidable in PTIME.

Algorithm idea

Principle enumerate words of increasing length to find a counterexample

Data structures

- a stack to store words w such that all aw need be checked
- a set Gen of independent vectors of $\mathbb{R}^{Q \cup Q'}$

Quantitative language equivalence

Quantitative language equivalence

Quantitative language equivalence is decidable in PTIME.

Algorithm idea

Principle enumerate words of increasing length to find a counterexample

Data structures

- a stack to store words w such that all aw need be checked
- a set *Gen* of independent vectors of $\mathbb{R}^{Q \cup Q'}$

Iteration if w is not a counterexample
and if $v = \mathbf{P}_w \mathbf{1}_F - \mathbf{P}'_w \mathbf{1}_{F'}$ is not generated by *Gen*
then add w to the stack and add $v - \mathrm{Proj}_{Gen}(v)$ to *Gen*

Quantitative language equivalence

Quantitative language equivalence

Quantitative language equivalence is decidable in PTIME.

Algorithm idea

Principle enumerate words of increasing length to find a counterexample

Data structures

- a stack to store words w such that all aw need be checked
- a set Gen of independent vectors of $\mathbb{R}^{Q \cup Q'}$

Iteration if w is not a counterexample
and if $v = \mathbf{P}_w \mathbf{1}_F - \mathbf{P}'_w \mathbf{1}_{F'}$ is not generated by Gen
then add w to the stack and add $v - \mathrm{Proj}_{Gen}(v)$ to Gen

Correctness is non trivial
$|Q| + |Q'|$ bounds the number of iterations and the size of a witness.

Boolean language equivalence

Boolean language equivalence

The problem, given a PA $\mathcal{A}$ of telling whether $L_{=\frac{1}{2}}(\mathcal{A}) = \{\varepsilon\}$ is undecidable.

Boolean language equivalence

Boolean language equivalence

The problem, given a PA $\mathcal{A}$ of telling whether $L_{=\frac{1}{2}}(\mathcal{A}) = \{\varepsilon\}$ is undecidable.

Proof sketch: reduction from PCP

- PCP instance: morphisms $\varphi_1 : A \to \{0,1\}^+$ and $\varphi_2 : A \to \{0,1\}^+$
- $v \in \{0,1\}^+$ defines a value $\mathrm{val}(v) = \sum_{i=1}^{n} \frac{v_i}{2^{n-i}}$

Boolean language equivalence

Boolean language equivalence

The problem, given a PA $\mathcal{A}$ of telling whether $L_{=\frac{1}{2}}(\mathcal{A}) = \{\varepsilon\}$ is undecidable.

Proof sketch: reduction from PCP

- PCP instance: morphisms $\varphi_1 : A \to \{0,1\}^+$ and $\varphi_2 : A \to \{0,1\}^+$
- $v \in \{0,1\}^+$ defines a value $\mathrm{val}(v) = \sum_{i=1}^{n} \frac{v_i}{2^{n-i}}$
- Define $\mathcal{A}_1$ such that $\mathbf{Pr}_{\mathcal{A}_1}(w) = \mathrm{val}(\varphi_1(w))$ and $\mathcal{A}_2$ such that $\mathbf{Pr}_{\mathcal{A}_2}(w) = 1 - \mathrm{val}(\varphi_2(w))$
- PA $\mathcal{A}$ starts in $\mathcal{A}_1$ or $\mathcal{A}_2$ with equal probability, thus

$$\mathbf{Pr}_{\mathcal{A}}(w) = \frac{1}{2}\Big(\mathrm{val}(\varphi_1(w)) + (1 - \mathrm{val}(\varphi_2(w)))\Big)$$

Boolean language equivalence

Boolean language equivalence

The problem, given a PA $\mathcal{A}$ of telling whether $L_{=\frac{1}{2}}(\mathcal{A}) = \{\varepsilon\}$ is undecidable.

Proof sketch: reduction from PCP

- PCP instance: morphisms $\varphi_1 : A \to \{0,1\}^+$ and $\varphi_2 : A \to \{0,1\}^+$
- $v \in \{0,1\}^+$ defines a value $\mathrm{val}(v) = \sum_{i=1}^{n} \frac{v_i}{2^{n-i}}$
- Define $\mathcal{A}_1$ such that $\mathbf{Pr}_{\mathcal{A}_1}(w) = \mathrm{val}(\varphi_1(w))$ and $\mathcal{A}_2$ such that $\mathbf{Pr}_{\mathcal{A}_2}(w) = 1 - \mathrm{val}(\varphi_2(w))$
- PA $\mathcal{A}$ starts in $\mathcal{A}_1$ or $\mathcal{A}_2$ with equal probability, thus

$$\mathbf{Pr}_{\mathcal{A}}(w) = \frac{1}{2}\Big(\mathrm{val}(\varphi_1(w)) + (1 - \mathrm{val}(\varphi_2(w)))\Big)$$

$$\mathbf{Pr}_{\mathcal{A}}(w) = \frac{1}{2} \quad \Longleftrightarrow \quad \varphi_1(w) = \varphi_2(w)$$

Qualitative problems for PA

Non-emptiness of (almost-)sure language
Input: $\mathcal{A}$ PA
Output: yes iff $\exists w,\ \mathbf{Pr}_{\mathcal{A}}(w) = 1$

Qualitative problems for PA

Non-emptiness of (almost-)sure language
Input: $\mathcal{A}$ PA
Output: yes iff $\exists w,\ \mathbf{Pr}_{\mathcal{A}}(w) = 1$

- almost-sure reachability for PA

Non-emptiness of almost-sure language is PSPACE-complete.

