Tutorial on Abstraction Heuristics

Daniel Fišer

Department of Computer Science,
Faculty of Electrical Engineering,
danfis@danfis.cz

March 29, 2020

Abstraction Heuristics

Before you proceed with this tutorial, read Section 4 of the tutorial notes on classical planning https://cw.fel.cvut.cz/wiki/_media/courses/be4m36pui/notes-cp.pdf

Example 1

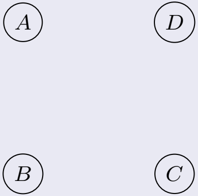
Compute the synchronized product of the following two transition systems $\mathcal{T}^1 = \langle S^1, L, T^1, I^1, G^1 \rangle$ and $\mathcal{T}^2 = \langle S^2, L, T^2, I^2, G^2 \rangle$, where

- $L = \{a, b, c, d, e\}$,
- $S^1 = \{A, B, C, D\}$,
- $T^1 = \{(A, a, B), (B, b, C), (C, c, A), (A, d, A), (A, e, D)\}$,
- $I^1 = \{A\}$, $G^1 = \{A, C\}$,
- $S^2 = \{X, Y, Z\}$,
- $T^2 = \{(X, a, Y), (X, a, Z), (Y, b, Z), (Z, c, Y), (Z, d, Y), (Z, e, Z)\}$,
- $I^2 = \{X\}$, and $G^2 = \{X\}$.

Note that both transition systems has the same set of labels. This is required in order to compute the synchronized product.

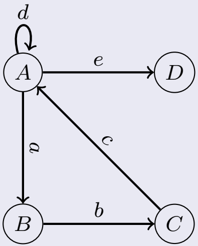
Example 1

Transition systems can be depicted as graphs where vertices correspond to states and edges correspond to transitions.
So for states $S^1 = \{A, B, C, D\}$, we get the following vertices.



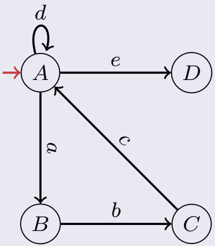
Example 1

Transition systems can be depicted as graphs where vertices correspond to states and edges correspond to transitions.
So for states $S^1 = \{A, B, C, D\}$, we get the following vertices.
For transitions $T^1 = \{(A, a, B), (B, b, C), (C, c, A), (A, d, A), (A, e, D)\}$, we get the following directed labeled edges.



Example 1

Transition systems can be depicted as graphs where vertices correspond to states and edges correspond to transitions.
So for states $S^1 = \{A, B, C, D\}$, we get the following vertices.
For transitions $T^1 = \{(A, a, B), (B, b, C), (C, c, A), (A, d, A), (A, e, D)\}$, we get the following directed labeled edges.
We mark the initial state A.



Example 1

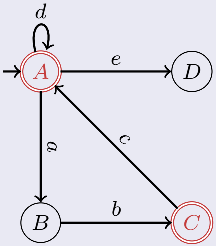
Transition systems can be depicted as graphs where vertices correspond to states and edges correspond to transitions.

So for states $S^1 = \{A, B, C, D\}$, we get the following vertices.

For transitions $T^1 = \{(A, a, B), (B, b, C), (C, c, A), (A, d, A), (A, e, D)\}$, we get the following directed labeled edges.

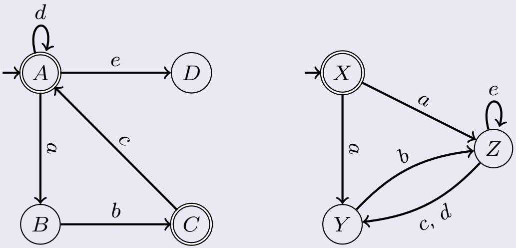
We mark the initial state A.

And goal states $G^1 = \{A, C\}$.



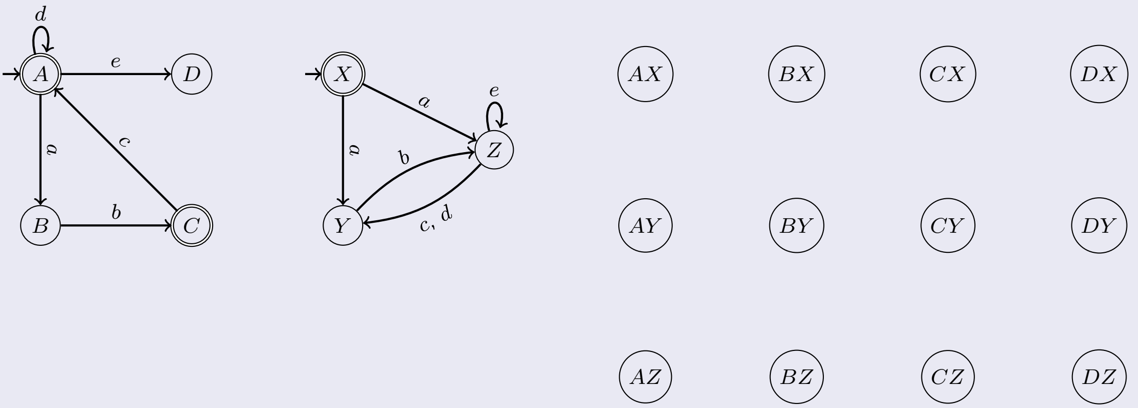
Example 1

Similarly, we construct the second transition system T^2.



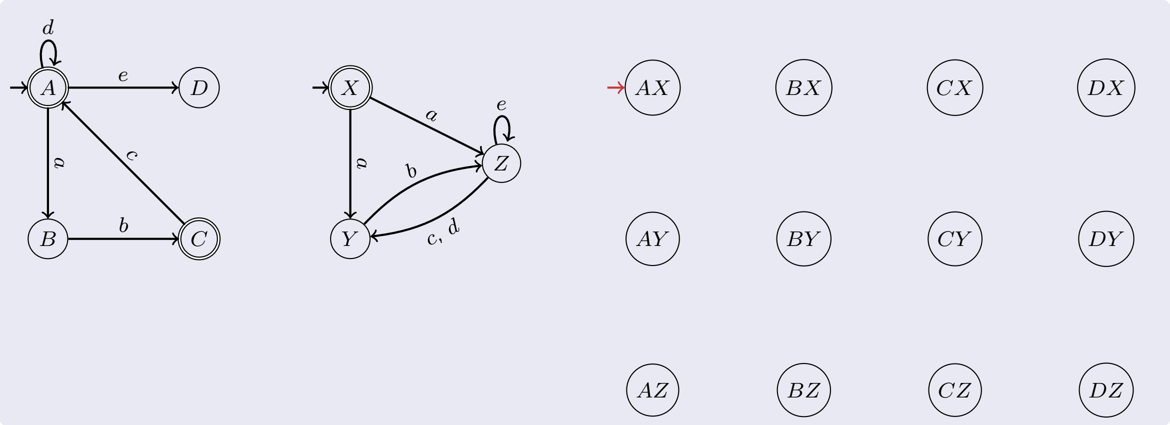
Example 1

The most straighforward way to compute the synchronized product $T^1 \otimes T^2$ is the following. We start with constructing the states of the product: this is the cartesian product of states S^1 and S^2.

