Risk-Aware Stochastic Shortest Path

Tobias Meggendorfer

IST Austria

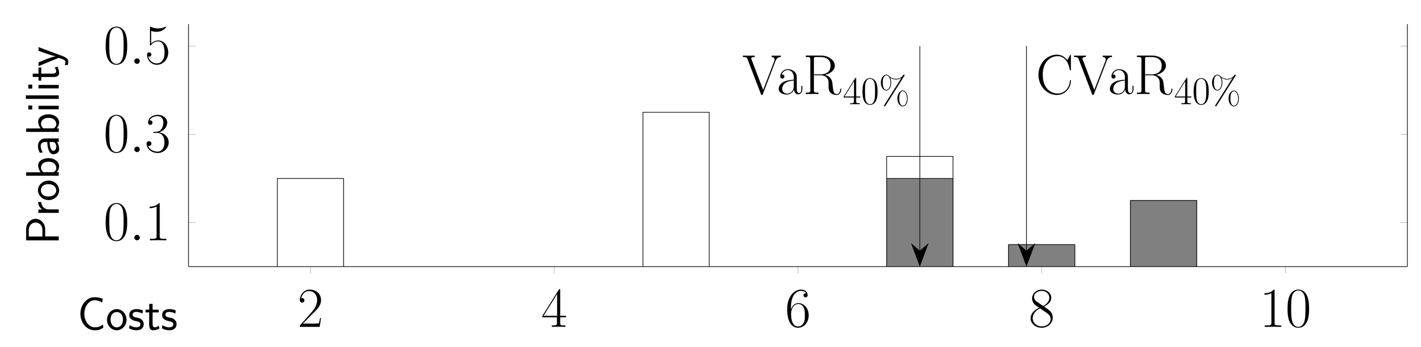
At a Glance

- Stochastic Shortest Path: Minimize cost to target region in stochastic environment
- Risk-Aware: Instead of expectation, consider *risk measure*
- Problem of this work[1]: Minimize *conditional value-at-risk* of total cost until target
- Two solutions (linear programming and value iteration)

Motivation – Control Mars Rover

- Risk assessment / aversion imperative to safety-critical systems
- Want both good average performance and little risk of failures
- Maximizing expectation not good enough
  - Failure may occur with little probability ⇒ little impact on expectation
  - High risk, high reward behaviour incentivized
- Completely avoiding bad behaviour (worst-case) neither
  - Any reasonable model has *some* probability of complete failure
  - Only really "safe" strategy: Do not move at all
- Thus: Consider probabilistic risk measure

The Conditional Value-at-Risk



- $\mathrm{VaR}_p(X)$ (the *Value-at-Risk*): "What is a reasonable bad case?"
- $\mathrm{CVaR}_p(X)$: "What happens in the average bad case?"

Definition: Let X be a random variable and $p \in (0, 1)$. Then

$$\mathrm{VaR}_p(X) := \min\{v \in \mathbb{N}_0 \mid \Sigma_{x=v+1}^{\infty} X(x) \leq t\}.$$

With $v = \mathrm{VaR}_p(X)$ and $\mathfrak{V} := \{X > v\}$:

$$\mathrm{CVaR}_p(X) := \tfrac{1}{p}(\mathbb{P}[\mathfrak{V}] \cdot \mathbb{E}[X \mid \mathfrak{V}] + (p - \mathbb{P}[\mathfrak{V}]) \cdot v).$$

Interesting properties:

- Interpolation between worst-case ($p \to 0$) and expectation ($p \to 1$)
- Robust to changes in X and p caused by, e.g., modelling errors
- *Coherent* risk measure

Model and Goal

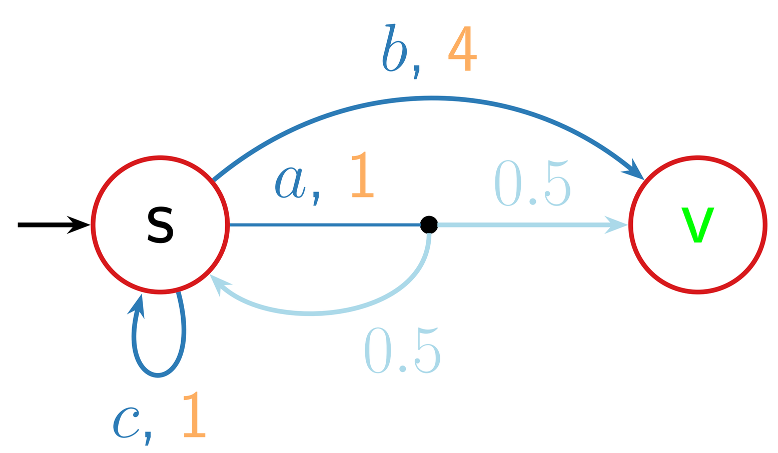
Markov Decision Process (MDP)

