Cost Partitioning Heuristics for Stochastic Shortest-Path Problems

Thorsten Klößner, Florian Pommerening, Thomas Keller, Gabriele Röger

kloessner@cs.uni-saarland.de, {florian.pommerening, tho.keller, gabriele.roeger}@unibas.ch

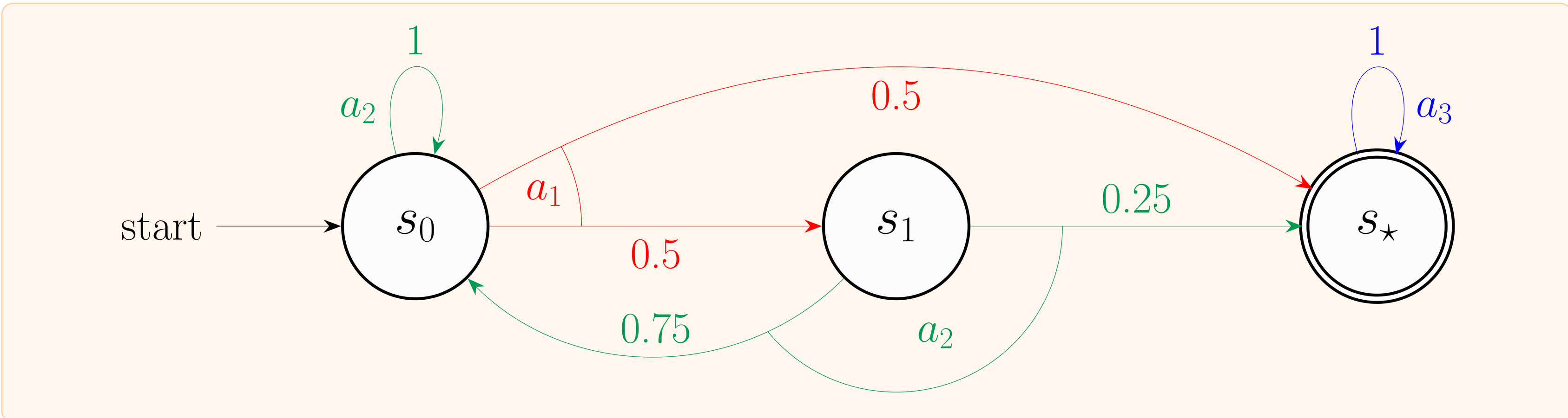
Universität des Saarlandes
Universität Basel

Stochastic Shortest-Path Problems

A *Stochastic Shortest-Path Problem* (SSP, Bertsekas and Tsitsiklis (1991)) consists of

- A finite set of *states* S
- A finite set of *actions* A
- A *transition probability function* $T : S \times A \times S \to [0, 1]$
- An *initial state* $s_0 \in S$
- A *goal state* $s_\star \in S$
- A *state-dependent cost function* $c : S \times A \to \mathbb{R}_0^+$

For every state-action pair $\langle s, a\rangle$ either a is *applicable* in s, i.e. $T(s, a, \cdot)$ is a probability distribution, or $T(s, a, t) = 0$ for all $t \in S$. $A(s)$ is the set of actions applicable in s.



Policy $\pi : S \to A$ specifies behaviour and is *optimal* for a state s if

- The goal state is reached with certainty when starting in s and executing π
- Among all such policies, π has the lowest expected cost-to-goal

Optimal state value $J^\star(s)$ is the expected cost-to-goal of an optimal policy for s

Heuristic Search

An optimal policy can be found by SSP heuristic search algorithms.

Requires an *admissible* heuristic $h : S \to \mathbb{R}$: $h(s) \le J^\star(s)$ for all $s \in S$.