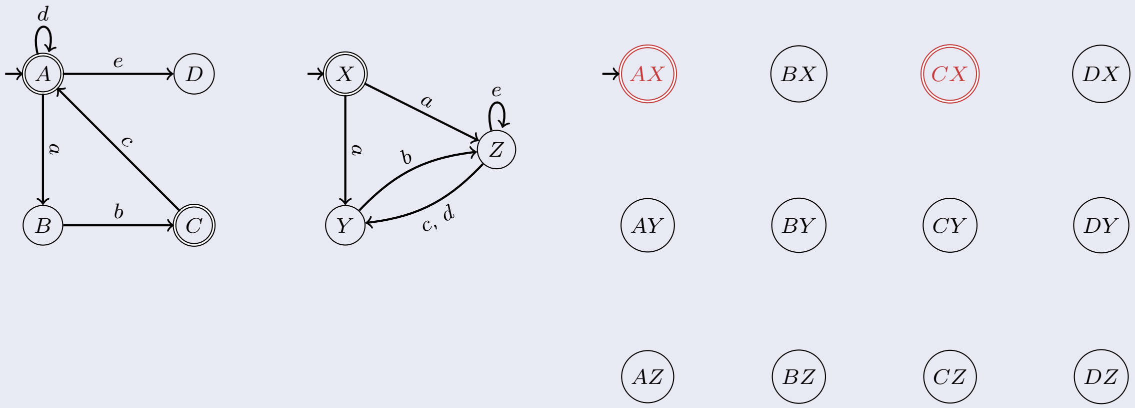


Example 1

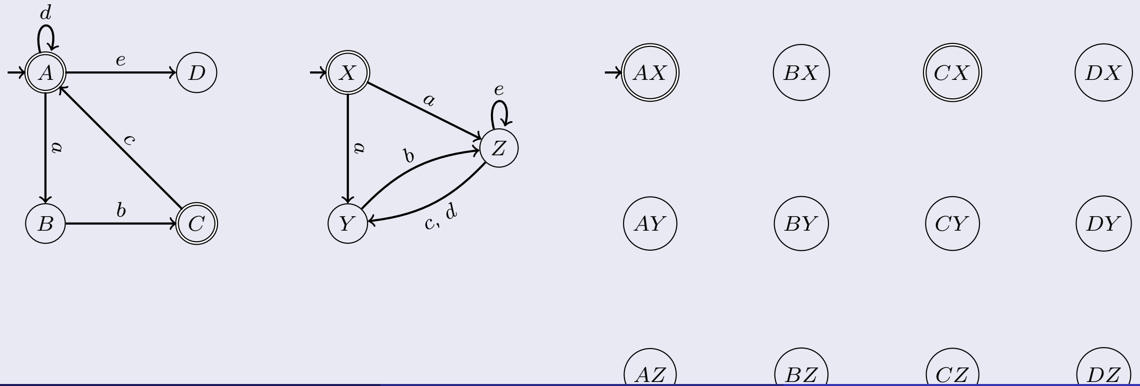
Now, we identify the initial state: it the cartesian product of I^1 and I^2.



Now, we identify the initial state: it the cartesian product of I^1 and I^2.
And the goal states: the cartesian product of G^1 and G^2.

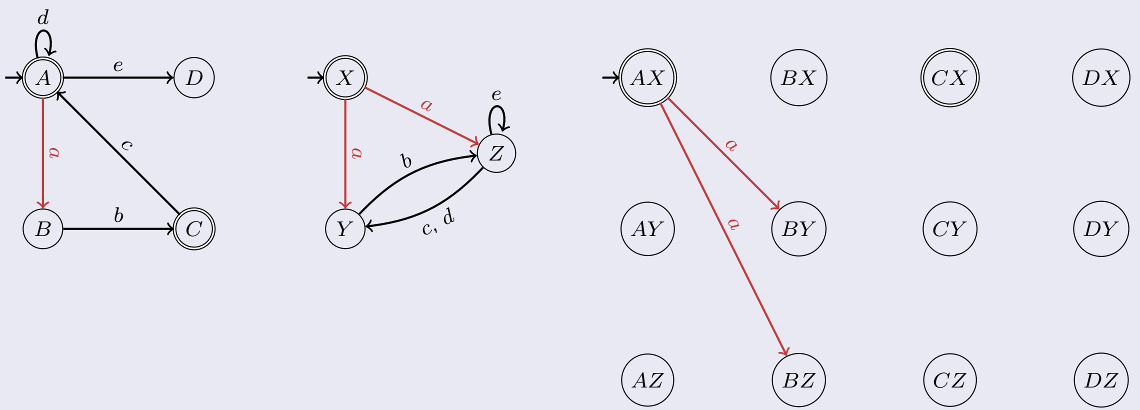


Now we go over all labels L and construct the transitions in the synchronized product: we will have a transition between two states if there was a transition with the same label in both transition systems T^1 and T^2: $T = \{((s_1, s_2), l, (t_1, t_2)) \mid (s_1, l, s_2) \in T^1, (s_2, l, t_2) \in T^2\}$.



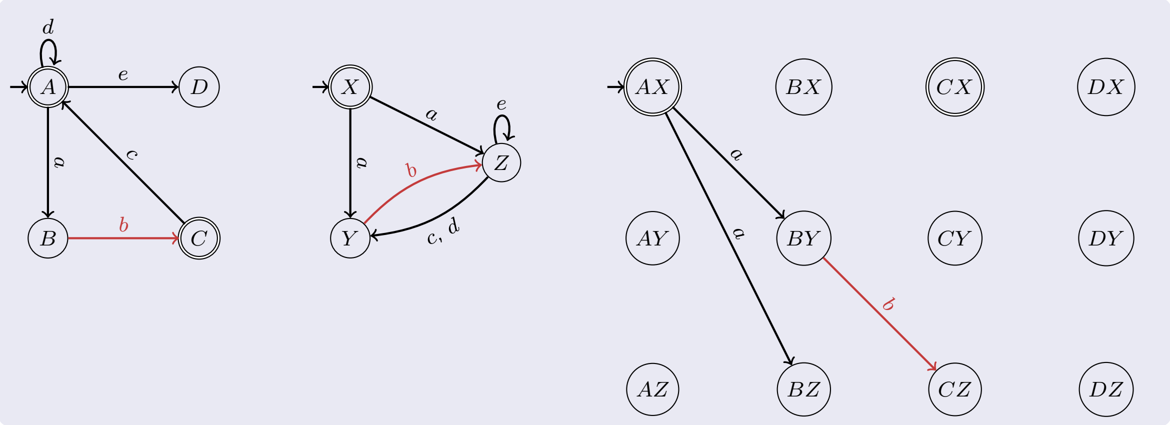
Example 1

So, for the label a, we get the following transitions.



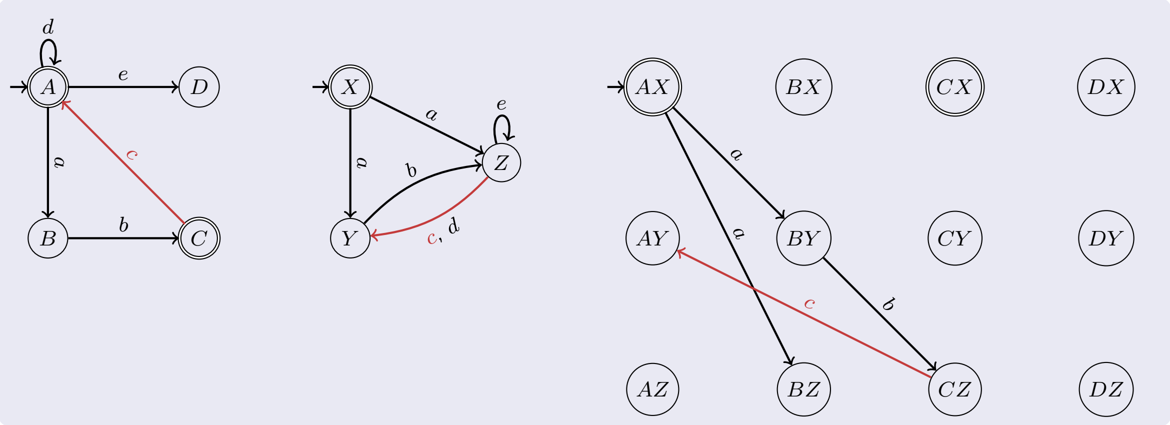
Example 1

For the label b...