- decidable in PSPACE
  - complement final states $F' = Q \setminus F$
  - consider $\mathcal{A}'$ as an NFA
  - $L(\mathcal{A}') \neq A^*$ iff $L_{=1}(\mathcal{A}) \neq \emptyset$

Qualitative problems for PA

Non-emptiness of (almost-)sure language

Input: $\mathcal{A}$ PA
Output: yes iff $\exists w, \ \mathbf{Pr}_{\mathcal{A}}(w) = 1$

- almost-sure reachability for PA

Non-emptiness of almost-sure language is PSPACE-complete.

- decidable in PSPACE
  - complement final states $F' = Q \setminus F$
  - consider $\mathcal{A}'$ as an NFA
  - $L(\mathcal{A}') \neq A^*$ iff $L_{=1}(\mathcal{A}) \neq \emptyset$

Non-emptiness of limit-sure language

Input: $\mathcal{A}$ PA
Output: yes iff $\exists (w_n)_{n \in \mathbb{N}}, \ \lim_{n \to \infty} \mathbf{Pr}_{\mathcal{A}}(w_n) = 1$

Qualitative problems for PA

Non-emptiness of (almost-)sure language

Input: $\mathcal{A}$ PA

Output: yes iff $\exists w,\ \mathbf{Pr}_{\mathcal{A}}(w) = 1$

- almost-sure reachability for PA

> Non-emptiness of almost-sure language is PSPACE-complete.

- decidable in PSPACE
  - complement final states $F' = Q \setminus F$
  - consider $\mathcal{A}'$ as an NFA
  - $L(\mathcal{A}') \neq A^*$ iff $L_{=1}(\mathcal{A}) \neq \emptyset$

Non-emptiness of limit-sure language

Input: $\mathcal{A}$ PA

Output: yes iff $\exists (w_n)_{n \in \mathbb{N}},\ \lim_{n \to \infty} \mathbf{Pr}_{\mathcal{A}}(w_n) = 1$

- limit-sure reachability for PA
- *value 1 problem*

> Non-emptiness of limit-sure language is undecidable.

1. Probabilistic automata
   - Presentation
   - Stochastic languages
   - Decision problems

2. Partially observable MDP
   - Presentation
   - POMDP analysis
   - Application to control for fault diagnosis

3. Conclusion

1. Probabilistic automata
   - Presentation
   - Stochastic languages
   - Decision problems

2. Partially observable MDP
   - Presentation
   - POMDP analysis
   - Application to control for fault diagnosis

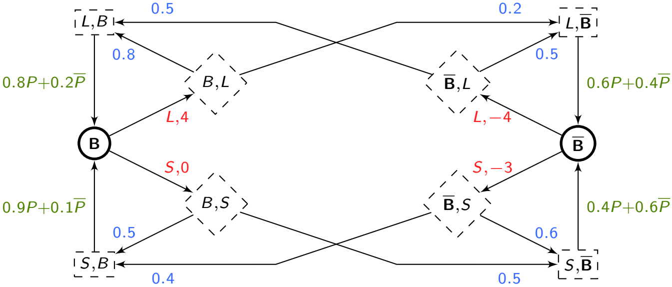
3. Conclusion

A first POMDP example

A company sells a product, either luxury (L) or standard (S).
Consumers may be sensitive to brands (**B**) or not ($\overline{\mathbf{B}}$)
but the company does not know this information...
... and only knows whether the product is purchased (P) or not ($\overline{P}$).

A first POMDP example

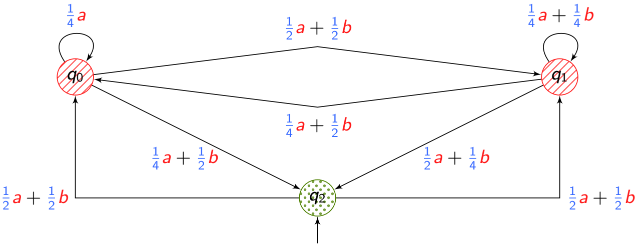
A company sells a product, either luxury (L) or standard (S).
Consumers may be sensitive to brands (**B**) or not ($\overline{\mathbf{B}}$)
but the company does not know this information...
... and only knows whether the product is purchased (P) or not ($\overline{P}$).



States: **B**, $\overline{\mathbf{B}}$; Actions: L, S; Observations: P, $\overline{P}$;

- probabilities: $p(\mathbf{B}|\mathbf{B}, L) = 0.8$;
- rewards: $\text{rew}(\mathbf{B}, L) = 4$;
- observations: $o(P|L, \mathbf{B}) = 0.8$

A second POMDP example



States : $\{q_0, q_1, q_2\}$; Actions : $\{a, b\}$; Observations : $\{$, $\}$

- probabilities: $p(q_1|q_0, a) = \frac{1}{2}$
- rewards: 0 everywhere
- observations: $o(q_0) = o(q_1) =$

POMDP

Deterministic observation POMDP

A POMDP $\mathcal{M} = (S, \Omega, A, o, p, \mathsf{rew}, \mathsf{rew_f})$ is defined by:

- S a finite set of states;
- Ω a finite set of observations;
- A a finite set of actions;

POMDP

Deterministic observation POMDP

A POMDP $\mathcal{M} = (S, \Omega, A, o, p, \mathrm{rew}, \mathrm{rew_f})$ is defined by:

- S a finite set of states;
- Ω a finite set of observations;
- A a finite set of actions;
- $o : S \to \Omega$ the observation function; $o(s) \in \Omega$ is the observation associated with state s;
- $p : S \times A \to \mathrm{Dist}(S)$ the transition function; $p(s'|s, a)$ is the probability that the next state be s' when action a occurs from s;
- $\mathrm{rew} : S \times A \to \mathbb{Q}$ the reward function; $\mathrm{rew}(s, a)$ is the reward associated with action a from state s.
- $\mathrm{rew_f} : S \to \mathbb{Q}$ the final reward function; $\mathrm{rew_f}(s)$ is the reward associated when ending in state s.

