Principles of AI Planning
10. Planning as search: abstractions

Coming up with heuristics in a principled way

General procedure for obtaining a heuristic
Solve an easier version of the problem.

Two common methods:
- relaxation: consider less constrained version of the problem
- abstraction: consider smaller version of real problem

In previous chapters, we have studied relaxation, which has been very successfully applied to satisficing planning. Now, we study abstraction, which is one of the most prominent techniques for optimal planning.

Abstracting a transition system

Abstracting a transition system means dropping some distinctions between states, while preserving the transition behaviour as much as possible.

- An abstraction of a transition system $\mathcal{I}$ is defined by an abstraction mapping $\alpha$ that defines which states of $\mathcal{I}$ should be distinguished and which ones should not.
- From $\mathcal{I}$ and $\alpha$, we compute an abstract transition system $\mathcal{I}'$ which is similar to $\mathcal{I}$, but smaller.
- The abstract goal distances (goal distances in $\mathcal{I}'$) are used as heuristic estimates for goal distances in $\mathcal{I}$.
Abstractions: informally

Introduction

Practical requirements

Multiple abstractions

Outlook

Abstractions: formally

Summary

Abstracting a transition system: example

Example (15-puzzle)

A 15-puzzle state is given by a permutation \(\langle b, t_1, \ldots, t_{15} \rangle\) of \(\{1, \ldots, 16\}\), where \(b\) denotes the blank position and the other components denote the positions of the 15 tiles.

One possible abstraction mapping ignores the precise location of tiles 8–15, i.e., two states are distinguished if and only if they differ in the position of the blank or one of the tiles 1–7:

\[
\alpha(\langle b, t_1, \ldots, t_{15} \rangle) = \langle b, t_1, \ldots, t_7 \rangle
\]

The heuristic values for this abstraction correspond to the cost of moving tiles 1–7 to their goal positions.

Abstraction example: 15-puzzle

real state space

- \(16! = 20922789888000 \approx 2 \cdot 10^{13}\) states
- \(\frac{16!}{2} = 10461394944000 \approx 10^{13}\) reachable states

abstract state space

- \(16 \cdot 15 \cdots 9 = 518918400 \approx 5 \cdot 10^8\) states
- \(16 \cdot 15 \cdots 9 = 518918400 \approx 5 \cdot 10^8\) reachable states
Given $T$ and $\alpha$, how do we compute $T'$?

**Requirement**

We want to obtain an admissible heuristic. Hence, $h^*(\alpha(s))$ (in the abstract state space $T'$) should never overestimate $h^*(s)$ (in the concrete state space $T$).

An easy way to achieve this is to ensure that all solutions in $T$ also exist in $T'$:

- If $s$ is a goal state in $T$, then $\alpha(s)$ is a goal state in $T'$.
- If $T$ has a transition from $s$ to $t$, then $T'$ has a transition from $\alpha(s)$ to $\alpha(t)$.

---

**Practical requirements for abstractions**

To be useful in practice, an abstraction heuristic must be efficiently computable. This gives us two requirements for $\alpha$:

- For a given state $s$, the abstract state $\alpha(s)$ must be efficiently computable.
- For a given abstract state $\alpha(s)$, the abstract goal distance $h^*(\alpha(s))$ must be efficiently computable.

There are different ways of achieving these requirements:

- **Pattern database heuristics** (Culberson & Schaeffer, 1996)
- merge-and-shrink abstractions (Dräger, Finkbeiner & Podelski, 2006)
- structural patterns (Katz & Domshlak, 2008)
- Cartesian abstractions (Ball, Podelski & Rajamani, 2001; Seipp & Helmert, 2013)

---

**Example (15-puzzle)**

In the running example:

