The Complexity of Finding Memoryless POMDP Policies

Sebastian Junges

Including work with: Bernd Becker, Nils Jansen, Joost-Pieter Katoen, Guillermo Perez, Tim Quatmann, Ralf Wimmer, Leonore Winterer, Tobias Winkler

Radboud University, RWTH Aachen University, University of Freiburg, University of Antwerp
Outline

Step 1:
- Relate POMDPs + memoryless policies
- to pMCs

Step 2:
- Discuss pMCs
- And relate them to the existential theory

For a formal treatment:


Monty Hall Problem
Monty Hall Problem

Proposal: Change if the car is behind the other door.

Strategy depends on unobservable information
Monty Hall Problem: Humans are bad in reasoning under uncertainty

Should you change now?
Monty-hall POMDP (aggregated)
Monty-hall POMDP (aggregated)
Monty-hall POMDP (aggregated)
Monty-hall POMDP (aggregated)

states indistinguishable

Door selected (player)

Door with goat opened (game master)

switch?

New door selected (player)
Randomisation and memory

POMDP: Reach red state without visiting the dragon.

Start in 1 or 5:
Memoryless policy has to randomise in \{2,4\}

Start in 6 or 7:
no memoryless policy
store whether we have been in 3

same observations:
- \{2,4\}
- \{6,7,8\}
Markov chain (MC): Arriving before 10am

- MCs are Markov Decision Processes with one action in every state
the complexity of finding memoryless POMDP policies

Sebastian Junges

OWLS 2020

Markov decision processes (MDP): Arriving before 10am

• Every state/action maps to a distribution over successors

\[
\begin{array}{c}
\text{asleep} \\
\downarrow 0.9 \\
\text{break-fast} \\
\downarrow 0.1 \\
\text{1} \\
\end{array}
\]

\[
\begin{array}{c}
\text{break-fast} \\
\downarrow 0.2 \\
\text{0.8} \\
\downarrow 0.3 \\
\text{0.7} \\
\end{array}
\]

what is the minimal probability to reach this state?

• MCs are Markov Decision Processes with one action in every state
POMDPs

MDPs with ‘observable colours’

Given any POMDP is there an observation-based policy s.t. the probability reaching $\bullet > \lambda$. 

Distribution over actions

Distribution over actions

The complexity of Finding Memoryless POMDP Policies
Sebastian Junges
OWLS 2020
The complexity of Finding Memoryless POMDP Policies
Sebastian Junges
OWLS 2020

Solving POMDPs is undecidable

Given any POMDP is there an observation-based policy s.t. the probability reaching $\bullet > \lambda$

But cannot be avoided as the world is a POMDP most of the time

AI — A Modern Approach
Partially observable MDPs (POMDPs)

Given any POMDP is there a memoryless strategy s.t. the probability reaching the target state $> \lambda$ in PSPACE, NP-HARD [Vlassis et al, 2012]

POMDP memoryless strategy: colours to distributions over actions
POMDPs with memoryless strategies

Maps observations to distributions over actions

Strategy is uniquely described by values for $x, y_1, y_2$
Applying the memoryless strategy

Induced Markov Chain with unknown probabilities

\[ \begin{align*}
    s_0 & \xrightarrow{a_1, a_2, a_3} s_1 \\
    s_0 & \xrightarrow{a_2} s_2 \\
    s_2 & \xrightarrow{y_2 \cdot 0.5} s_3 \\
    s_0 & \xrightarrow{y_2 \cdot 0.5 + (1 - y_1 - y_2) \cdot 1} s_3 \\
    s_0 & \xrightarrow{x} s_1 \\
    s_0 & \xrightarrow{(1 - y_1 - y_2) \cdot 1} s_3 \\
    s_1 & \xrightarrow{0.5 \cdot x} s_3 \\
    s_1 & \xrightarrow{y_1 \cdot 1} s_2 \\
    s_2 & \xrightarrow{1 - x + 0.5 \cdot x} s_1 \\
    s_3 & \xrightarrow{1 - x} s_2 \\
    s_3 & \xrightarrow{x} s_2
\end{align*} \]
Given any POMDP is there a memoryless strategy s.t. the probability reaching $\bullet > \lambda$.

Given any parametric MC is there a parameter valuation s.t. the probability reaching $\bullet > \lambda$. 