Strategies

To obtain a stochastic process, a *strategy* rules out non-determinism.

Strategies

A *strategy* is a function $\nu : (A\Omega)^* \to \mathrm{Dist}(A)$ mapping each *history* $\rho \in (A\Omega)^*$ with a distribution over actions; $\nu(\rho, a)$ is the probability that a is chosen given history ρ.

Strategies

To obtain a stochastic process, a *strategy* rules out non-determinism.

Strategies

A *strategy* is a function $\nu : (A\Omega)^* \to \text{Dist}(A)$ mapping each *history* $\rho \in (A\Omega)^*$ with a distribution over actions; $\nu(\rho, a)$ is the probability that a is chosen given history ρ.

Induced Markov chain

Let $\mathcal{M}$ be a POMDP, ν a strategy and $\pi \in \text{Dist}(S)$ an initial distribution. The Markov chain $\mathcal{M}^{\pi}_{\nu}$ induced by $\mathcal{M}$, ν et π is defined by:

- $(A\Omega)^* \times S$ its (infinite) state space;
- π_0 the initial distribution such that $\pi_0(\varepsilon, s) = \pi(s)$ and π_0 is null for other states;
- $\mathbf{P}$ the transition matrix such that:
  $\mathbf{P}[(\rho, s), (\rho ao(s'), s')] = \nu(\rho, a)p(s'|s, a)$, and $\mathbf{P}$ is zero elsewhere.

POMDP subclasses

Two very particular cases:

- $\Omega = S$: the agent knows the state of the system; (full observation) Markov decision process.
- $|\Omega| = 1$: observation is useless; *blind* POMDP.

POMDP subclasses

Two very particular cases:

- $\Omega = S$: the agent knows the state of the system; (full observation) Markov decision process.
- $|\Omega| = 1$: observation is useless; *blind* POMDP.

PA vs POMDP

Probabilistic automata form a subclass of POMDP.

word in probabilistic automaton $\Longleftrightarrow$ pure strategy in blind POMDP

Consequence: All hardness results lift from PA to POMDP.

1. Probabilistic automata
   - Presentation
   - Stochastic languages
   - Decision problems

2. Partially observable MDP
   - Presentation
   - POMDP analysis
   - Application to control for fault diagnosis

3. Conclusion

Finite-horizon analysis

Expected total payoff

The *expected total payoff* at time t, under strategy ν is

$$u_t^\nu = \sum_{i=0}^{t-1} \mathbb{E}^\nu(\mathsf{rew}(X_i, Y_i)) + \mathbb{E}^\nu(\mathsf{rew_f}(X_t))$$

where X_i (resp. Y_i) is the random variable of state (action) at step i.
The *optimal expected total payoff* at time t is

$$u_t^\star = \sup_\nu u_t^\nu$$

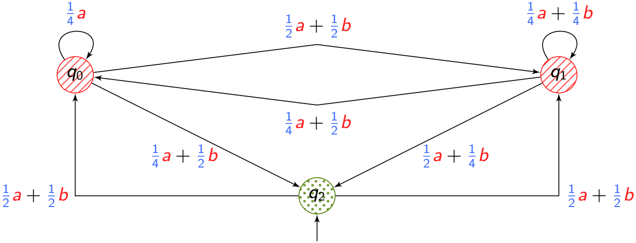
Finite-horizon analysis

One can compute a set of indices Z_t, a family of vectors $\{\mathbf{r}_z\}_{z\in Z_t}$, a family of polyedra $\{\mathbf{D}_z\}_{z\in Z_t}$ such that

- $\bigcup_{z\in Z_t} \mathbf{D}_z$ is the set of distributions over states
- for every initial distribution π, $\pi \in \mathbf{D}_z \Rightarrow u_t^\star(\pi) = \pi\mathbf{r}_z$

Finite-horizon analysis on an example

$\mathrm{rew}_f(q_2) = 1$ and all other rewards are 0



Objective: for $t = 1$, determine Z, $(\mathbf{D}_z)_{z \in Z}$ and $(\mathbf{r}_z)_{z \in Z}$ such that

$$\pi \in \mathbf{D}_z \Rightarrow u_t^\star(\pi) = \pi \mathbf{r}_z$$

$Z = \{a, b\}$
$\mathbf{D}_a = \{(x_0, x_1, x_2) \mid x_0 + x_1 + x_2 = 1 \wedge x_0 \leq x_1\}$ $\quad \mathbf{r}_a = (\frac{1}{4}, \frac{1}{2}, 0)$
$\mathbf{D}_b = \{(x_0, x_1, x_2) \mid x_0 + x_1 + x_2 = 1 \wedge x_0 \geq x_1\}$ $\quad \mathbf{r}_b = (\frac{1}{2}, \frac{1}{4}, 0)$

Infinite-horizon problems

Objectives

Reachability F visited at least once:

$$\Diamond F = \{q_0 q_1 q_2 \cdots \in S^\omega \mid \exists n,\ q_n \in F\}$$

Safety always stay in F:

$$\Box F = \{q_0 q_1 q_2 \cdots \in S^\omega \mid \forall n,\ q_n \in F\}$$

Büchi F visited an infinite number of times:

$$\Box\Diamond F = \{q_0 q_1 q_2 \cdots \in S^\omega \mid \forall m\ \exists n \geq m,\ q_n \in F\}$$

Goal: For φ an objective, evaluate $\sup_\nu \mathbb{P}^\nu(\mathcal{M} \models \varphi)$.