- Standard Model for single actor in random environment
- Comprises: States, Actions, and Transition Probabilities
- Evolution: In state, choose action, draw successor from distribution

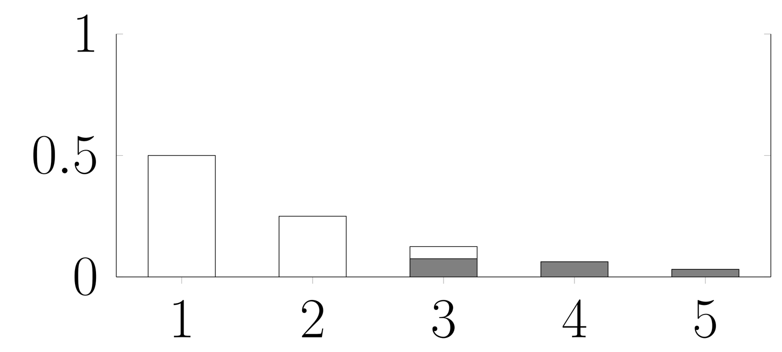
Stochastic Shortest Path (SSP)

- Additionally cost per action and target
- Goal: Minimize risk of total cost until target is reached

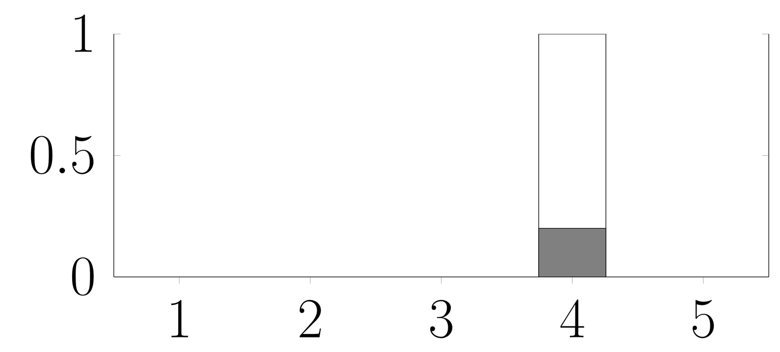
Example



- Action a preferred over b in expectation (2 vs. 4)
- But a is "risky" – significant chance to still be in s after e.g. 6 steps
- The distributions:



a: $\mathrm{CVaR}_{20\%} = 4.25$



b: $\mathrm{CVaR}_{20\%} = 4$

- Action b preferred over a for CVaR with small enough p!

Difficulties

- VaR of optimal CVaR may be exponential
- Optimal policies may require exponential memory
- Trade-off between moving to target efficiently and "risky" actions

Solution Insight

- CVaR ≙ SSP of worst p outcomes
- Assume VaR is n. At step n:
  - $1 - p$ probability mass reached target
  - remaining p somewhere else in the system

⇒ CVaR is weighted average of SSP for remaining part

Linear Programming

- Idea: Given "guess" for VaR, can find optimal strategies with LP
- Try out all possible VaRs
- EXPTIME algorithm (exp. many LP of exp. size)

Value Iteration

- Trade-off problem ⇒ Pareto sets
- Define $\mathfrak{P}_n^s \subseteq [0, 1] \times [0, \infty)$: contains (p, e) iff at step n
  1. goal can be reached with prob. $\geq p$
  2. remaining expected time to reach goal $\leq e$
- Central results of the paper:
  - Can derive achievable CVaR from $\mathfrak{P}_n^s$
  - $\mathfrak{P}_n^s$ is convex polygon
  - $\mathfrak{P}_{n+1}^s$ obtainable from combination of $\mathfrak{P}_n^{s'}$
- Minkowski sum of convex polygons in 2D: PTIME
- EXPTIME worst case, but comparatively fast in practice

[1] T. Risk-aware stochastic shortest path. AAAI'22, 2022.