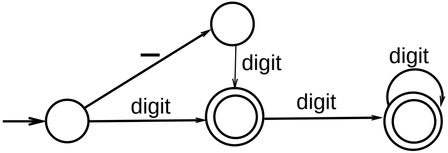
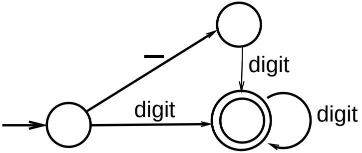
MINIMIZING THE NUMBER OF STATES

- The DFA constructed by the procedure mentioned earlier does not result in a minimum DFA.
- The DFA that we had seen earlier:

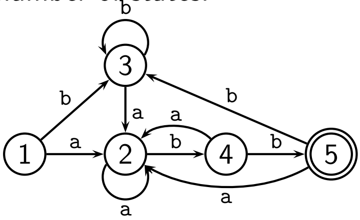


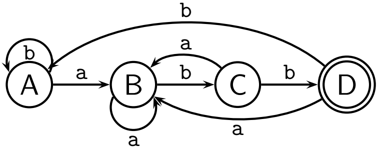
is not minimum. There is another DFA for the same regular expression with lesser number of states.



MINIMIZING THE NUMBER OF STATES

- The DFA constructed for $(b|\epsilon)(a|b)^*abb$.
- There is another DFA for the same regular expression with lesser number of states.



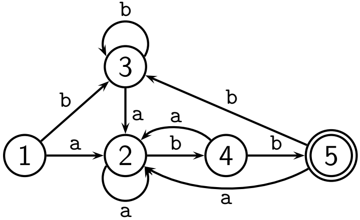


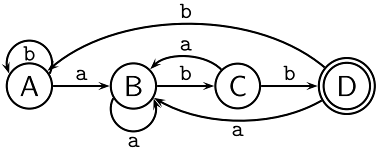
MINIMIZING THE NUMBER OF STATES

- For a typical language, the number of states of the DFA is in order of hundreds.
- Therefore we should try to minimize the number of states.

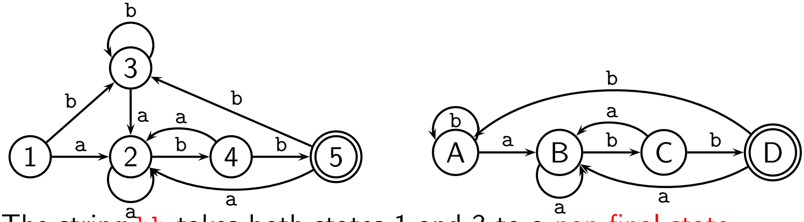
MINIMIZING THE NUMBER OF STATES

- The second DFA has been obtained by merging states 1 and 3 of the first DFA.
- Under what conditions can this merging take place?





MINIMIZING THE NUMBER OF STATES

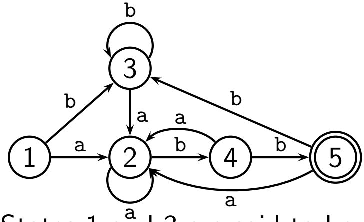


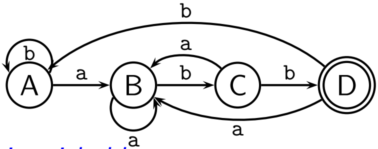
- The string bb takes both states 1 and 3 to a non-final state.
- The string aba takes both states 1 and 3 to a non-final state.
- The string ϵ takes both states 1 and 3 to a non-final state.
- The string bbabb takes both states 1 and 3 to a final state.

Observation:

Any string that takes state 1 to a final state also takes 3 to a final state. Conversely, any string that takes state 1 to a non-final state also takes 3 to a non-final state.

MINIMIZING THE NUMBER OF STATES





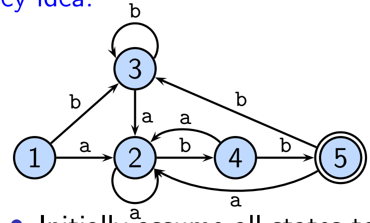
- States 1 and 3 are said to be *indistinguishable*.
- Minimimization strategy:
  - Find indistinguishable states.
  - Merge them.
- Question: How does one find indistingushable states?

MINIMIZING THE NUMBER OF STATES

Key idea:

MINIMIZING THE NUMBER OF STATES

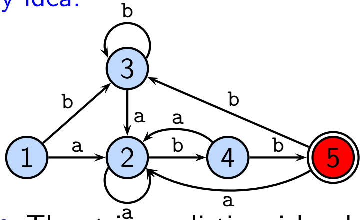
Key idea:



- Initially assume all states to be indistinguishable. Put them in a single set.

MINIMIZING THE NUMBER OF STATES

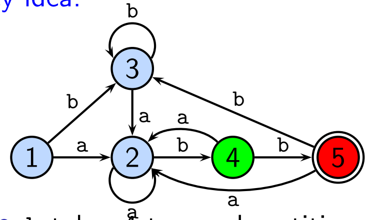
Key idea:



- The string ϵ distinguishes between final states and non-final states. Create two partitions.

MINIMIZING THE NUMBER OF STATES

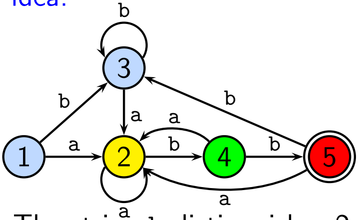
Key idea:



- b takes 4 to a red partition and retains other blue states in blue partition. Put 4 in a separate partition.

MINIMIZING THE NUMBER OF STATES

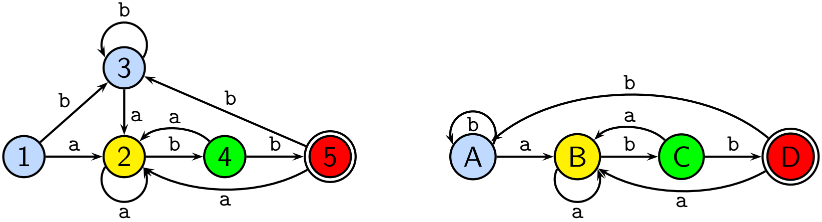
Key idea:



- The string b distinguishes 2 from other states in the blue partition.

MINIMIZING THE NUMBER OF STATES

Key idea:



- No other partition possible. Merge all states in the same partition.

Summary of the Method

1. Construct an initial partition $\pi = \{S - F, F_1, \ldots, F_n,\}$, where $F = F_1 \cup F_2 \cup \ldots F_n s$, and each F_i is the set of final states for some token i.
2. for each set G in π do
   partition G into subsets such that two states
   s and t of G are in the same subset if and only if
   for all input symbols a, states s and t have transitions
   onto states in the same set of π;
   replace G in π_{new} by the set of all subsets formed
3. If $\pi_{new} = \pi$, let $\pi_{final} := \pi$ and continue with step 4. Otherwise repeat step 2 with $\pi := \pi_{new}$.
4. Merge states in the same set of the partition.
5. Remove any dead states.