Infinite-horizon problems

Objectives

Reachability F visited at least once:

$$\Diamond F = \{q_0 q_1 q_2 \cdots \in S^\omega \mid \exists n,\ q_n \in F\}$$

Safety always stay in F:

$$\Box F = \{q_0 q_1 q_2 \cdots \in S^\omega \mid \forall n,\ q_n \in F\}$$

Büchi F visited an infinite number of times:

$$\Box\Diamond F = \{q_0 q_1 q_2 \cdots \in S^\omega \mid \forall m\ \exists n \geq m,\ q_n \in F\}$$

Goal: For φ an objective, evaluate $\sup_\nu \mathbb{P}^\nu(\mathcal{M} \models \varphi)$.

Deterministic strategies are sufficient!

Let $\mathcal{M}$ be a POMDP, and $\varphi \subseteq S^\omega$ a Borelian objective. For every strategy ν, there exists a deterministic strategy ν' such that

$$\mathbb{P}^\nu(\mathcal{M} \models \varphi) \leq \mathbb{P}^{\nu'}(\mathcal{M} \models \varphi).$$

Undecidability of infinite-horizon quantitative analysis

Undecidability of quantitative reachability

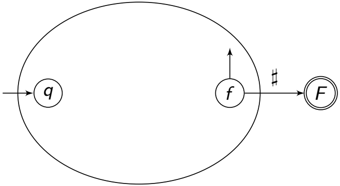
The problem of the existence of a strategy ensuring the reachability objective $\Diamond F$ with probability at least p is undecidable for POMDP.

Undecidability of infinite-horizon quantitative analysis

Undecidability of quantitative reachability

The problem of the existence of a strategy ensuring the reachability objective $\Diamond F$ with probability at least p is undecidable for POMDP.

Reduction from the emptiness problem for PA.
Only subtlety: synchronize paths!



deterministic strategies in $\mathcal{M}$: $\nu_w = w\sharp$, where w word for the PA $\mathcal{A}$

$$\mathbb{P}^{\nu_w}(\mathcal{M} \models \Diamond F) = \mathbb{P}_{\mathcal{A}}(w)$$

Undecidability of qualitative infinite-horizon analysis

Undecidability of positive repeated reachability

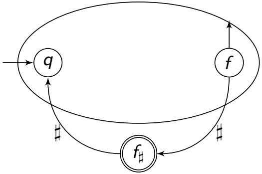
The problem of the existence of a strategy ensuring the repeated reachability objective $\Box\Diamond F$ with probability > 0 is undecidable for POMDP.

Undecidability of qualitative infinite-horizon analysis

Undecidability of positive repeated reachability

The problem of the existence of a strategy ensuring the repeated reachability objective $\Box\Diamond F$ with probability > 0 is undecidable for POMDP.

Reduction from the value 1 problem for PA.



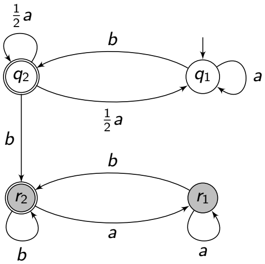
deterministic strategies in $\mathcal{M}$: $\nu_{\mathbf{w}} = w_1\sharp\sharp w_2\sharp\sharp w_3\cdots$, where w_i words for PA $\mathcal{A}$

$$\mathbb{P}^{\nu_{\mathbf{w}}}(\mathcal{M} \models \Box\Diamond f_\sharp) > 0 \Longleftrightarrow \prod_i \mathbb{P}_{\mathcal{A}}(w_i) > 0$$

$$\mathrm{val}(\mathcal{A}) = 1 \quad\Longleftrightarrow\quad \exists (w_i)_{i\in\mathbb{N}} \prod_i \mathbb{P}_{\mathcal{A}}(w_i) > 0 \quad\Longleftrightarrow\quad \exists \nu_{\mathbf{w}}\, \mathbb{P}^{\nu_{\mathbf{w}}}(\mathcal{M} \models \Box\Diamond f_\sharp) > 0$$

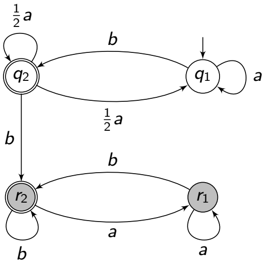
Combination of infinite-horizon objectives

Infinite memory is needed for combined objectives!
Goal: $\Box\Diamond\{q_2, r_2\}$ almost surely and $\Box\{q_1, q_2\}$ with positive probability.



Combination of infinite-horizon objectives

Infinite memory is needed for combined objectives!
Goal: $\Box\Diamond\{q_2, r_2\}$ almost surely and $\Box\{q_1, q_2\}$ with positive probability.



Undecidability of combined qualitative objectives

The problem of the existence of a strategy ensuring

- a safety objective $\Box G$ with probability > 0, and
- a Büchi objective $\Box\Diamond F$ with probability $= 1$

is undecidable for POMDP.