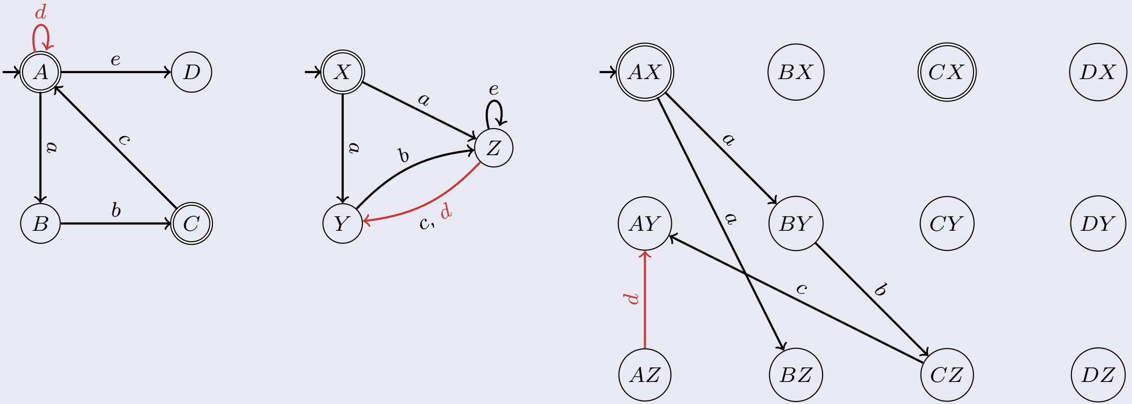
Example 1

For the label c...



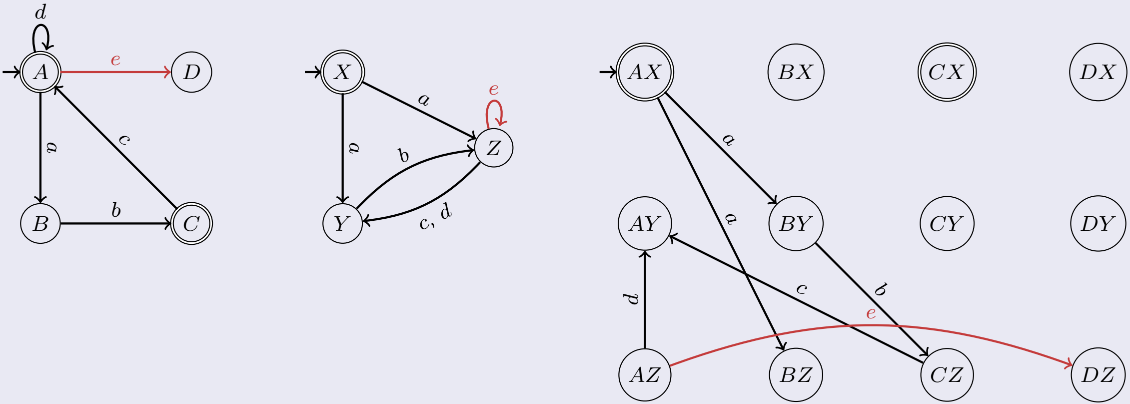
Example 1

For the label d...

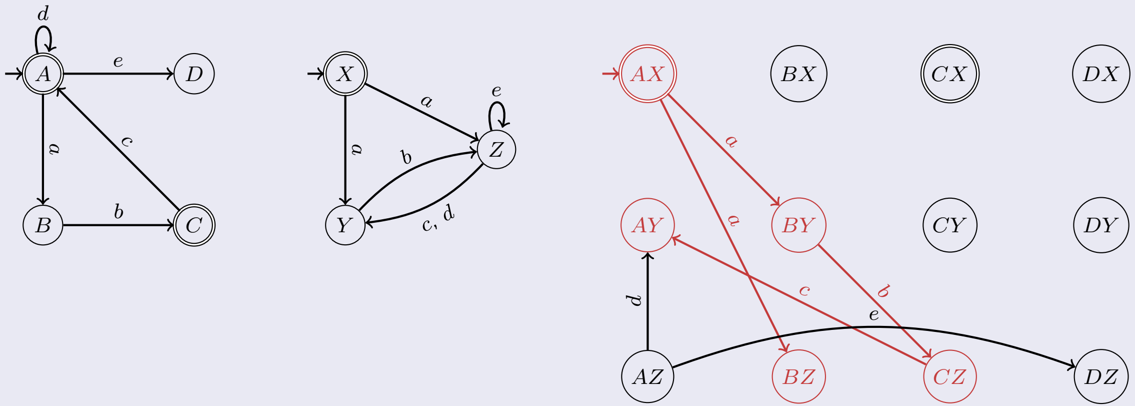


Example 1

For the label e...



Now note that the only interesting part of the resulting transition system is the part that is reachable from the initial state.



Example 2

Now we compute merge&shrink heuristic for the example planning task from our first tutorial (depicted on Figure 1 in the notes).
We will not use any particular shrink or merge strategies (there are many), but rather demonstrate the principle behind merging (i.e., computing synchronized products of abstractions) and shrinking (i.e., further abstracting of abstractions).

Example 2

Recall that we modeled our example planning task in FDR as follows. We have a planning task $\mathcal{P} = \langle \mathcal{V}, \mathcal{O}, s_{\text{init}}, s_{\text{goal}} \rangle$ with:

- Variables $\mathcal{V} = \{b, t, p\}$ (position of the boat b, the truck t, and the package p).
- with domains $D_b = \{A, B\}$, $D_t = \{B, C\}$, $D_p = \{A, B, C, b, t\}$.
- The initial state is $s_{\text{init}} = \{b{=}A, t{=}C, p{=}A\}$.
- The goal is $s_{\text{goal}} = \{p{=}C\}$.
- And we have operators for moving boat and truck, and for loading and unloading package. We will abbreviate the operators:
  - MbAB, MbBA for moving the boat between A and B,
  - MtBC, MtCB for moving the truck between B and C,
  - LbA, LbB, UbA, UbB for loading and unloading the package from and to the boat in the city A or B,
  - and LtB, LtC, UtB, UtC for loading and unloading the truck.

Example 2

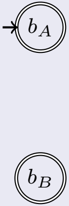
Also recall from the lecture, that we want to create an abstract transition system T with an abstraction function α. And for a given state s we will take $\mathrm{h}^{\star}(\alpha(s))$ in T (i.e., the cost of the cheapest path in T from $\alpha(s)$ to the nearest abstract goal) as the heuristic estimate for our original problem.

Example 2

We start with atomic projections to each invidivual variable. We will abbreviate assignements to variables with subscripts, so for example $b{=}A$ will be written as b_A or $p{=}b$ as p_b.

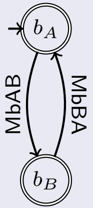
Example 2

So, the projection to the variable b has only two abstract states. The state b_A is the initial abstract state, and both of the states are goal states, because the goal does not mention the position of the boat.