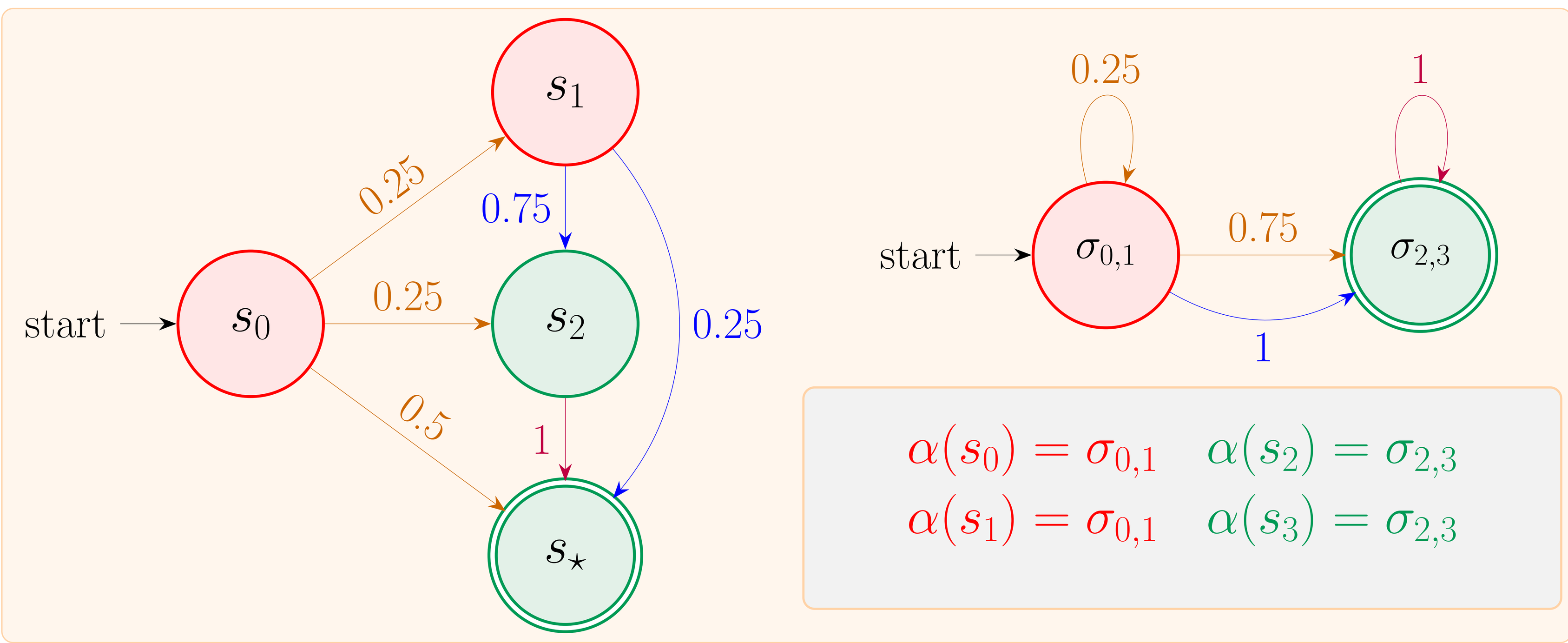
We assume that h is also dependent on the cost function c of the SSP, and write $h(s, c)$.

h is *generally admissible* if $h(s, c') \le J^\star(s)$ for all cost functions c'.

Abstraction Heuristics

Abstractions are induced by a function $\alpha : S \to S_\alpha$, where S_α are the *abstract states*.

The *abstraction mapping* α induces an abstract SSP with state set S_α.



Induces the *abstraction heuristic* h^α. $h^\alpha(s, c')$ is the optimal state value of $\alpha(s)$ in the abstract SSP obtained from the original SSP with cost function c'.

Theorem. h^α is generally admissible.

Cost Partitioning for SSPs

A cost partition is a tuple $\langle c_1, \dots, c_n\rangle$ that splits c into subadditive parts:

$$c_1(s, a) + \dots + c_n(s, a) \le c(s, a) \qquad \forall s \in S\ \forall a \in A(s)$$

Operator cost partitions assign only state-independent costs.

$$c_i(s, a) = c_i(t, a) \qquad \forall i \in \{1, \dots, n\}\ \forall s, t \in S\ \forall a \in A(s) \cup A(t)$$

For heuristics $\mathcal{H} = \langle h_1, \dots, h_n\rangle$ and cost partition $P = \langle c_1, \dots, c_n\rangle$, define

$$h^{\mathcal{H},P}(s) = h_n(s, c_1) + \dots + h_n(s, c_n)$$

as the *cost partitioning heuristic* for $\mathcal{H}$ and P.

Theorem. If $h \in \mathcal{H}$ are generally admissible, then $h^{\mathcal{H},P}$ is admissible (for any P).

Optimal Cost Partitioning

Problem: Given $\mathcal{H}$ and a state $s \in S$, how do we maximize $h^{\mathcal{H},P}(s)$?

Define the *optimal cost partitioning heuristic* for $\mathcal{H}$ as

$$h^{\text{OTCP}}_{\mathcal{H}}(s) := \sup_{P \text{ cost partition}} \left\{ h^{\mathcal{H},P}(s) \right\}.$$

Likewise, define the optimal *operator* cost partitioning heuristic $h^{\text{OOCP}}_{\mathcal{H}}$.