Decidability of qualitative infinite-horizon analysis

Decidability of positive reachability

The problem of the existence of a strategy ensuring a reachability objective $\Diamond F$ with probability > 0 is NLOGSPACE-complete for POMDP.

Decidability of qualitative infinite-horizon analysis

Decidability of positive reachability

The problem of the existence of a strategy ensuring a reachability objective $\Diamond F$ with probability > 0 is NLOGSPACE-complete for POMDP.

- Equivalent to reachability in graphs.
- Purely random strategy works: uniform randomization on all actions at each step.

Decidability of qualitative infinite-horizon analysis (2)

Decidability of almost-sure safety

The problem of the existence of a strategy ensuring a safety objective $\Box G$ with probability $= 1$ is EXPTIME-complete for POMDP.

Decidability of qualitative infinite-horizon analysis (2)

Decidability of almost-sure safety

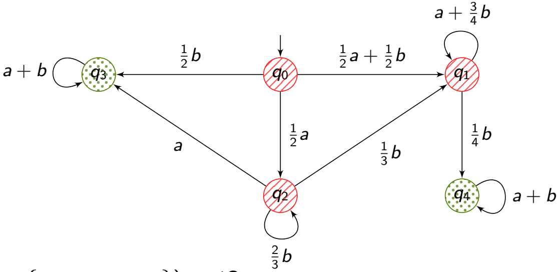
The problem of the existence of a strategy ensuring a safety objective $\Box G$ with probability $= 1$ is EXPTIME-complete for POMDP.

Beliefs

The *belief* of the agent is the set of possible states, given the sequence of observations so far.

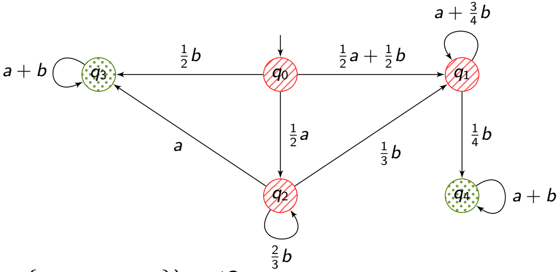
Necessary and sufficient condition: agent maintains its belief included in G.
One builds the *belief game*.

Belief game on an example

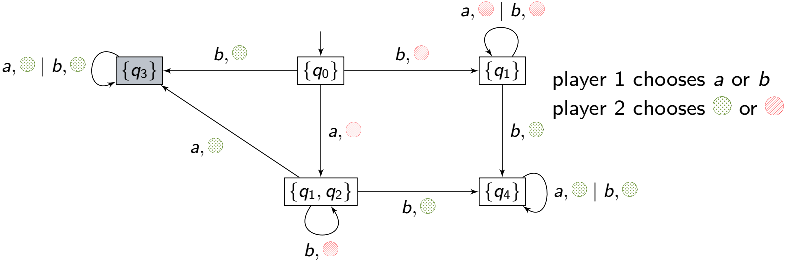


$$\exists \nu \ \mathbb{P}^{\nu}(\mathcal{M} \models \Box\{q_0, q_1, q_2, q_4\}) = 1?$$

Belief game on an example



$\exists \nu \ \mathbb{P}^{\nu}(\mathcal{M} \models \Box\{q_0, q_1, q_2, q_4\}) = 1?$



Decidability of qualitative infinite-horizon analysis (3)

Decidability positive safety

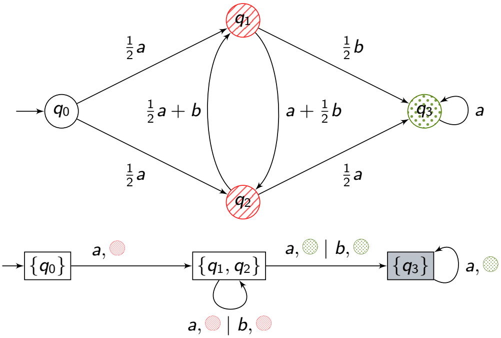
The problem of the existence of a strategy ensuring a safety objective $\Box G$ with positive probability is EXPTIME-complete for POMDP.

Decidability of qualitative infinite-horizon analysis (3)

Decidability positive safety

The problem of the existence of a strategy ensuring a safety objective $\Box G$ with positive probability is EXPTIME-complete for POMDP.

Positional strategies on belief game are not enough...



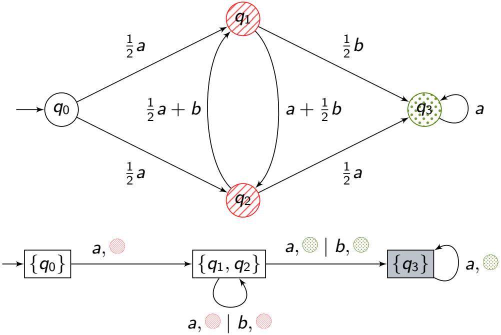
Yet, choosing a, then bet the system lies in q_1, and alternate a and b for ever, guarantees a probability $\frac{1}{2}$ for $\Box\{q_0, q_1, q_2\}$.

Decidability of qualitative infinite-horizon analysis (3)

Decidability positive safety

The problem of the existence of a strategy ensuring a safety objective $\Box G$ with positive probability is EXPTIME-complete for POMDP.

Positional strategies on belief game are not enough...



... but almost! It is necessary and sufficient to reach a belief $C \subseteq S$ such that there exists a state $s \in C$ and a strategy ensuring to surely stay in G from s.