- $T$ has the unique goal state $\langle 16, 1, 2, \ldots, 15 \rangle$. $\rightsquigarrow$ $T'$ has the unique goal state $\langle 16, 1, 2, \ldots, 7 \rangle$.
- Let $x$ and $y$ be neighboring positions in the $4 \times 4$ grid. $T$ has a transition from $\langle x, t_1, \ldots, t_{i-1}, y, t_{i+1}, \ldots, t_{16} \rangle$ to $\langle y, t_1, \ldots, t_{i-1}, x, t_{i+1}, \ldots, t_{15} \rangle$ for all $i \in \{1, \ldots, 15\}$. $\rightsquigarrow$ $T'$ has a transition from $\langle x, t_1, \ldots, t_{i-1}, y, t_{i+1}, \ldots, t_7 \rangle$ to $\langle y, t_1, \ldots, t_{i-1}, x, t_{i+1}, \ldots, t_7 \rangle$ for all $i \in \{1, \ldots, 7\}$. $\rightsquigarrow$ Moreover, $T''$ has a transition from $\langle x, t_1, \ldots, t_7 \rangle$ to $\langle y, t_1, \ldots, t_7 \rangle$ if $y \notin \{t_1, \ldots, t_7\}$.
Multiple abstractions

- One important practical question is how to come up with a suitable abstraction mapping $\alpha$.
- Indeed, there is usually a huge number of possibilities, and it is important to pick good abstractions (i.e., ones that lead to informative heuristics).
- However, it is generally not necessary to commit to a single abstraction.

Combining multiple abstractions

Maximizing several abstractions:
- Each abstraction mapping gives rise to an admissible heuristic.
- By computing the maximum of several admissible heuristics, we obtain another admissible heuristic which dominates the component heuristics.
- Thus, we can always compute several abstractions and maximize over the individual abstract goal distances.

Adding several abstractions:
- In some cases, we can even compute the sum of individual estimates and still stay admissible.
- Summation often leads to much higher estimates than maximization, so it is important to understand when it is admissible.

Maximizing several abstractions: example

Example (15-puzzle)
- mapping to tiles 1–7 was arbitrary
  - can use any subset of tiles
- with the same amount of memory required for the tables for the mapping to tiles 1–7, we could store the tables for nine different abstractions to six tiles and the blank
- use maximum of individual estimates

Adding several abstractions: example

1st abstraction: ignore precise location of 8–15
2nd abstraction: ignore precise location of 1–7
  - Is the sum of the abstraction heuristics admissible?
### Adding several abstractions: example

<table>
<thead>
<tr>
<th></th>
<th>1st abstraction: ignore precise location of 8–15</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2nd abstraction: ignore precise location of 1–7</td>
</tr>
</tbody>
</table>

⇝ The sum of the abstraction heuristics is **not admissible.**

<table>
<thead>
<tr>
<th></th>
<th>1st abstraction: ignore precise location of 8–15 and blank</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2nd abstraction: ignore precise location of 1–7 and blank</td>
</tr>
</tbody>
</table>

⇝ The sum of the abstraction heuristics is **admissible.**

---

### Our plan for the next lectures

In the following, we take a deeper look at abstractions and their use for admissible heuristics.

- In the rest of this chapter, we **formally introduce** abstractions and abstraction heuristics and study some of their most important properties.
- In the **following chapter**, we discuss one particular class of abstraction heuristics in detail, namely pattern database heuristics.
Transition systems

Reminder from Chapter 2:

Definition (transition system)
A transition system is a 5-tuple $\mathcal{T} = (S, L, T, s_0, S_*)$ where
- $S$ is a finite set of states,
- $L$ is a finite set of (transition) labels,
- $T \subseteq S \times L \times S$ is the transition relation,
- $s_0 \in S$ is the initial state, and
- $S_* \subseteq S$ is the set of goal states.

We say that $\mathcal{T}$ has the transition $(s, \ell, s')$ if $(s, \ell, s') \in T$.
We also write this $s \xrightarrow{\ell} s'$, or $s \rightarrow s'$ when not interested in $\ell$.