POMDPs and parametric Markov chains
Markov chains

Knuth-Yao die

In every state, flip a coin

Final states: die outcomes
Markov chains

Knuth-Yao die (two fair coins)
Markov chains

Knuth-Yao die (unfair blue coin)
Markov chains

Knuth-Yao die (unfair coins)
Parametric Markov chains (pMCs)

Knuth-Yao die (symbolic probabilities)

Transition probabilities are polynomials

*in this talk: \{x, 1 - x\}
Parametric Markov chains (pMCs)

Knuth-Yao die (instantiation)
\( \mathcal{M}[x \mapsto 0.6, y \mapsto 0.52] \)
Parametric Markov chains (pMCs)

Knuth-Yao die (with specification)

“What is the probability to reach the red state?”
or
“What is the probability above/below some threshold?”
Parametric Markov chains (pMCs)

Knuth-Yao die (trade-offs)
Problem statement: Parameter synthesis

Given: a parametric MC $\mathcal{M}$ with parameters $x$

Find: $\text{val}: x \rightarrow [0,1]$

such that: $\mathcal{M}[\text{val}] \models \varphi$, i.e., a red state is reached with probability at least/at most $\lambda$
Given any POMDP
is there a memoryless strategy s.t.
the probability reaching $\bullet > \lambda$

Given any parametric MC
is there a parameter valuation s.t.
the probability reaching $\bullet > \lambda$
A pMC is simple iff

1. $P(s, s') \in \{x, 1-x \mid x \text{ parameter}\} \cup \mathbb{Q}$ for all $s, s'$.  
2. $\sum_{s'} P(s, s') = 1$ for all $s$.
POMDPs and parametric Markov chains

Given any simple POMDP is there a memoryless policy s.t. the probability reaching $\bullet > \lambda$

Given any simple pMC is there a parameter valuation s.t. the probability reaching $\bullet > \lambda$

Given any POMDP is there a memoryless policy s.t. the probability reaching $\bullet > \lambda$
Simple POMDPs

Every state is of either type 1 or type 2

Nondeterministic choice, Dirac Distributions

Unique choice, Distribution over successor states
Encoding feasibility in Existential Theory of the Reals (ETR)

Does a valuation exist s.t. a red state is reached with probability is more than 1/6?

\[ \exists p_i \exists x, y : \]
\[ 0 < x < 1, \ 0 < y < 1 \]
\[ p_1 > \frac{1}{6} \]

Yes, iff the constraints are satisfiable
Solving systems of polynomials — in general — is exponential in number of variables

states + parameters
Eliminating state-variables

Results in a rational function $f(x)$ over the parameters $x$

State elimination (as in NFAs) or Gaussian elimination w/ polynomials

For a pMC with $k$ parameters, $n$ states and linear polynomials as probabilities:

- The rational function can be exponential in $k$ (even for acyclic pMCs)
- For any fixed $k$, the computation can be done in polynomial time in $n$
Result of state elimination

41 States - 138 Transitions - 2 Parameters

The numerator has 408 terms,
The denominator is the product of 48 linear polynomials

5 seconds
Solving polynomial inequality — in general — is exponential in number of variables
Recap: Complexity theory

- SAT: Given a Boolean formula, is it satisfiable?
- Has the LP a feasible solution?
- Given a polynomial, has it a real root?
- Given a set of polynomials, do they have a common root?

Complexity classes:
- P
- NP
- PSPACE
- coNP
- ETR
- coETR
- coNP
- PTIME
Recap: Complexity theory

Given a polynomial, has it a real root?

SAT: Given a Boolean formula, is it satisfiable?

Has the linear program a feasible solution?

Given a set of polynomials, do they have a common root?

PSPACE

ETR

coETR

NP

coNP

PTIME

TAUT: Given a Boolean formula, is it a tautology?

Given a polynomial, has it a real root?
How difficult is parameter synthesis?

Given: a parametric MC $\mathcal{M}$ with parameters $x$
exists: val: $x \rightarrow [0,1]$ s.t.: in $\mathcal{M}[\text{val}]$ a red state is reached with probability [relation] $\lambda$

<table>
<thead>
<tr>
<th>model</th>
<th>relation</th>
<th>result</th>
</tr>
</thead>
<tbody>
<tr>
<td>pMC</td>
<td>$\leq \geq$</td>
<td>ETR-complete</td>
</tr>
<tr>
<td></td>
<td>$\ll \gg$</td>
<td>NP-hard in ETR</td>
</tr>
</tbody>
</table>

[CONCUR’19]

The complexity of Finding Memoryless POMDP Policies
Sebastian Junges
OWLS 2020
Encoding polynomial inequalities as pMC

Given any polynomial \( f \)
is there a variable valuation \( \text{val} \) s.t. \( f(\text{val}) \geq \kappa \)

Given any pMC
is there a parameter valuation s.t. the probability reaching \( \bullet \geq \lambda \)

\[-2x^2y + y \geq 5\]

\[2 \cdot ((1 - x)xy + (1 - x)y + (1 - y) - 1) + y \geq 5\]

\[2 \cdot (1 - x)xy + 2 \cdot (1 - x)y + 2 \cdot (1 - y) + y \geq 7\]

\[\frac{8}{8}\]

Probability of reaching \( \bullet \) at least 7/8

\[1 - x \quad x \quad y\]

\[2/8 \quad 1 - x \quad y\]

\[2/8 \quad 2/8 \quad 1 - y\]

\[1/8 \quad y\]
How difficult is parameter synthesis?