Decidability of infinite-horizon qualitative analysis

Decidability almost sure (repeated) reachability

The problem of the existence of a strategy ensuring a reachability objective $\Diamond F$ almost surely is EXPTIME-complete for POMDP.

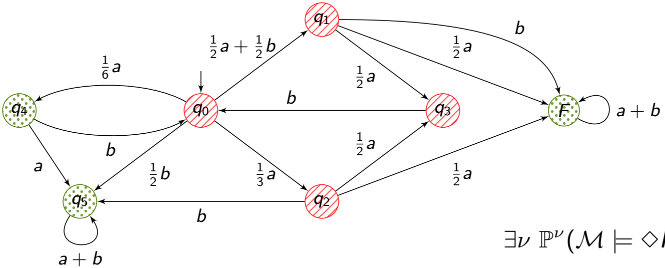
Idea: one needs to reach a belief included in F; every observation deviating from this path must still lead to a winning belief, to be able to try again to reach F.

Win is the biggest set of beliefs such that:

$$\mathrm{Win} = \{C \mid \exists C \xrightarrow{a_1,o_1} C_1 \cdots \xrightarrow{a_n,o_n} C_n \subseteq F$$

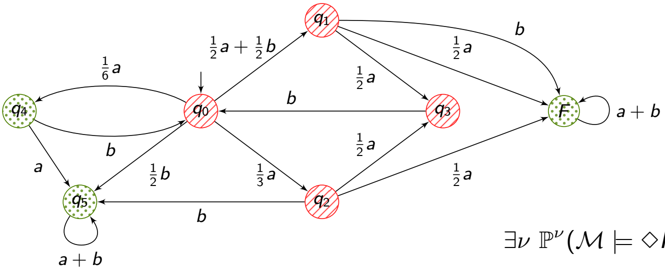
$$\text{and } \forall o'_k C \xrightarrow{a_1,o_1} C_1 \cdots \xrightarrow{a_k,o'_k} C'_k \in \mathrm{Win}\}$$

Decision algorithm on an example

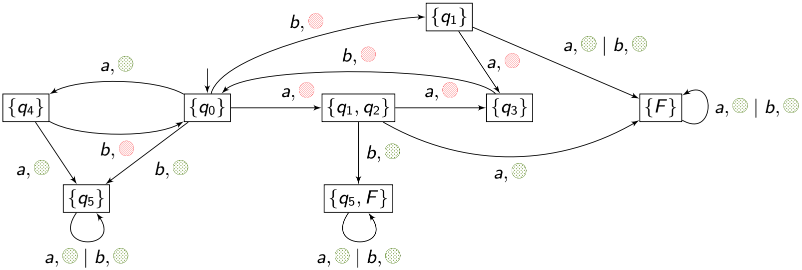


$$\exists \nu \; \mathbb{P}^{\nu}(\mathcal{M} \models \Diamond F) = 1?$$

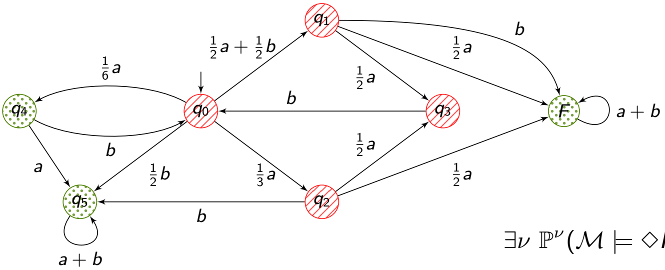
Decision algorithm on an example



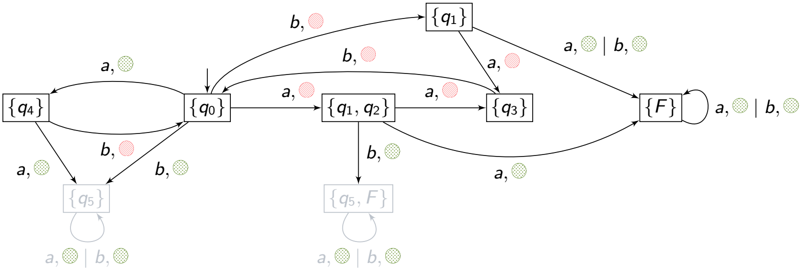
$$\exists \nu \ \mathbb{P}^{\nu}(\mathcal{M} \models \Diamond F) = 1?$$



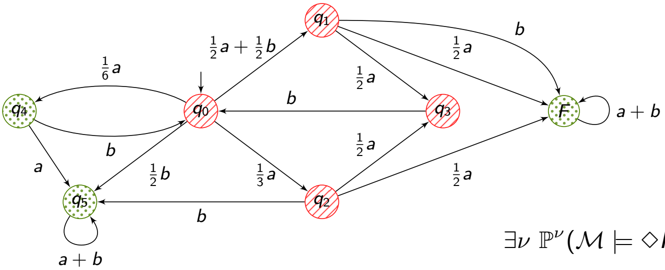
Decision algorithm on an example



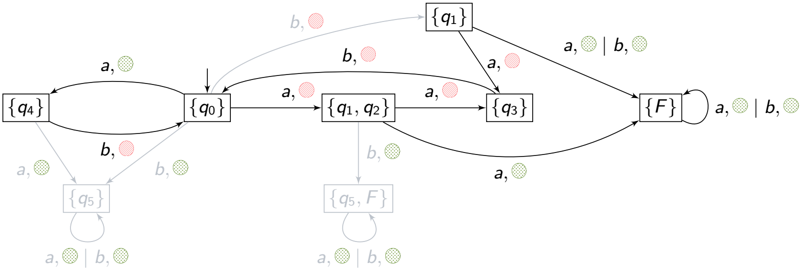
$\exists \nu \; \mathbb{P}^{\nu}(\mathcal{M} \models \Diamond F) = 1?$