Transition systems: example

Example task: one package, two trucks

Example (one package, two trucks)
Consider the following FDR planning task $\langle V, I, O, \gamma \rangle$:
- $V = \{p, t_A, t_B\}$ with
  - $\Delta_p = \{L, R, A, B\}$
  - $\Delta_{t_A} = \Delta_{t_B} = \{L, R\}$
- $I = \{p \mapsto L, t_A \mapsto R, t_B \mapsto R\}$
- $O = \{\text{pickup}_{i,j} \mid i \in \{A, B\}, j \in \{L, R\}\}$
  - $\cup \{\text{drop}_{i,j} \mid i \in \{A, B\}, j \in \{L, R\}\}$
  - $\cup \{\text{move}_{i,j,j'} \mid i, j, j' \in \{L, R\}, j \neq j'\}$, where
  - $\text{pickup}_{i,j} = \langle t_i = j \land p = j, p := i \rangle$
  - $\text{drop}_{i,j} = \langle t_i = j \land p = i, p := j \rangle$
  - $\text{move}_{i,j,j'} = \langle t_i = j, t_j := j' \rangle$
- $\gamma = (p = R)$
Abstractions

Transition systems: Abstractions [informally]

Abstractions: formally

Transition systems: Abstractions

Abstraction

Abstraction heuristics

Additivity

Refinements

Equivalence

Practice

Summary

Abstractions: terminology

Let $\mathcal{I}$ and $\mathcal{I}'$ be transition systems and $\alpha$ a function such that $\mathcal{I}'$ is an abstraction of $\mathcal{I}$ with abstraction mapping $\alpha$.

- $\mathcal{I}$ is called the concrete transition system.
- $\mathcal{I}'$ is called the abstract transition system.
- Similarly: concrete/abstract state space, concrete/abstract transition, etc.

We say that:

- $\mathcal{I}'$ is an abstraction of $\mathcal{I}$ (without mentioning $\alpha$)

- $\alpha$ is an abstraction mapping on $\mathcal{I}$ (without mentioning $\mathcal{I}'$)

Note: For a given $\mathcal{I}$ and $\alpha$, there can be multiple abstractions $\mathcal{I}'$, and for a given $\mathcal{I}$ and $\mathcal{I}'$, there can be multiple abstraction mappings $\alpha$.

Abstractions: example

Definition (abstraction, abstraction mapping)

Let $\mathcal{I} = \langle S, L, T, s_0, S_\star \rangle$ and $\mathcal{I}' = \langle S', L', T', s'_0, S'_\star \rangle$ be transition systems with the same label set $L = L'$, and let $\alpha : S \rightarrow S'$ be a surjective function.

We say that $\mathcal{I}'$ is an abstraction of $\mathcal{I}$ with abstraction mapping $\alpha$ (or: abstraction function $\alpha$) if

- $\alpha(s_0) = s'_0$,
- for all $s \in S$, we have $\alpha(s) \in S'$, and
- for all $\langle s, \ell, t \rangle \in T$, we have $\langle \alpha(s), \ell, \alpha(t) \rangle \in T'$.
Abstractions

Induced abstractions

Definition (induced abstractions)

Let $T = (S, L, T, s_0, S_\star)$ be a transition system, and let $\alpha : S \rightarrow S'$ be a surjective function.

The abstraction (of $T$) induced by $\alpha$, in symbols $T^\alpha$, is the transition system $T^\alpha = (S', L', T', s'_0, S'_\star)$ defined by:

- $T' = \{ (s, \ell, t) \in T \}$
- $s'_0 = \alpha(s_0)$
- $S'_\star = \{ \alpha(s) \mid s \in S_\star \}$

Note: It is easy to see that $T^\alpha$ is an abstraction of $T$. It is the "smallest" abstraction of $T$ with abstraction mapping $\alpha$.

Strictly homomorphic abstractions: example

This abstraction is a strictly homomorphic abstraction of the concrete transition system $T$. 

Abstractions: informally
Abstractions: formally
Transition systems
Abstractions
Abstraction heuristics
Additivity
Refinements
Equivalence
Practice
Summary

Abstractions: informally
Abstractions: formally
Transition systems
Abstractions
Abstraction heuristics
Additivity
Refinements
Equivalence
Practice
Summary

Induced abstractions: terminology

Let $T$ and $T'$ be transition systems and $\alpha$ be a function such that $T' = T^\alpha$ (i.e., $T'$ is the abstraction of $T$ induced by $\alpha$).

