The Core of MDP

Focus on what is important

Jan Křetínský    Tobias Meggendorfer
16. Sept. 2020

Technical University of Munich
Preface – Context and Motivation

- Context: Probabilistic systems (MDP)
- But: Approach generally applicable

- Problem: (Practical) Systems usually
  1. very large (billions of states)
  2. with many barely relevant states

⇒ Identify relevant states + restrict computation to these
Markov Decision Processes

- States + Actions → Probabilities
- Strategies remove non-determinism
- Objectives formalize goals
Example

- Goal: Reach success
- Maximal probability: $1 - 10^{-9}$
- Let’s modify the example
State of the Art

- Exact probability: We don’t know!
- But: We can make an educated guess ...

⇒ Don’t need to investigate unknown region (already known)
- Actually: Don’t need “??” region for any property
- ⇒ Region not important
The Core

- $S_\varepsilon$ is $\varepsilon$-core if system remains inside with prob. $\geq 1 - \varepsilon$
- Reachability restricted to "core states" $\Rightarrow$ error at most $\varepsilon$
- 0-core = set of reachable states
- In paper:
  - Sampling / Learning approach to identify cores
  - Possibility of heuristic guidance
  - Connection to other objectives / systems
  - Step-bounded variant with further applications
Conclusion

- Cores are an **intrinsic** property of the system (⇒ reusable)
- Potentially **significantly** smaller and easier to understand
- Many applications, from analysis to design aid
- Key Idea: Only likely reached states are interesting

---

Now “??” region is “interesting” (for infinite horizon)

In paper:
- On-the-fly MEC quotient
- Finite-horizon cores