For abstraction heuristics $\mathcal{H} = \langle h^{\alpha_1}, \dots, h^{\alpha_n}\rangle$, $h^{\text{OTCP}}_{\mathcal{H}}(\bar{s})$ can be computed by an LP:

$$
\begin{array}{lll}
\textbf{Maximize} & \mathsf{y}_{\alpha_1(\bar{s})} + \dots + \mathsf{y}_{\alpha_n(\bar{s})} & \\
\textbf{subject to} & \mathsf{y}_{\alpha_i(s_\star)} = 0 & 1 \le i \le n \\
& \mathsf{y}_{\alpha_i(s)} \le \mathsf{c}_{\alpha sa} + \sum_{t \in S} T(s, a, t)\mathsf{y}_{\alpha(t)} & \forall s \in S, a \in A(\alpha(s)), 1 \le i \le n \\
& \mathsf{c}_{\alpha_1 sa} + \dots + \mathsf{c}_{\alpha_n sa} \le c(s, a) & \forall s \in S, a \in A
\end{array}
$$

The optimal objective value of this LP is $h^{\text{OTCP}}_{\mathcal{H}}(\bar{s})$!

Caveat: Size of the linear program is $\mathcal{O}(\mathbf{n} \cdot |\mathbf{S}||\mathbf{A}|)$

If we only want to compute $h^{\text{OOCP}}_{\mathcal{H}}(s)$, the LP size can be reduced to $\Theta(\sum_{\mathbf{1 \le i \le n}} |\mathbf{S}_{\alpha_i}||\mathbf{A}|)$!

Relation to Occupation Measure Heuristics

Occupation Measure Heuristics (Trevizan, Thiébaux, and Haslum, 2017) h^{pom} and h^{roc} are closely related to optimal cost partitioning over abstraction heuristics

Projection occupation measure heuristic h^{pom} combines atomic projections (abstractions using single state variables) using a linear program

Interestingly, this LP is the dualization of the LP computing $h^{\text{OOCP}}_{\mathcal{H}}$ for atomic abstractions!

Theorem. If $\mathcal{H}$ is the set of all atomic projection heuristics, $h^{\text{pom}} = h^{\text{OOCP}}_{\mathcal{H}}$.

Even holds for h^{roc} under a syntactical restriction we call *transition normal form*.

Theorem. For problems in transition normal form, $h^{\text{pom}} = h^{\text{roc}}$.

Relation to Approximate Linear Programming (ALP)

ALP was introduced for discounted-reward infinite-horizon MDPs by (Guestrin et al., 2003)

For SSPs, approximates $J^\star(s)$ with a linear combination of *basis functions* $f_1, \dots, f_n$:

$$h(s) := w_1 f_1(s) + \dots + w_n f_n(s) \approx J^\star(s) \qquad \text{where } f_i : S \to \mathbb{R}$$

The weights $w_1, \dots, w_n$ are optimized in a linear program for a weighted sum of heuristic values $\sum_{s \in S} \rho(s)h(s)$ where the *state relevance function* ρ specifies the state weights.

Idea: What if we encode abstraction mappings $\alpha_1, \dots, \alpha_n$ into the basis functions?

For abstraction mapping α_i and abstract state $\sigma \in S_{\alpha_i}$, define the indicator function

$$f^{\alpha_i}_\sigma(s) = \begin{cases} 1 & \text{if } \alpha(s) = \sigma, \\ 0 & \text{otherwise.} \end{cases}$$

Theorem. ALP over the basis functions $f^{\alpha_i}_\sigma$ computes a transition cost partitioning P that maximizes $\sum_{s \in S} \rho(s)h^{\mathcal{H},P}(s)$ for the abstraction heuristics $\mathcal{H} = \langle h^{\alpha_1}, \dots, h^{\alpha_n}\rangle$.

Generalizes the link to potential heuristics (Pommerening et al., 2015) in classical planning!

References

Bertsekas, D. P.; and Tsitsiklis, J. N. 1991. An Analysis of Stochastic Shortest Path Problems. Mathematics of Operations Research, 16: 580–595.

Guestrin, C.; Koller, D.; Parr, R.; and Venkataraman, S. 2003. Efficient Solution Algorithms for Factored MDPs. 19: 399–468.

Pommerening, F.; Helmert, M.; Röger, G.; and Seipp, J. 2015. From Non-Negative to General Operator Cost Partitioning. In Bonet, B.; and Koenig, S., eds., Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI 2015), 3335–3341. AAAI Press.

Trevizan, F. W.; Thiébaux, S.; and Haslum, P. 2017. Occupation Measure Heuristics for Probabilistic Planning. In Barbulescu, L.; Frank, J.; Mausam; and Smith, S. F., eds., Proceedings of the Twenty-Seventh International Conference on Automated Planning and Scheduling (ICAPS 2017), 306–315. AAAI Press.