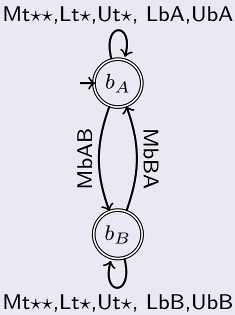
Example 2

So, the projection to the variable b has only two abstract states. The state b_A is the initial abstract state, and both of the states are goal states, because the goal does not mention the position of the boat.
The movement of the boat corresponds to the actions MbAB and MbBA that change b_A to b_B and vice versa.



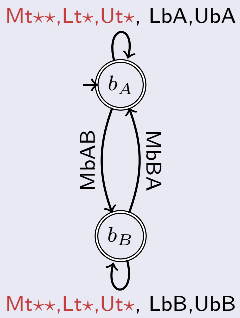
Example 2

Now it is important to remember that an abstraction must preserve all transitions. So the rest of the operators must be in the loops. (We use a similar notation that is used in the lecture slides.)



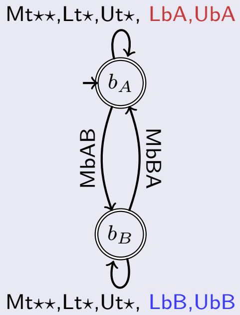
Example 2

Now it is important to remember that an abstraction must preserve all transitions. So the rest of the operators must be in the loops. (We use a similar notation that is used in the lecture slides.)
Note that the truck can be moved, loaded or unloaded regardless of the position of the boat.



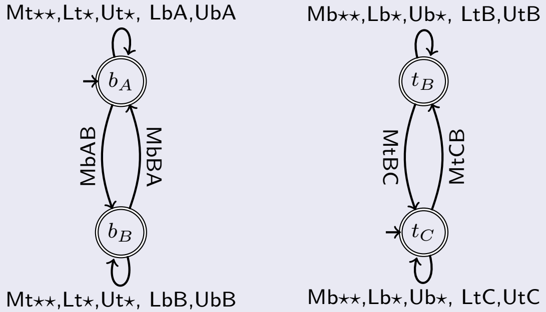
Example 2

Now it is important to remember that an abstraction must preserve all transitions. So the rest of the operators must be in the loops. (We use a similar notation that is used in the lecture slides.)
But the boat can be loaded in city A only if it is positioned in the city A and similarly for the city B.



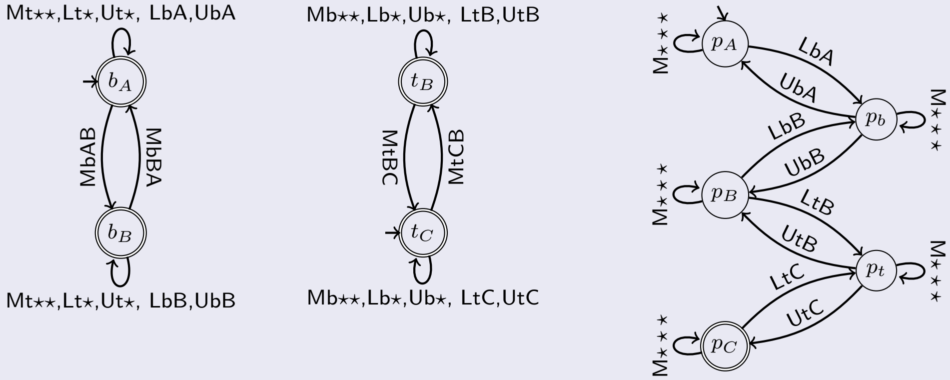
Example 2

Similarly, we can construct the projection to the variable t (note where are goal states and where is the initial state).



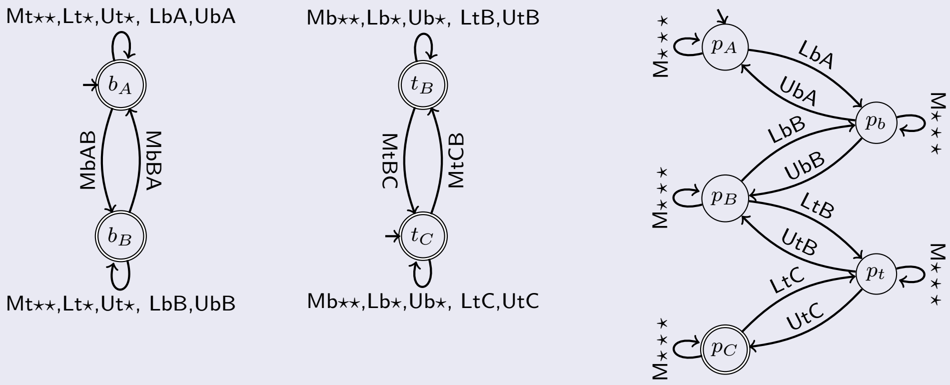
Example 2

And for the variable p...



Example 2

And for the variable p...
Note that the movement of both boat and truck is independent of the assignment to the variable p.

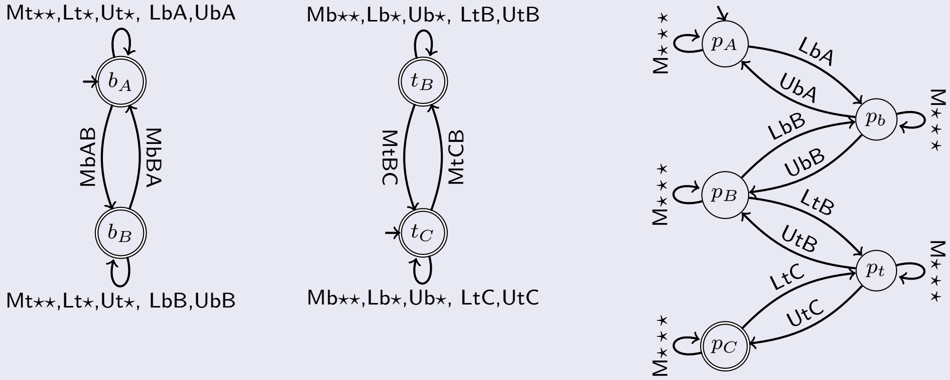


Example 2

And for the variable p...

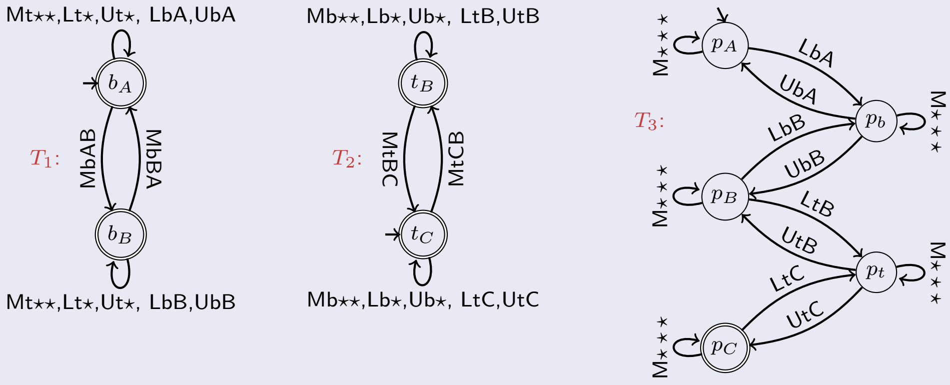
Note that the movement of both boat and truck is independent of the assignment to the variable p.

Also note, that here we have only one goal state.