- $\alpha$ is called a strict homomorphism from $T$ to $T'$, and $T'$ is called a strictly homomorphic abstraction of $T$.
- If $\alpha$ is bijective, it is called an isomorphism between $T$ and $T'$, and the two transition systems are called isomorphic.

Note: Most arcs represent many parallel transitions.
Strictly homomorphic abstractions: example

If we add any goal states or transitions, it is still an abstraction of $\mathcal{I}$, but no longer a strictly homomorphic one.

Abstraction heuristics

Definition (abstr. heur. induced by an abstraction)
Let $\Pi$ be an FDR planning task with state space $S$, and let $\mathcal{A}$ be an abstraction of $\mathcal{I}(\Pi)$ with abstraction mapping $\alpha$. The abstraction heuristic induced by $\mathcal{A}$ and $\alpha$, $h^{\mathcal{A},\alpha}$, is the heuristic function $h^{\mathcal{A},\alpha} : S \rightarrow \mathbb{N}_0 \cup \{\infty\}$ which maps each state $s \in S$ to $h^{\mathcal{A},\alpha}(\alpha(s))$ (the goal distance of $\alpha(s)$ in $\mathcal{A}$).

Note: $h^{\mathcal{A},\alpha}(s) = \infty$ if no goal state of $\mathcal{A}$ is reachable from $\alpha(s)$

Definition (abstr. heur. induced by strict homomorphism)
Let $\Pi$ be an FDR planning task and $\alpha$ a strict homomorphism on $\mathcal{I}(\Pi)$. The abstraction heuristic induced by $\alpha$, $h^\alpha$, is the abstraction heuristic induced by $\mathcal{I}(\Pi)^\alpha$ and $\alpha$, i.e., $h^\alpha := h^{\mathcal{I}(\Pi)^\alpha,\alpha}$.
Consistency of abstraction heuristics

Theorem (consistency and admissibility of $h^\alpha$)
Let $\Pi$ be an FDR planning task, and let $\mathcal{A}$ be an abstraction of $\mathcal{I}(\Pi)$ with abstraction mapping $\alpha$.
Then $h^\alpha$ is safe, goal-aware, admissible and consistent.

Proof.
We prove goal-awareness and consistency; the other properties follow from these two.
Let $\mathcal{I} = \mathcal{I}(\Pi) = \langle S, L, T, s_0, S_\mathcal{F}\rangle$ and $\mathcal{A} = \langle S', L', T', s'_0, S'_\mathcal{F}\rangle$.
Goal-awareness: We need to show that $h^\alpha$ $\alpha$ $s$ $0$ for all $s \in S_\mathcal{F}$, so let $s \in S$. Then $\alpha(s) \in S'_\mathcal{F}$ by the definition of abstractions and abstraction mappings, and hence $h^\alpha(s) = h^\alpha_\mathcal{A}(\alpha(s)) = 0$.

Consistency of abstraction heuristics (ctd.)

Proof (ctd.)
Consistency: Let $s, t \in S$ such that $t$ is a successor of $s$. We need to prove that $h^\alpha(s) \leq h^\alpha(t) + 1$.
Since $t$ is a successor of $s$, there exists an operator $o$ with $app_o(s) = t$ and hence $\langle s, o, t \rangle \in T$.
By the definition of abstractions and abstraction mappings, we get $\langle \alpha(s), o, \alpha(t) \rangle \in T'$, so $\alpha(t)$ is a successor of $\alpha(s)$ in $\mathcal{A}$.
Therefore, $h^\alpha(s) = h^\alpha_\mathcal{A}(\alpha(s)) \leq h^\alpha_\mathcal{A}(\alpha(t)) + 1 = h^\alpha(t) + 1$,
where the inequality holds because the shortest path from $\alpha(s)$ to the goal in $\mathcal{A}$ cannot be longer than the shortest path from $\alpha(s)$ to the goal via $\alpha(t)$. □

Orthogonality of abstraction mappings

Definition (orthogonal abstraction mappings)
Let $\alpha_1$ and $\alpha_2$ be abstraction mappings on $\mathcal{I}$.
We say that $\alpha_1$ and $\alpha_2$ are orthogonal if for all transitions $\langle s, t \rangle$ of $\mathcal{I}$, we have $\alpha_1(s) \neq \alpha_2(t)$ for at most one $i \in \{1, 2\}$.

