Planning and Optimization

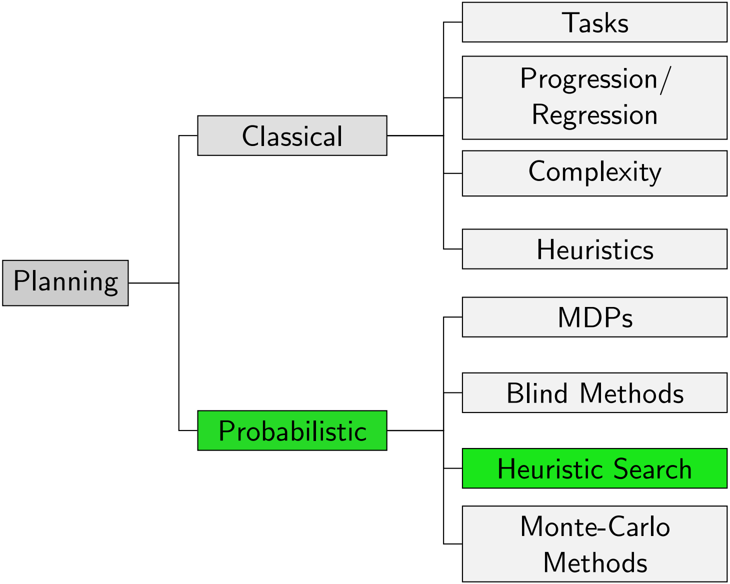
G1. Heuristic Search: AO* & LAO* Part I

Gabriele Röger and Thomas Keller

Universität Basel

December 3, 2018

Content of this Course



Heuristic Search

Heuristic Search: Recap

Heuristic Search Algorithms

Heuristic search algorithms use heuristic functions to (partially or fully) determine the order of node expansion.

(From Lecture 15 of the AI course last semester)

Best-first Search: Recap

Best-first Search

A best-first search is a heuristic search algorithm
that evaluates search nodes with an evaluation function f
and always expands a node n with minimal $f(n)$ value.

(From Lecture 15 of the AI course last semester)

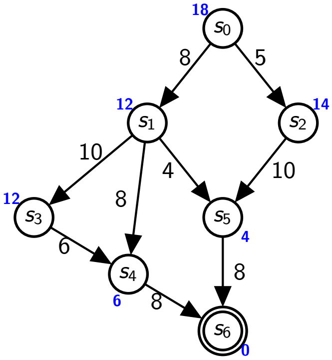
A^*Search: Recap

A^*Search

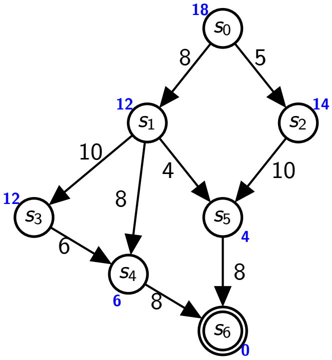
A^* is the best-first search algorithm with evaluation function $f(n) = g(n) + h(n.\text{state})$.

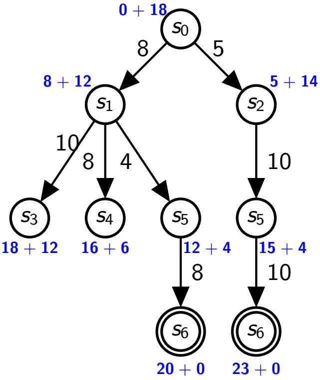
(From Lecture 15 of the AI course last semester)

A* Search (With Reopening): Example



A* Search (With Reopening): Example





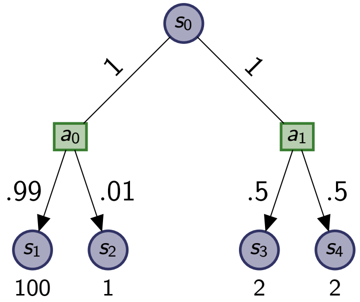
Motivation

From A* to AO*

- Equivalent of A* in (acyclic) probabilistic planning is AO*
- Even though we know A* and foundations of probabilistic planning, the generalization is far from straightforward:
  - e.g., in A*, $g(n)$ is cost from root n_0 to n
  - equivalent in AO* is expected cost from n_0 to n

Expected Cost to Reach State

Consider the following expansion of state s_0:



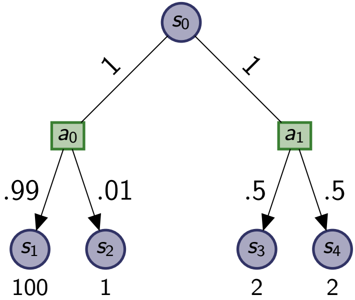
Expected cost to reach any of the leaves is infinite or undefined (neither is reached with probability 1).