Decision algorithm on an example



$$\exists \nu \ \mathbb{P}^{\nu}(\mathcal{M} \models \Diamond F) = 1?$$



1. Probabilistic automata
   - Presentation
   - Stochastic languages
   - Decision problems

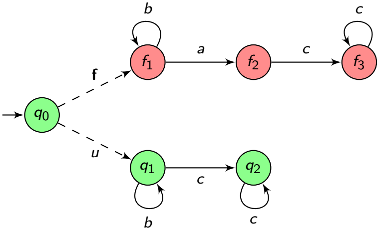
2. Partially observable MDP
   - Presentation
   - POMDP analysis
   - Application to control for fault diagnosis

3. Conclusion

Fault diagnosis

Goal: determine whether a fault **f** occurred, based on the observed events.

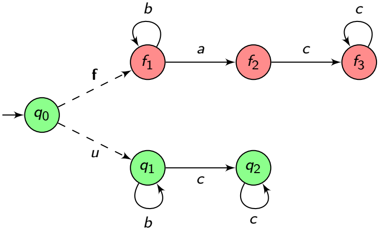
$\Sigma_o = \{a, b, c\}$ observable ; $\Sigma_u = \{\mathbf{f}, u\}$ non-observable



Fault diagnosis

Goal: determine whether a fault **f** occurred, based on the observed events.

$\Sigma_o = \{a, b, c\}$ observable ; $\Sigma_u = \{\mathbf{f}, u\}$ non-observable

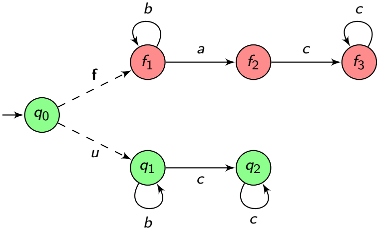


c^+	✓	surely correct
ac^+	✗	surely faulty
b^+	?	ambiguous

Fault diagnosis

Goal: determine whether a fault **f** occurred, based on the observed events.

$\Sigma_o = \{a, b, c\}$ observable ; $\Sigma_u = \{\mathbf{f}, u\}$ non-observable



c^+	✓	surely correct
ac^+	✗	surely faulty
b^+	?	ambiguous

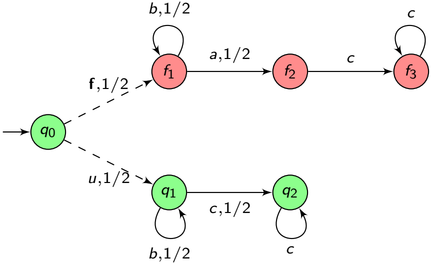
Diagnosability

A system is diagnosable if all its observed sequences are unambiguous.

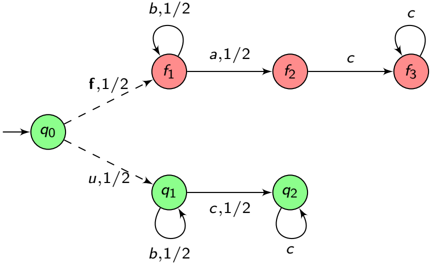
Decidability of diagnosis

The diagnosability problem is NLOGSPACE-complete.

Fault diagnosis for probabilistic systems

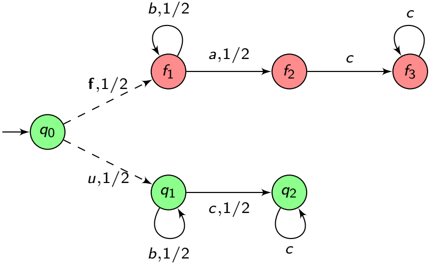


Fault diagnosis for probabilistic systems



b^+ ambiguous but...

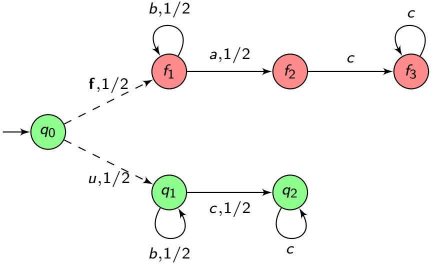
Fault diagnosis for probabilistic systems



b^+ ambiguous but...

$$\lim_{n\to\infty} \mathbb{P}(\mathbf{f}b^n + ub^n) = 0$$

Fault diagnosis for probabilistic systems



b^+ ambiguous but...

$$\lim_{n\to\infty} \mathbb{P}(\mathbf{f} b^n + ub^n) = 0$$

Almost-sure diagnosability

A probabilistic system is diagnosable if the probability of ambiguous observed sequences is null.

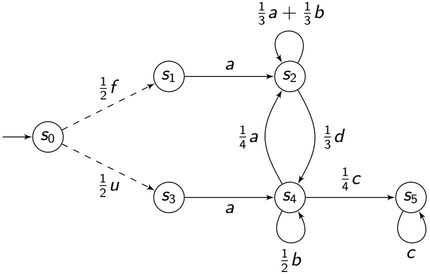
Decidability of almost-sure diagnosis

The almost-sure diagnosis problem is PSPACE-complete.

Active diagnosis

Goal: control the system so that its set of ambiguous sequences has null measure.