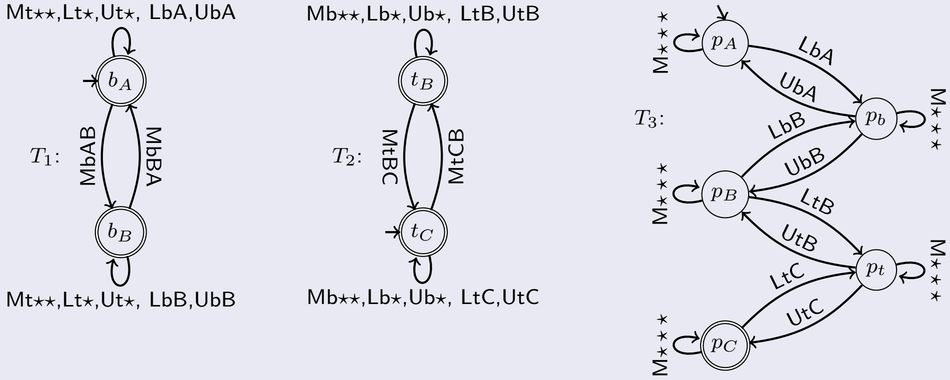
Example 2

Now we have a set of three atomic abstractions, T_1, T_2, T_3 (line 1 of Algorithm 3), and we can proceed with the computation of one abstraction of the planning task.



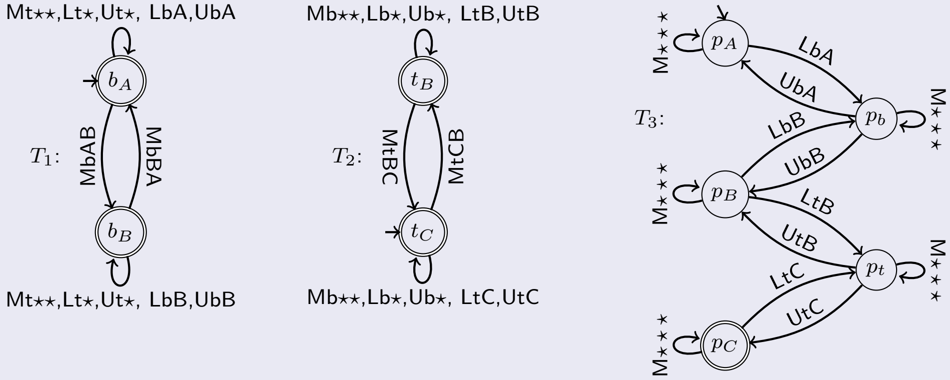
Example 2

Next, we need to choose two abstractions (line 3). (As was already said, in this tutorial, we will not follow any particular strategy in any step, but rather use intuition to demonstrate how merge&shrink works.)



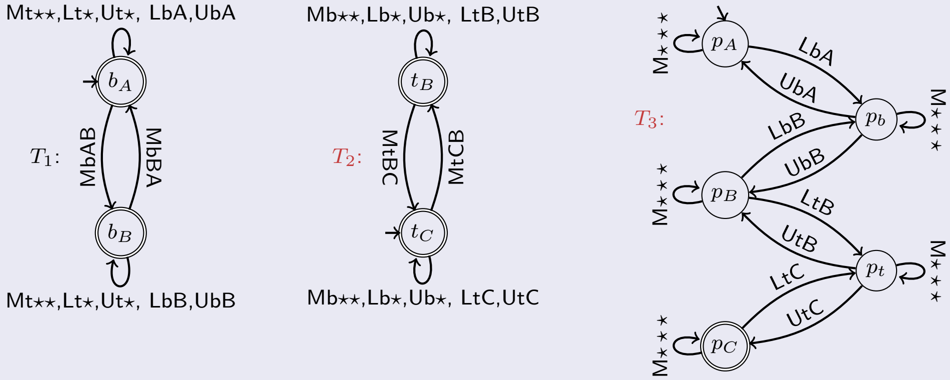
Example 2

If we choose T_1 and T_2 and we compute the synchronized product, we get an abstraction describing position of both boat and truck (assuming we do not shrink anything, because there is not much to shrink).



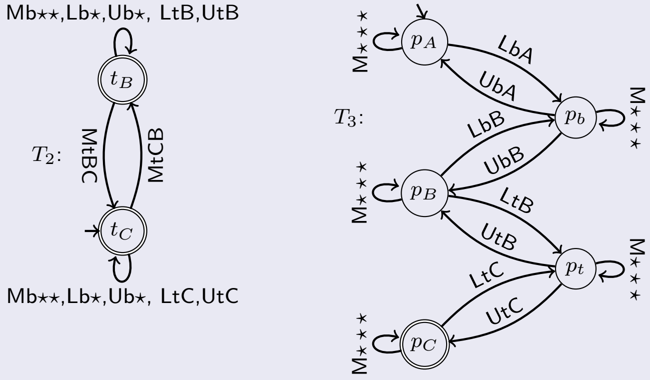
Example 2

If we choose T_1 and T_2 and we compute the synchronized product, we get an abstraction describing position of both boat and truck (assuming we do not shrink anything, because there is not much to shrink).
This is boring, so let's try to select T_2 and T_3.



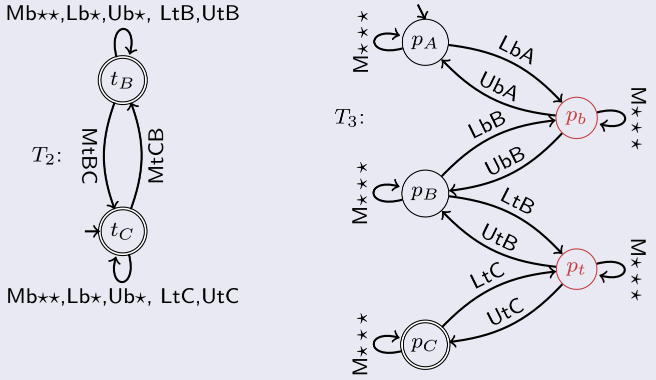
Example 2

Now we have to decide whether to shrink these abstractions. Shrinking means further abstracting.



Example 2

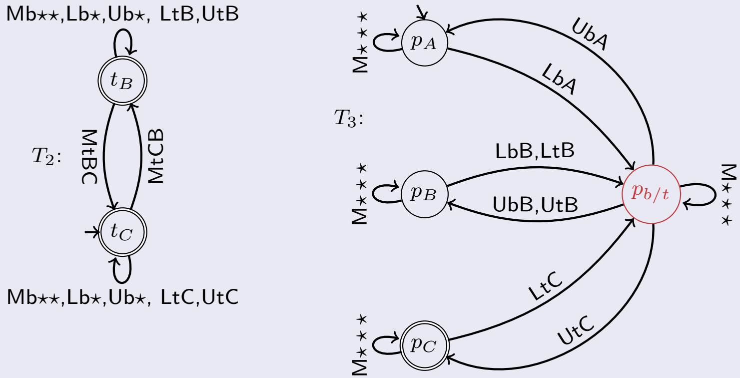
Now we have to decide whether to shrink these abstractions. Shrinking means further abstracting.
So for example, we can choose to shrink T_3 and decide not to distinguish whether the package is loaded in the boat or in the truck. (Note again, that we need to preserve all transitions.)



Example 2

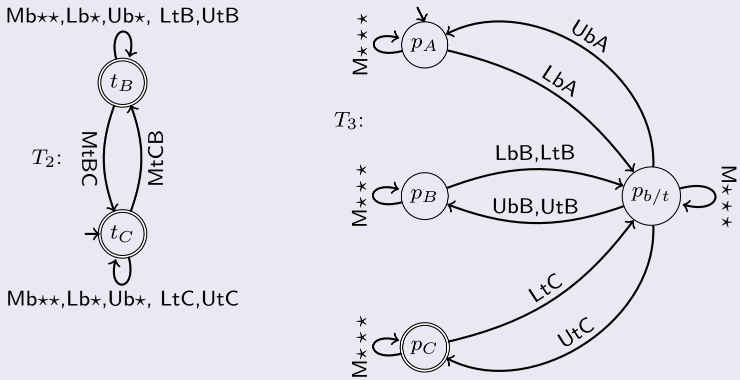
Now we have to decide whether to shrink these abstractions. Shrinking means further abstracting.