Affecting transition labels

Definition (affecting transition labels)
Let $\mathcal{A}$ be a transition system, and let $\ell$ be one of its labels.
We say that $\ell$ affects $\mathcal{A}$ if $\mathcal{A}$ has a transition $\langle s, \ell, t \rangle$ with $s \neq t$.

Theorem (affecting labels vs. orthogonality)
Let $\mathcal{A}_1$ be an abstraction of $\mathcal{I}$ with abstraction mapping $\alpha_1$.
Let $\mathcal{A}_2$ be an abstraction of $\mathcal{I}$ with abstraction mapping $\alpha_2$.
If no label of $\mathcal{I}$ affects both $\mathcal{A}_1$ and $\mathcal{A}_2$, then $\alpha_1$ and $\alpha_2$ are orthogonal.
(Easy proof omitted.)
Orthogonal abstraction mappings: example

Are the abstraction mappings orthogonal?

Orthogonality and additivity

Theorem (additivity for orthogonal abstraction mappings)

Let $h_{\alpha_i}$, $\alpha_i$, ..., $h_{\alpha_0}$, $\alpha_0$ be abstraction heuristics for the same planning task $\Pi$ such that $\alpha_i$ and $\alpha_j$ are orthogonal for all $i \neq j$. Then $\sum_{i=1}^{n} h_{\alpha_i}$ is a safe, goal-aware, admissible and consistent heuristic for $\Pi$. 

Orthogonality and additivity: example

transition system $T$

state variables: first package, second package, truck
Orthogonality and additivity: example

Abstractions:
- informally
- formally
- Transition systems
- Abstractions
- Abstraction heuristics
- Refinements
- Equivalence
- Practice
- Summary

Orthogonality and additivity: example

orthogonality and additivity: proof (ctd.)

Proof (ctd.)

Consistency: Let $s, t \in S$ such that $t$ is a successor of $s$. Let $L := \sum_{i=1}^{n} h^{\alpha_i(s)}$ and $R := \sum_{i=1}^{n} h^{\alpha_i(t)}$.

We need to prove that $L \leq R + 1$.

Since $t$ is a successor of $s$, there exists an operator $o$ with $app_o(s) = t$ and hence $(s, o, t) \in T$.

Because the abstraction mappings are orthogonal, $\alpha_i(s) \neq \alpha_i(t)$ for at most one $i \in \{1, \ldots, n\}$.

Case 1: $\alpha_i(s) = \alpha_i(t)$ for all $i \in \{1, \ldots, n\}$.

Then $L = \sum_{i=1}^{n} h^{\alpha_i(s)} = \sum_{i=1}^{n} h^{\alpha_i(t)} = \sum_{i=1}^{n} h^{\alpha_i(t)} = R \leq R + 1$. 

abstraction $\mathcal{A}_1$

mapping: only consider state of first package

abstraction $\mathcal{A}_2$ (orthogonal to $\mathcal{A}_1$)

mapping: only consider state of second package
Abstractions of abstractions

**Theorem (transitivity of abstractions)**

Let $\tau$, $\tau'$ and $\tau''$ be transition systems.

- If $\tau'$ is an abstraction of $\tau$ and $\tau''$ is an abstraction of $\tau'$, then $\tau''$ is an abstraction of $\tau$.
- If $\tau'$ is a strictly homomorphic abstraction of $\tau$ and $\tau''$ is a strictly homomorphic abstraction of $\tau'$, then $\tau''$ is a strictly homomorphic abstraction of $\tau$.

Proof (ctd.)

**Case 2: $\alpha_i(s) \neq \alpha_i(t)$ for exactly one $i \in \{1, \ldots, n\}$.**

Let $k \in \{1, \ldots, n\}$ such that $\alpha_k(s) \neq \alpha_k(t)$.