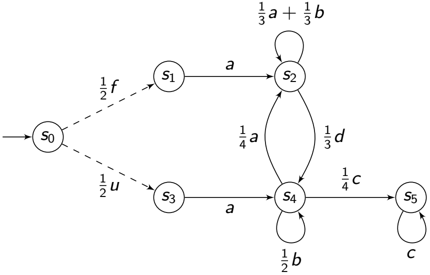
$\Sigma_o = \Sigma_c = \{a, b, c, d\}$ observable and controllable;
$\Sigma_u = \Sigma_e = \{\mathbf{f}, u\}$ unobservable and uncontrollable



Active diagnosis

Goal: control the system so that its set of ambiguous sequences has null measure.

$\Sigma_o = \Sigma_c = \{a, b, c, d\}$ observable and controllable;
$\Sigma_u = \Sigma_e = \{\mathbf{f}, u\}$ unobservable and uncontrollable

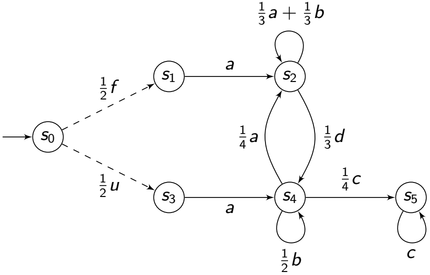


$aadc^\omega$ ambiguous
$\mathbb{P}(faadc^\omega + uaadc^\omega) > 0$

Active diagnosis

Goal: control the system so that its set of ambiguous sequences has null measure.

$\Sigma_o = \Sigma_c = \{a, b, c, d\}$ observable and controllable;
$\Sigma_u = \Sigma_e = \{\mathbf{f}, u\}$ unobservable and uncontrollable



$aadc^\omega$ ambiguous
$\mathbb{P}(faadc^\omega + uaadc^\omega) > 0$

forbid a after the first a

Controller: decides which actions are allowed, based on observations
$\sigma : \Sigma_{\text{obs}}^* \to 2^{\Sigma_c}$

Problem resolution

Decidability of active almost-sure diagnosis

The active diagnosis problem for probabilistic systems is EXPTIME-complete.

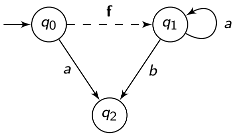
Problem resolution

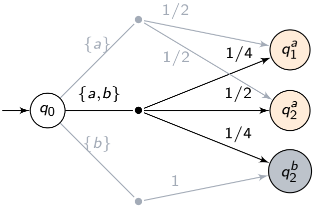
Decidability of active almost-sure diagnosis

The active diagnosis problem for probabilistic systems is EXPTIME-complete.

Idea of EXPTIME-algorithm

- characterize unambiguous sequences by a deterministic Büchi automaton $\mathcal{B}$
- build the product of probabilistic LTS with $\mathcal{B}$: new pLTS
- transform it into POMDP $\mathcal{P}$
  each action is a subset of controllable events
  the observations are observable events





- decide whether there exists a strategy ensuring almost-surely the Büchi condition in $\mathcal{P}$.

Conclusion

POMDP partially observable Markov decision processes

- finite-horizon optimization
- infinite-horizon optimization unfeasible
- qualitative infinite-horizon analysis mostly feasible
- application to active diagnosis of stochastic systems

Conclusion

POMDP partially observable Markov decision processes

- finite-horizon optimization
- infinite-horizon optimization unfeasible
- qualitative infinite-horizon analysis mostly feasible
- application to active diagnosis of stochastic systems

PA probabilistic automata

- particular case of POMDP
- expressiveness
- langagues equivalence, equality, value 1

Partial observation + Probabilities + Control: a challenging combination

References

1. M. O. Rabin. Probabilistic automata. Information and Control, 6(3) :230- 245, 1963.
2. A. Paz. Introduction to probabilistic automata. Academic Press Inc., 1971.
3. O. Madani, S. Hanks, and A. Condon. On the undecidability of probabilistic planning and related stochastic optimization problems. Artificial Intelligence, 147(1-2) :5-34, 2003.
4. H. Gimbert and Y. Oualhadj. Probabilistic automata on finite words: Decidable and undecidable problems. In Proceedings of ICALP'10, volume 6199 of Lecture Notes in Computer Science, pages 527-538. Springer, 2010.
5. C. Baier, N. Bertrand, and M. Größer. Probabilistic ω-automata. J. ACM, 59(1), 2012.

6. K. J. Aström. Optimal control of Markov decision processes with incomplete state estimation. Journal of Mathematical Analysis and Applications, 10 :174-205, 1965.
7. A.R. Cassandra, M.L. Littman, and N.L. Zhang. Incremental pruning: A simple, fast, exact method for partially observable Markov decision processes. In Proceedings of UAI'87, pages 54-61. Morgan Kaufmann, 1997.
8. K. Chatterjee, L. Doyen, H. Gimbert, and T.A. Henzinger. Randomness for free. In Proceedings of MFCS'10, volume 6281 of Lecture Notes in Computer Science, pages 246-257. Springer, 2010.
9. K. Chatterjee, L. Doyen, and T. Henzinger. Qualitative analysis of partially-observable Markov decision processes. In Proceedings of MFCS'10, volume 6281 of Lecture Notes in Computer Science, pages 258-269. Springer, 2010.
10. N. Bertrand, E. Fabre, S. Haar, S. Haddad, and L. Helouet. Active diagnosis for probabilistic systems. In Proceedings of FoSSaCS'14, volume 8412 of Lecture Notes in Computer Science, pages 29-42. Springer, 2014.