So for example, we can choose to shrink T_3 and decide not to distinguish whether the package is loaded in the boat or in the truck. (Note again, that we need to preserve all transitions.)



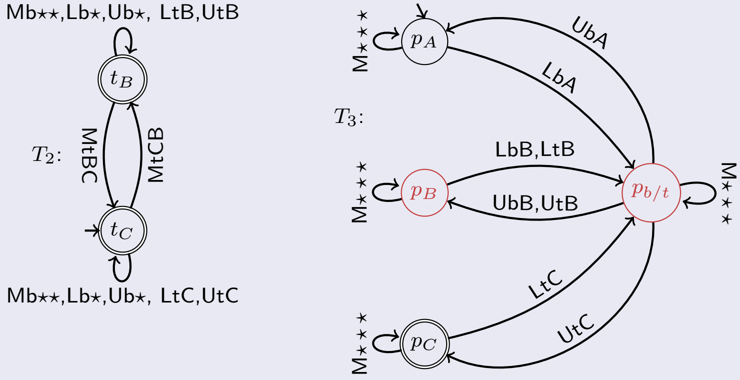
Example 2

Now we have two states in T_2 and four in T_3, so the resulting merge (synchronized product) would have eight states. So, let's try to shrink T_3 even more.



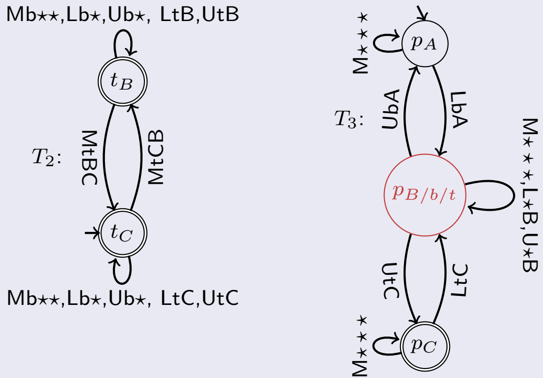
Example 2

Now, we shrink T_3 by abstracting states p_B and $p_{b/t}$ into one abstract state.



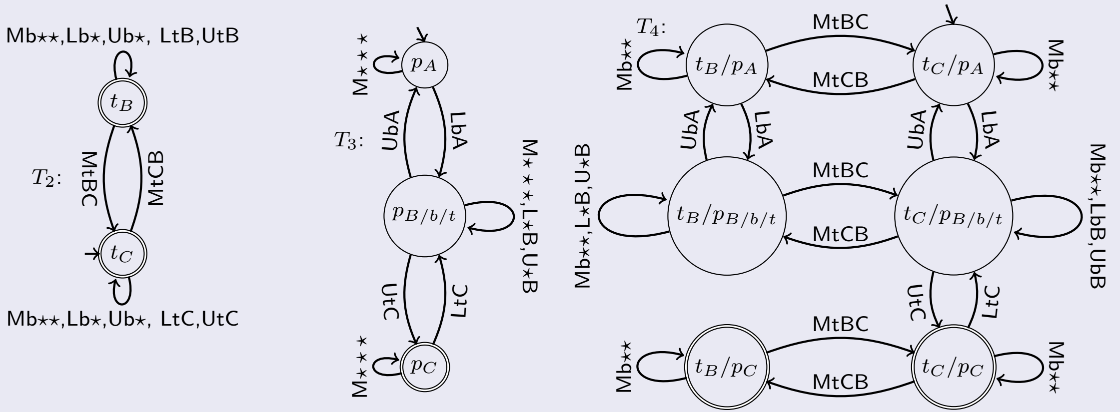
Example 2

Now, we shrink T_3 by abstracting states p_B and $p_{b/t}$ into one abstract state.



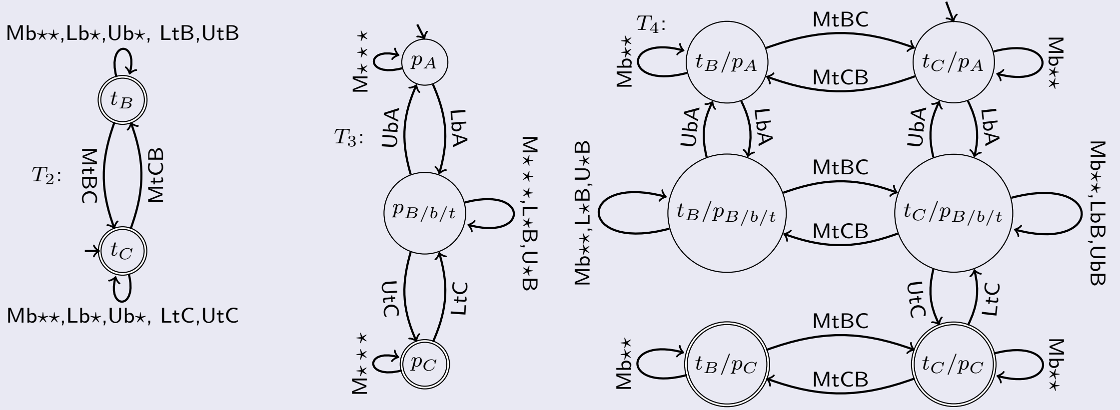
Example 2

Now, we can replace T_2 and T_3 with the merge (synchronized product) $T_4 = T_2 \otimes T_3$ (line 5).



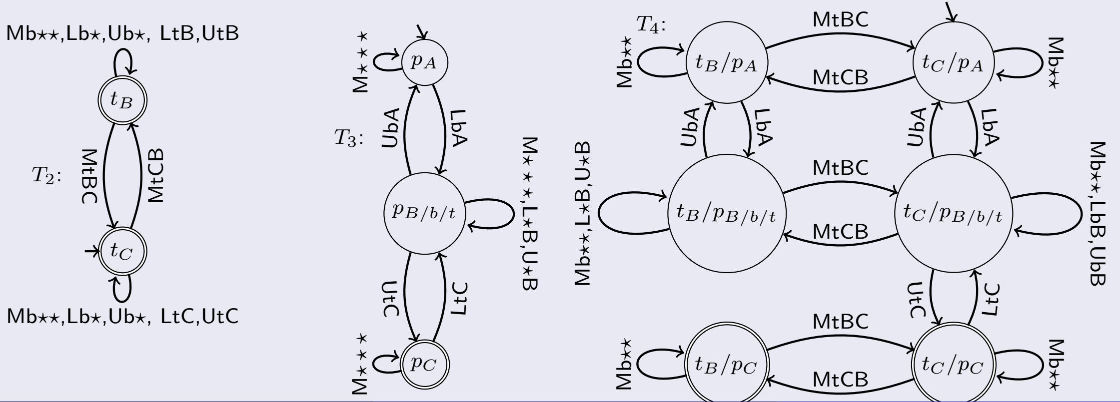
Example 2

Find the intial state and goal states in T_4 and make sure you understand why these states.