Then $L = \sum_{i=1}^n h^{\alpha_i} \alpha_i(s)$

\[
= \sum_{i \in \{1, \ldots, n\} \setminus \{k\}} h^{\alpha_i} \alpha_i(s) + h^{\alpha_k} \alpha_k(s)
\]

\[
\leq \sum_{i \in \{1, \ldots, n\} \setminus \{k\}} h^{\alpha_i} \alpha_i(t) + h^{\alpha_k} \alpha_k(t) + 1
\]

\[
= \sum_{i=1}^n h^{\alpha_i} \alpha_i(t) + 1
\]

where the inequality holds because $\alpha_i(s) = \alpha_i(t)$ for all $i \neq k$ and $h^{\alpha_k} \alpha_k$ is consistent.
Abstractions of abstractions: example

Transition system $\mathcal{T}''$ as an abstraction of $\mathcal{T}'$

November 30th, 2018 B. Nebel, R. Mattmüller – AI Planning 44 / 61

Abstractions of abstractions: proof

Proof.

Let $\mathcal{T} = \langle S, L, T, s_0, S_* \rangle$, let $\mathcal{T}' = \langle S', L, T', s'_0, S'_* \rangle$ be an abstraction of $\mathcal{T}$ with abstraction mapping $\alpha$, and let $\mathcal{T}'' = \langle S'', L, T'', s''_0, S''_* \rangle$ be an abstraction of $\mathcal{T}'$ with abstraction mapping $\alpha'$.

We show that $\mathcal{T}''$ is an abstraction of $\mathcal{T}$ with abstraction mapping $\beta := \alpha' \circ \alpha$, i.e., that

1. $\beta(s_0) = s''_0$,
2. for all $s \in S_*$, we have $\beta(s) \in S''_*$, and
3. for all $\langle s, \ell, t \rangle \in T$, we have $\langle \beta(s), \ell, \beta(t) \rangle \in T''$.

Moreover, we show that if $\alpha$ and $\alpha'$ are strict homomorphisms, then $\beta$ is also a strict homomorphism.

...
Abstractions of abstractions: proof (ctd.)

Proof (ctd.)
2. For all \( s \in S_\alpha \), we have \( \beta(s) \in S'_\alpha \):

Let \( s \in S_\alpha \). Because \( \mathcal{I}' \) is an abstraction of \( \mathcal{I} \) with mapping \( \alpha \), we have \( \alpha(s) \in S'_\alpha \). Because \( \mathcal{I}'' \) is an abstraction of \( \mathcal{I}' \) with mapping \( \alpha' \) and \( \alpha(s) \in S'_\alpha \), we have \( \alpha'(\alpha(s)) \in S''_\alpha \).

Hence \( \beta(s) = \alpha'(\alpha(s)) \in S''_\alpha \).

Strict homomorphism if \( \alpha \) and \( \alpha' \) strict homomorphisms:
Let \( s'' \in S''_\alpha \). Because \( \alpha' \) is a strict homomorphism, there exists a state \( s' \in S_\alpha \) such that \( \alpha'(s') = s'' \). Because \( \alpha \) is a strict homomorphism, there exists a state \( s \in S_\alpha \) such that \( \alpha(s) = s' \).

Thus \( s'' = \alpha'(\alpha(s)) = \beta(s) \) for some \( s \in S_\alpha \).

\( \cdots \)

Abstractions of abstractions: proof (ctd.)

Proof (ctd.)
3. For all \( (s, t) \in T \), we have \( \langle \beta(s), \ell; \beta(t) \rangle \in T'' \):

Let \( (s, t) \in T \). Because \( \mathcal{I}'' \) is an abstraction of \( \mathcal{I} \) with mapping \( \alpha \), we have \( \langle \alpha(s), \ell, \alpha(t) \rangle \in T' \). Because \( \mathcal{I}''' \) is an abstraction of \( \mathcal{I}' \) with mapping \( \alpha' \) and \( \langle \alpha(s), \ell, \alpha(t) \rangle \in T' \), we have \( \langle \alpha'(\alpha(s)), \ell, \alpha'(\alpha(t)) \rangle \in T'' \).