Given: a parametric MC $\mathcal{M}$ with parameters $x$

exists: $\text{val}: x \rightarrow [0,1]$

s.t.: in $\mathcal{M}[\text{val}]$ a red state is reached with probability [relation] $\lambda$

Solving polynomial inequality — in general — is exponential in number of variables

Exponential in parameters

Parameters

In P

[Hutschenreiter et al. 2017]
What about parametric MDPs?

Given: a parametric MDP \( \mathcal{M} \) with parameters \( \mathbf{x} \)

exists: \( \text{val}: \mathbf{x} \rightarrow [0, 1] \) such that for all \( \sigma: S \rightarrow \text{Act} \): \( \mathcal{M}_\sigma[\text{val}] \models \phi \)

Selecting an action in every state
How difficult is parameter synthesis?

The complexity landscape for parameter synthesis (simplified)

<table>
<thead>
<tr>
<th>Model</th>
<th>Relation</th>
<th>Parameters</th>
<th>Arbitrarily Many</th>
<th>A-priori Fixed</th>
</tr>
</thead>
<tbody>
<tr>
<td>pMC</td>
<td>≤ ≥</td>
<td>ETR-complete</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt; &gt;</td>
<td>NP-hard in ETR</td>
<td></td>
<td>[Hutschenreiter et al. 2017]</td>
</tr>
<tr>
<td>pMDP</td>
<td>≤ ≤ &gt; ≥</td>
<td>ETR-complete</td>
<td></td>
<td>in NP</td>
</tr>
</tbody>
</table>

ETR encoding as extension of the standard LP for MDPs

How to eliminate state variables?
Graph preservation

$x \mapsto 0$ is not graph preserving

<table>
<thead>
<tr>
<th>$\mathcal{M}$</th>
<th>$\mathcal{M}[x \mapsto 0]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1 - x$</td>
<td>$1$</td>
</tr>
<tr>
<td>$x$</td>
<td></td>
</tr>
<tr>
<td>$\xrightarrow{}$</td>
<td>$\xrightarrow{}$</td>
</tr>
</tbody>
</table>

Practical methods exploit continuity.

The (presented) complexity analysis is unaffected by this observation.
Problem statement: Parameter synthesis

Given: a parametric MDP $\mathcal{M}$ with parameters $\mathbf{x}$

Find: $\text{val}: \mathbf{x} \to [0,1]$

such that: $\mathcal{M}_\sigma[\text{val}] \models \varphi$, i.e., a red state is reached with probability at least/at most $\lambda$
Practical parameter synthesis

Two settings

- Care for all points
- Only care for some satisfactory point
Several variants of encoding via SMT solvers [Hahn et al. 2011]

Parameter lifting: abstraction-refinement

Methods assume and exploit (to some extent) that the graph structure is fixed.

Several variants of encoding via SMT solvers [Hahn et al. 2011]

Parameter lifting: abstraction-refinement

Methods assume and exploit (to some extent) that the graph structure is fixed.
The complexity of Finding Memoryless POMDP Policies

Sebastian Junges

OWLS 2020
Related work ... necessarily incomplete here.

- **Infinite state systems**: e.g., Chakarov et al., Esparza et al., Kaminski et al., Zuck et al., etc.
- **Modal transition systems**: e.g., Benes et al., Delahaye et al.
- **Interval/Constraint MDPs**: e.g., Delahaye et al., Chatterjee et al., Chen et al., Hahn et al., Larsen et al.
- **Various Applications**: e.g., Aflaki et al., Calinescu et al., Fillieri et al., Polgreen et al., Rosenblum et al.
- **POMDPs with small strategies**: e.g., Chatterjee et al., Amato et al.
- **Quantitative Verification of Software Product Lines**: e.g., Ghezzi et al., Terbeek et al.
- **Parametric Continuous-Time MCs**: e.g., Ceska et al., Han et al.
Future challenges

The complexity of feasibility in pMDPs with one parameter

Robust strategies instead of (parameter) feasibility

Parameter Synthesis

feasibility = \exists \text{val}: x \rightarrow [0,1] \text{ such that for all } \sigma: S \rightarrow \text{Act}

robust strategies = \exists \sigma: S \rightarrow \text{Act} \text{ such that for all } \text{val}: x \rightarrow [0,1]

New challenges for verification:
Expensive (but powerful) abstraction techniques & Symbolic probabilistic model checking
Want to know more?

Want to know more?

For a formal treatment:

Sebastian Junges, Nils Jansen, Ralf Wimmer, Tim Quatmann, Leonore Winterer, Joost-Pieter Katoen, Bernd Becker:  
Finite-State Controllers of POMDPs using Parameter Synthesis. UAI 2018: 519-529

Sebastian Junges, Joost-Pieter Katoen, Guillermo A. Pérez, Tobias Winkler:  