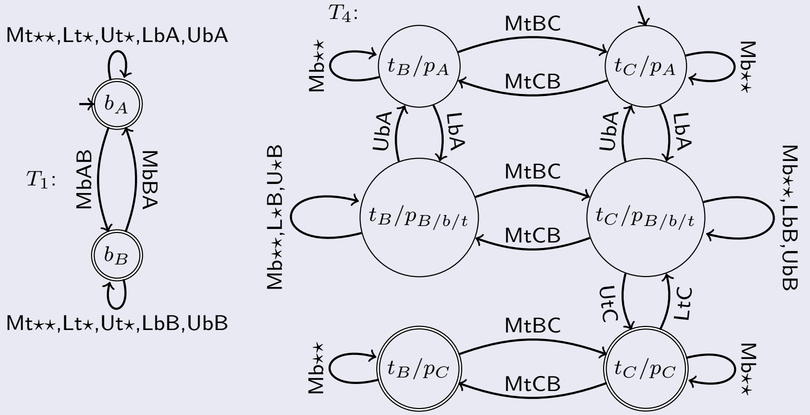
Example 2

Go over each transition in T_4 and make sure you understand why the transition is there.
Why T_4 does not have any transition between $t_B/p_{B/b/t}$ and t_B/p_C?
Why $t_B/p_{B/b/t}$ has LtB, UtB, LbB, and UbB on the loop, but $t_C/p_{B/b/t}$ has only LbB and UbB?



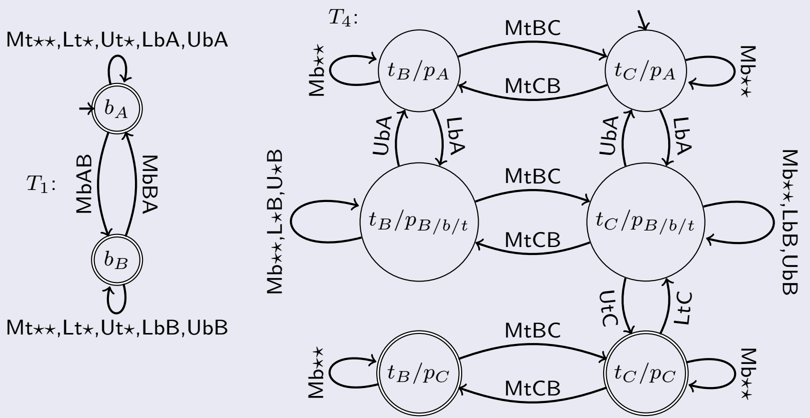
Example 2

Now we have only two abstractions T_1 and T_4 so we have to choose these two (line 3).

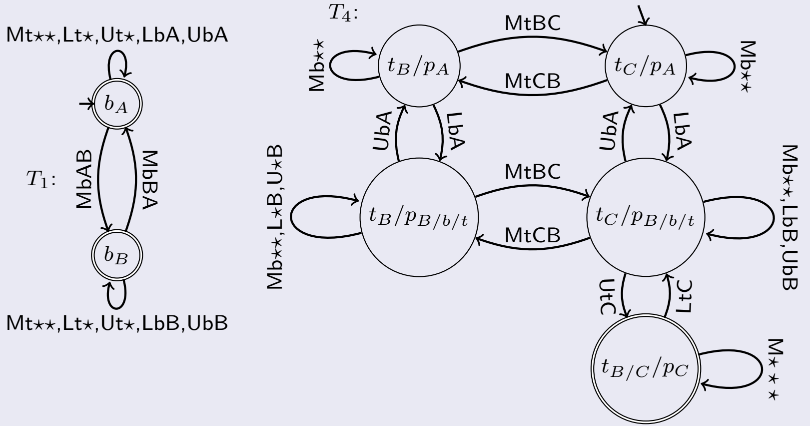


Example 2

We keep T_1 without shrinking, but we will shrink T_4 because it has too many states.

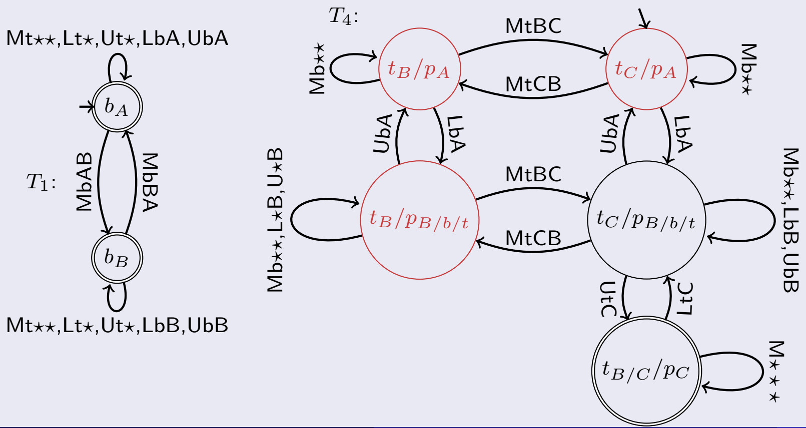


First, let's combine both goal states into one state.

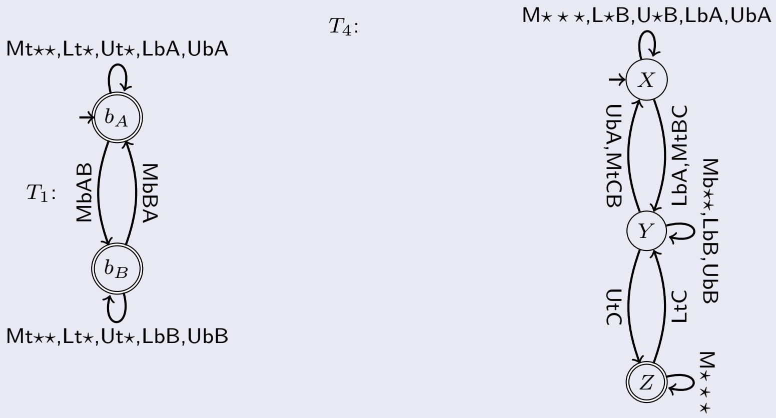


Example 2

Next, we can try to combine three states t_C/p_A, t_B/p_A, and $t_B/p_{B/b/t}$ into one state. And we rename this state to X so it fits on the slide, and we also rename $t_C/p_{B/b/t}$ to Y, and $t_{B/C}/p_C$ to Z.

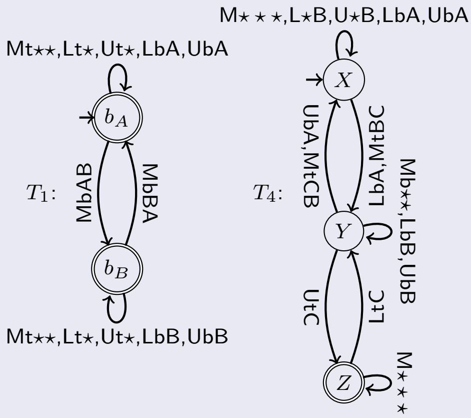


Next, we can try to combine three states t_C/p_A, t_B/p_A, and $t_B/p_{B/b/t}$ into one state. And we rename this state to X so it fits on the slide, and we also rename $t_C/p_{B/b/t}$ to Y, and $t_{B/C}/p_C$ to Z.



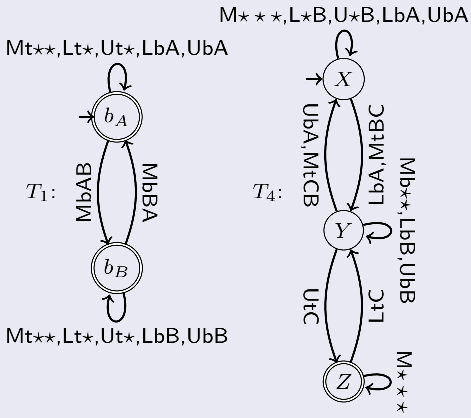
Example 2

Now make sure you understand how we got all labels, especially in the loop above X and between X and Y.