Hence \( \langle \beta(s), \ell, \beta(t) \rangle = \langle \alpha'(\alpha(s)), \ell, \alpha'(\alpha(t)) \rangle \in T'' \).

Strict homomorphism if \( \alpha \) and \( \alpha' \) strict homomorphisms:
Let \( (s'', t'' \in T'' \). Because \( \alpha' \) is a strict homomorphism, there exists a transition \( (s', t') \in T'' \) such that \( \alpha'(s') = s'' \) and \( \alpha'(t') = t'' \). Because \( \alpha \) is a strict homomorphism, there exists a transition \( (s, t) \in T \) such that \( \alpha(s) = s' \) and \( \alpha(t) = t' \).

Thus \( (s'', t'' \in T'' \). Because \( \alpha' \) is a strict homomorphism, there exists a transition \( (s, t) \in T \) such that \( \alpha(s) = s' \) and \( \alpha(t) = t' \).

\( \Box \)

Coarsenings and refinements

Terminology: Let \( \mathcal{I} \) be a transition system, let \( \mathcal{I}' \) be an abstraction of \( \mathcal{I} \) with abstraction mapping \( \alpha \), and let \( \mathcal{I}'' \) be an abstraction of \( \mathcal{I}' \) with abstraction mapping \( \alpha' \).

Then:
1. \( \langle \mathcal{I}', \alpha' \circ \alpha \rangle \) is called a coarsening of \( \langle \mathcal{I}', \alpha \rangle \), and
2. \( \langle \mathcal{I}', \alpha \rangle \) is called a refinement of \( \langle \mathcal{I}'', \alpha' \circ \alpha \rangle \).

Abstractions: informally

Heuristic quality of refinements

Theorem (heuristic quality of refinements)

Let \( h^{\mathcal{A},\alpha} \) and \( h^{\mathcal{B},\beta} \) be abstraction heuristics for the same planning task \( \Pi \) such that \( \langle \mathcal{A}, \alpha \rangle \) is a refinement of \( \langle \mathcal{B}, \beta \rangle \). Then \( h^{\mathcal{A},\alpha} \) dominates \( h^{\mathcal{B},\beta} \).

In other words, \( h^{\mathcal{A},\alpha}(s) \geq h^{\mathcal{B},\beta}(s) \) for all states \( s \in \Pi \).

Proof.

Since \( \langle \mathcal{A}, \alpha \rangle \) is a refinement of \( \langle \mathcal{B}, \beta \rangle \), there exists a mapping \( \alpha' \) such that \( \beta = \alpha' \circ \alpha \) and \( \mathcal{B} \) is an abstraction of \( \mathcal{A} \) with abstraction mapping \( \alpha' \).

For any state \( s \in \Pi \), we get \( h^{\mathcal{B},\beta}(s) = h^{\mathcal{B}}(\beta(s)) = h^{\mathcal{B}}(\alpha'(\alpha(s))) = h^{\mathcal{A},\alpha'}(\alpha(s)) \leq h^{\mathcal{A},\alpha}(s) = h^{\mathcal{B}}(\alpha(s)) \), where the inequality holds because \( h^{\mathcal{A},\alpha} \) is an admissible heuristic in the transition system \( \mathcal{A} \).
Isomorphic transition systems

Definition (isomorphic transition systems)
Let $\mathcal{I} = \langle S, L, T, s_0, S_\star \rangle$ and $\mathcal{I}' = \langle S', L', T', s'_0, S'_\star \rangle$ be transition systems.
We say that $\mathcal{I}$ is isomorphic to $\mathcal{I}'$, in symbols $\mathcal{I} \sim \mathcal{I}'$, if there exist bijective functions $\varphi : S \rightarrow S'$ and $\psi : L \rightarrow L'$ such that:

- $\varphi(s_0) = s'_0$, $s \in S$ iff $\varphi(s) \in S'$, and
- $\langle s, \ell, t \rangle \in T$ iff $\langle \varphi(s), \psi(\ell), \varphi(t) \rangle \in T'$.