From A^* to AO^*

- Equivalent of A^* in (acyclic) probabilistic planning is AO^*
- Even though we know A^* and foundations of probabilistic planning, the generalization is far from straightforward:
  - e.g., in A^*, $g(n)$ is cost from root n_0 to n
  - equivalent in AO^* is expected cost from n_0 to n
  - alternative could be expected cost from n_0 to n given n is reached

Expected Cost to Reach State Given It Is Reached

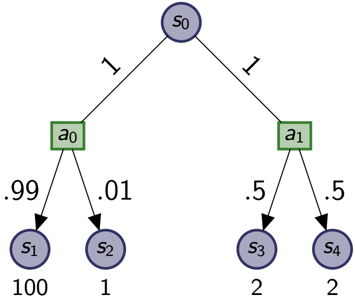
Consider the following expansion of state s_0:



Conditional probability is misleading: s_2 would be expanded, which isn't part of the best looking option

The Best Looking Action

Consider the following expansion of state s_0:



Conditional probability is misleading: s_2 would be expanded, which isn't part of the best looking option:
with state-value estimate $\hat{V}(s) := h(s)$, greedy action $a_{\hat{V}}(s) = a_1$

Expansion in Best Solution Graph

AO^* uses different idea:

- AO^* keeps track of best solution graph
- AO^* expands a state that can be reached from s_0 by only applying greedy actions
- $\Rightarrow$ no g-value equivalent required

Expansion in Best Solution Graph

AO* uses different idea:

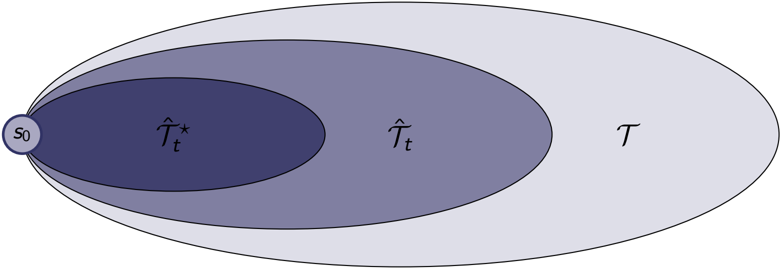
- AO* keeps track of best solution graph
- AO* expands a state that can be reached from s_0 by only applying greedy actions
- ⇒ no g-value equivalent required
- Equivalent version of A* built on this idea can be derived ⇒ A* with backward induction
- Since change is non-trivial, we focus on A* variant now
- and generalize later to acyclic probabilistic tasks (AO*)
- and probabilistic tasks in general (LAO*)

A* with Backward Induction

Transition Systems

A* with backward induction distinguishes three transition systems:

- The transition system $\mathcal{T} = \langle S, L, c, T, s_0, S^\star \rangle$
  $\Rightarrow$ given implicitly
- The explicated graph $\hat{\mathcal{T}}_t = \langle \hat{S}_t, L, c, \hat{T}_t, s_0, S^\star \rangle$
  $\Rightarrow$ the part of $\mathcal{T}$ explicitly considered during search
- The partial solution graph $\hat{\mathcal{T}}^\star_t = \langle \hat{S}^\star_t, L, c, \hat{T}^\star_t, s_0, S^\star \rangle$
  $\Rightarrow$ The part of $\hat{\mathcal{T}}_t$ that contains best solution



Explicated Graph

- Expanding a state s at time step t explicates all successors $s' \in \text{succ}(s)$ by adding them to explicated graph:

$$\hat{\mathcal{T}}_t = \langle \hat{S}_{t-1} \cup \text{succ}(s), L, c, \hat{T}_{t-1} \cup \{\langle s, l, s' \rangle \in T\}, s_0, S^\star \}$$

- Each explicated state is annotated with state-value estimate $\hat{V}_t(s)$ that describes estimated cost to a goal at time step t
- When state s' is explicated and $s' \notin \hat{S}_{t-1}$, its state-value estimate is initialized to $\hat{V}_t(s') := h(s')$
- We call leaf states of $\hat{\mathcal{T}}_t$ fringe states

Partial Solution Graph

- The partial solution graph $\hat{\mathcal{T}}_t^\star$ is the subgraph of $\hat{\mathcal{T}}_t$ that is spanned by the smallest set of states $\hat{S}_t^\star$ that satisfies:
  - $s_0 \in \hat{S}_t^\star$
  - if $s \in \hat{S}_t^\star$, $s' \in \hat{S}_t$ and $\langle s, a_{\hat{V}_t(s)}(s), s' \rangle \in \hat{T}_t$, then s' in $\hat{S}_t^\star$
- The partial solution graph forms a sequence of states $\langle s_0, \ldots, s_n \rangle$, starting with the initial state s_0 and ending in the greedy fringe state s_n

Backward Induction

- A* with backward induction does not maintain static open list
- State-value estimates determine partial solution graph
- Partial solution graph determines which state is expanded
- (Some) state-value estimates are updated in time step t by backward induction:

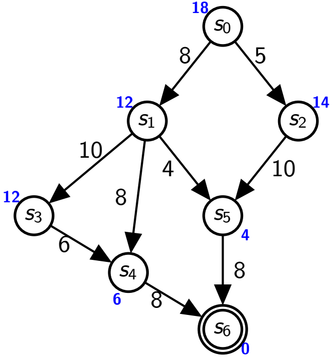
$$\hat{V}_t(s) = \min_{\langle s,l,s'\rangle \in \hat{T}_t(s)} c(l) + \hat{V}_t(s')$$

A* with backward induction

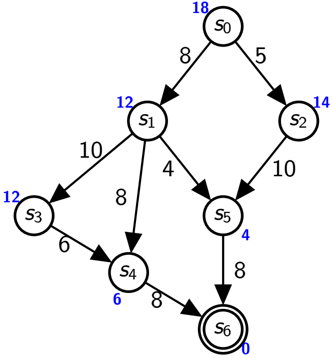
A* with backward induction for classical planning task $\mathcal{T}$

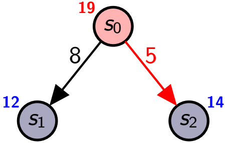
explicate s_0
while greedy fringe state $s \notin S_\star$:
    expand s
    perform backward induction of states in $\hat{\mathcal{T}}^\star_{t-1}$ in reverse order
return $\hat{\mathcal{T}}^\star_t$

A* with backward induction

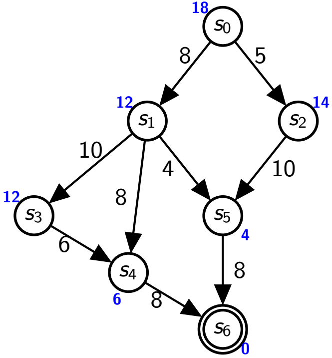


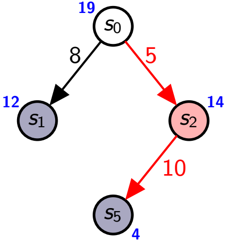
A* with backward induction



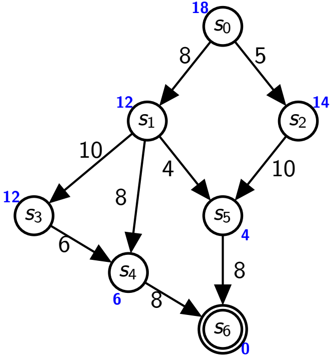


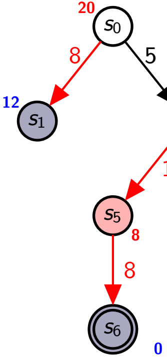
A* with backward induction



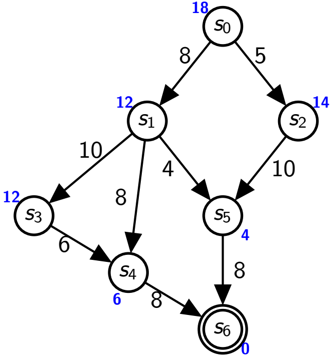


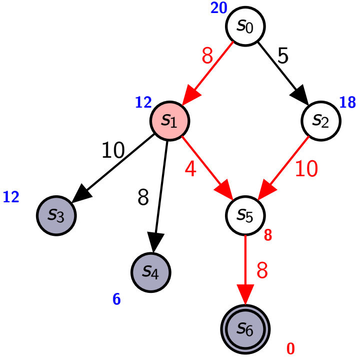
A* with backward induction



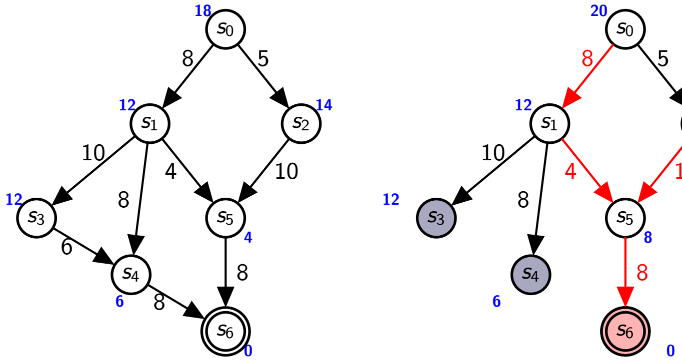


A* with backward induction





A* with backward induction



Equivalence of A* and A* with Backward Induction

Theorem

A and A* with Backward Induction expand the same set of states if run with identical admissible heuristic h and identical tie-breaking criterion.*

Proof Sketch.

The proof shows that

- there is always a unique state s in greedy fringe of A* with backward induction
- $f(s) = g(s) + h(s)$ is minimal among all fringe states
- $g(s)$ of fringe node s encoded in greedy action choices
- $h(s)$ of fringe node equal to $\hat{V}_t(s)$

Summary

Summary

- Non-trivial to generalize A^* to probabilistic planning
- For better understanding of AO^*, we change A^* towards AO^*
- Derived A^* with backward induction, which is similar to AO^*
- and expands identical states as A^*