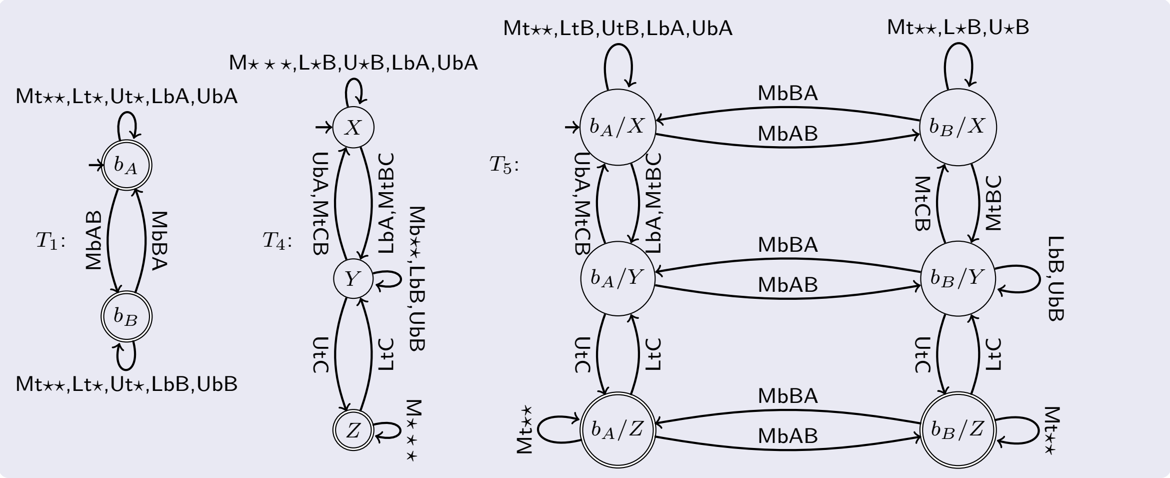


Example 2

Next, we compute the synchronized product $T_5 = T_1 \otimes T_4$.



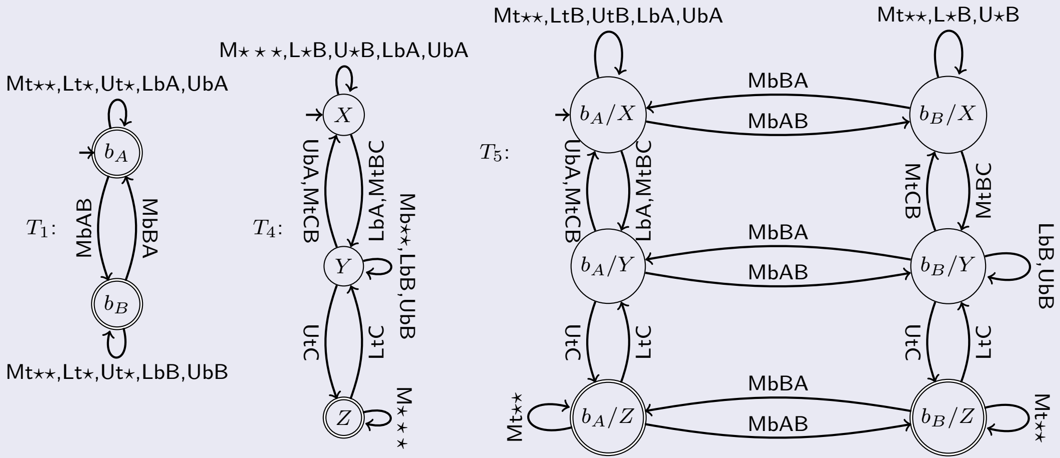
Example 2



T_1:
Mt⋆⋆,Lt⋆,Ut⋆,LbA,UbA
b_A
MbAB
MbBA
b_B
Mt⋆⋆,Lt⋆,Ut⋆,LbB,UbB
T_4:
M⋆ ⋆ ⋆,L⋆B,U⋆B,LbA,UbA
X
UbA,MtCB
LbA,MtBC
Y
Mb⋆⋆,LbB,UbB
UtC
LtC
Z
M⋆⋆⋆
T_5:
Mt⋆⋆,LtB,UtB,LbA,UbA
Mt⋆⋆,L⋆B,U⋆B
b_A/X
MbBA
MbAB
b_B/X
UbA,MtCB
LbA,MtBC
MtCB
MtBC
b_A/Y
MbBA
MbAB
b_B/Y
LbB,UbB
UtC
LtC
UtC
LtC
Mt⋆⋆
b_A/Z
MbBA
MbAB
b_B/Z
Mt⋆⋆

Example 2

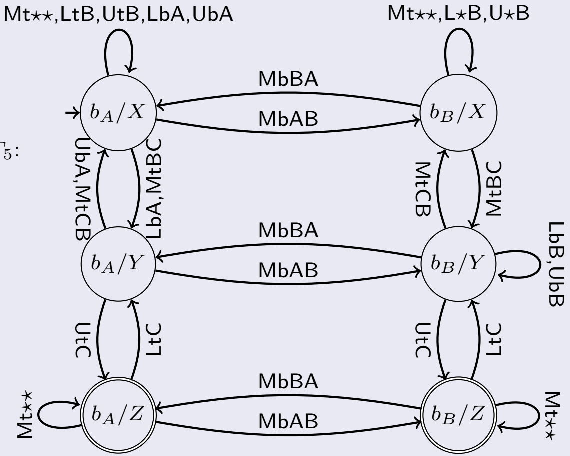
Again, make sure you understand how we labeled edges in T_5.



Example 2

Now we have only one abstraction, so we can terminate Algorithm 3 with T_5 being the resulting abstraction.

T_5:



Example 2

Assume that the cost of all operators is one.
Answer the following questions (recall that X corresponds to the state combining t_C/p_A, t_B/p_A, and $t_B/p_{B/b/t}$, Y corresponds to $t_C/p_{B/b/t}$, and Z corresponds to $t_{B/C}/p_C$):

- What is the heuristic value for the initial state ($h^{\mathrm{m\&s}}(s_{\mathrm{init}})$)?
- What is the heuristic value for the state $s = \{b{=}B, t{=}B, p{=}A\}$?
- What is the heuristic value for the state $s = \{b{=}B, t{=}C, p{=}B\}$?
- Go back and try to find better selection (line 3) and shrinking (line 4) strategies that will give us higher heuristic value for the initial state.

Example 2

Assume that the cost of all operators is one.

Answer the following questions (recall that X corresponds to the state combining t_C/p_A, t_B/p_A, and $t_B/p_{B/b/t}$, Y corresponds to $t_C/p_{B/b/t}$, and Z corresponds to $t_{B/C}/p_C$):

- What is the heuristic value for the initial state ($\mathrm{h}^{\mathrm{m\&s}}(s_{\mathrm{init}})$)? Answer: $\mathrm{h}^{\mathrm{m\&s}}(s_{\mathrm{init}}) = 2$
- What is the heuristic value for the state $s = \{b{=}B, t{=}B, p{=}A\}$? Answer: $\mathrm{h}^{\mathrm{m\&s}}(s) = 2$
- What is the heuristic value for the state $s = \{b{=}B, t{=}C, p{=}B\}$? Answer: $\mathrm{h}^{\mathrm{m\&s}}(s) = 1$
- Go back and try to find better selection (line 3) and shrinking (line 4) strategies that will give us higher heuristic value for the initial state.