November 30th, 2018
B. Nebel, R. Mattmüller – AI Planning 51 / 61

Graph-equivalent transition systems

Definition (graph-equivalent transition systems)
Let $\mathcal{I} = \langle S, L, T, s_0, S_\star \rangle$ and $\mathcal{I}' = \langle S', L', T', s'_0, S'_\star \rangle$ be transition systems.
We say that $\mathcal{I}$ is graph-equivalent to $\mathcal{I}'$, in symbols $\mathcal{I} \cong \mathcal{I}'$, if there exists a bijective function $\varphi : S \rightarrow S'$ such that:

- $\varphi(s_0) = s'_0$, $s \in S$ iff $\varphi(s) \in S'$, and
- $\langle s, \ell, t \rangle \in T$ for some $\ell \in L$ iff $\langle \varphi(s), \ell', \varphi(t) \rangle \in T'$ for some $\ell' \in L'$.

Note: There is no requirement that the labels of $\mathcal{I}$ and $\mathcal{I}'$ correspond in any way. For example, it is permitted that all transitions of $\mathcal{I}$ have different labels and all transitions of $\mathcal{I}'$ have the same label.

November 30th, 2018
B. Nebel, R. Mattmüller – AI Planning 52 / 61

Isomorphism vs. graph equivalence

- $(\sim)$ and $(\cong)$ are equivalence relations.
- Two isomorphic transition systems are interchangeable for all practical intents and purposes.
- Two graph-equivalent transition systems are interchangeable for most intents and purposes.
  In particular, their state distances are identical, so they define the same abstraction heuristic for corresponding abstraction functions.
  Isomorphism implies graph equivalence, but not vice versa.

November 30th, 2018
B. Nebel, R. Mattmüller – AI Planning 53 / 61

Using abstraction heuristics in practice

In practice, there are conflicting goals for abstractions:
- we want to obtain an informative heuristic, but
  want to keep its representation small.
  Abstractions have small representations if they have
  - few abstract states and
    a succinct encoding for $\alpha$.

November 30th, 2018
B. Nebel, R. Mattmüller – AI Planning 54 / 61
Counterexample: one-state abstraction

One-state abstraction: \( \alpha(s) := \text{const.} \)
+ very few abstract states and succinct encoding for \( \alpha \)
− completely uninformative heuristic

Counterexample: identity abstraction

Identity abstraction: \( \alpha(s) := s \).
+ perfect heuristic and succinct encoding for \( \alpha \)
− too many abstract states

Counterexample: perfect abstraction

Perfect abstraction: \( \alpha(s) := h^*(s) \).
+ perfect heuristic and usually few abstract states
− usually no succinct encoding for \( \alpha \)

Automatically deriving good abstraction heuristics

Abstraction heuristics for planning: main research problem
Automatically derive effective abstraction heuristics for planning tasks.

we will study one state-of-the-art approach in the next chapter.
An abstraction relates a transition system $T$ (e.g. of a planning task) to another (usually smaller) transition system $T'$ via an abstraction mapping $\alpha$.

Abstraction preserves all important aspects of $T$: initial state, goal states and (labeled) transitions. Hence, they can be used to define heuristics for the original system $T$: estimate the goal distance of $s$ in $T$ by the optimal goal distance of $\alpha(s)$ in $T'$.

Such abstraction heuristics are safe, goal-aware, admissible and consistent.

Strictly homomorphic abstractions are desirable as they do not include “unnecessary” abstract goal states or transitions (which could lower heuristic values).

Any surjection from the states of $T$ to any set induces a strictly homomorphic abstraction in a natural way.

Multiple abstraction heuristics can be added without losing properties like admissibility if the underlying abstraction mappings are orthogonal.

One sufficient condition for orthogonality is that abstractions are affected by disjoint sets of labels.

The process of abstraction is transitive: an abstraction can be abstracted further to yield another abstraction. Based on this notion, we can define abstractions that are coarsenings or refinements of others. A refinement can never lead to a worse heuristic. Practically useful abstractions are those which give informative heuristics, yet have a small representation.