Lecture 4.1: The fundamental theorem

Let $L \subseteq \Sigma^*$. Then the following three statements are equivalent.

1. L is accepted by some NFA.
2. L is accepted by some DFA.
3. L is a regular language.

Proof: Proof plan

3 => 1 => 2 => 3

<u>3 => 1</u>:

- Given: L is a regular language.
- Need to show: L is accepted by some NFA.

L can be represented by a regular expression.

1. $\emptyset$ and each $\sigma \in \Sigma$ are regular expression.
2. If α and β are regular expressions, then
$$(\alpha\beta), (\alpha \cup \beta), \alpha^*$$
are regular expressions.

Induction on the number of operations (Concatenation, union, Kleene star) in the regular expression that represents L.

The fundamental theorem

Basis step: Zero operation in the regular expression that represents L.

- The expression is either $\emptyset$ or σ for some $\sigma \in \Sigma$.
- Here are the NFA's that accepts L.

  Figure 2-16

Induction hypothesis: Suppose 3 => 1 when the regular expression that represents L contains n or less operations.

Induction step: Need to show that 3 => 1 when the regular expression that represents L contains $n+1$ operations.

- There exist regular expressions α and β such that

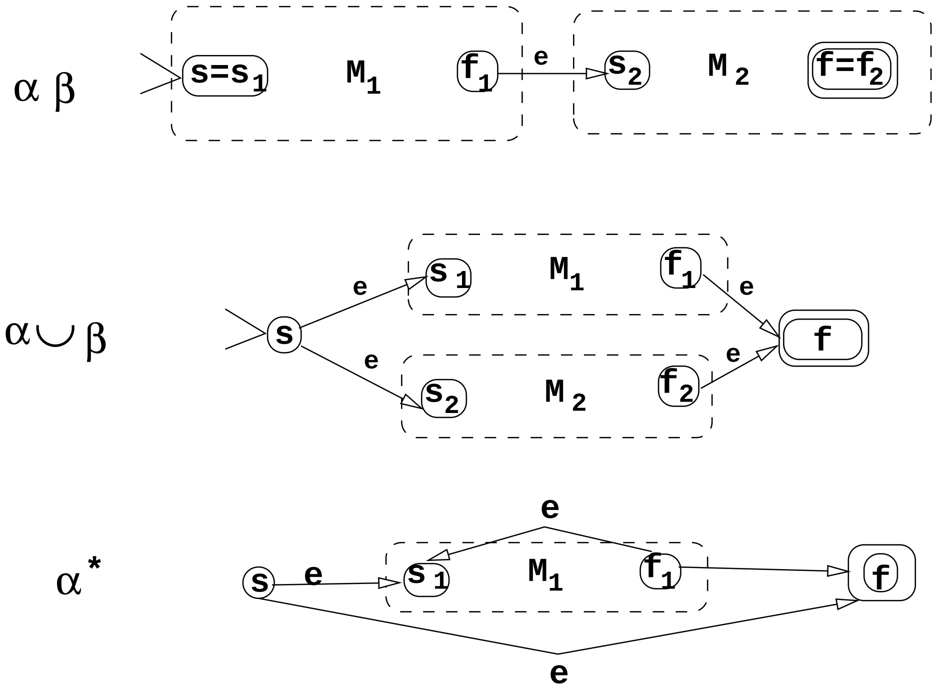
  Either $\alpha\beta$, or $\alpha\cup\beta$, or α^* that represent L, and

- Both α and β contain n or fewer operations.

- By the induction hypothesis, there exist

  1. NFA M_1 that accepts α, and
  2. NFA M_2 that accepts β.

- The following NFA's accept $\alpha\beta$, $\alpha\cup\beta$, or α^* respectively.


α β
s=s1
M1
f1
e
s2
M2
f=f2
α ∪ β
s
e
s1
M1
f1
e
f
e
s2
M2
f2
e
α*
e
s
e
s1
M1
f1
f
e

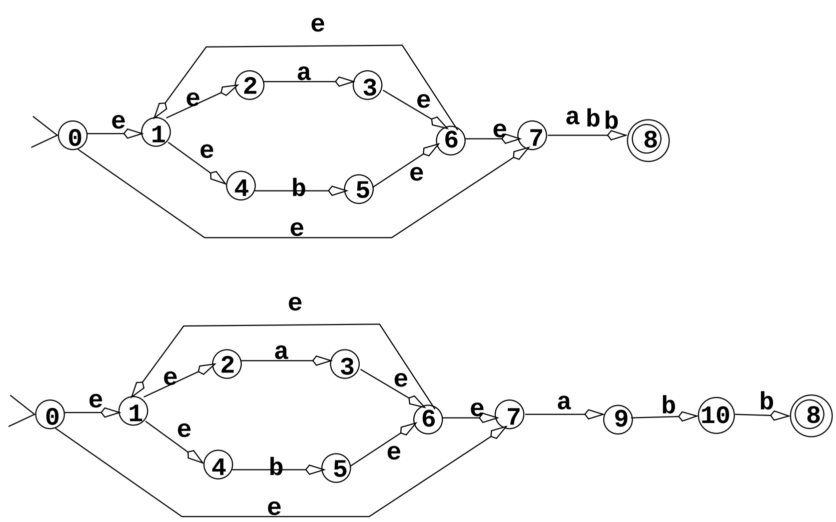
The fundamental theorem

<u>1 => 2</u>:

- Given: L is accepted by an NFA $M = \{K, \Sigma, \Delta, s, F\}$.
- Need to show: There exists a DFA M'' that accepts L.

Construct M'' from M:

1. Get rid of arcs labelled with string of length >1, yielding $M' = \{K', \Sigma, \Delta', s', F'\}$. Figure 2-9.



2. Define e-closure of q, $E(q)$, to be the set of states in a NFA that are *reachable* from q without reading any input (or via e-transitions *only*).

$$
\begin{array}{lll}
E(0) = \{0,1,2,4,7\} & E(4) = \{4\} & E(8) = \{8\} \\
E(1) = \{1,2,4\} & E(5) = \{5,6,7,1,2,4\} & E(9) = \{9\} \\
E(2) = \{2\} & E(6) = \{6,7,1,2,4\} & E(10) = \{10\} \\
E(3) = \{3,6,1,2,4\} & E(7) = \{7\} &
\end{array}
$$

3. *Subset construction*: $M'' = \{K'', \Sigma, \delta'', s'', F''\}$
   - $K'' = 2^{K'}, S'' = E(s')$,
   - $F'' = \{Q \in K'', Q \cap F' \neq \emptyset\}$,
   - $\delta''(Q, \sigma) = \cup\{E(p) : q \rightarrow^{\sigma} p \text{ for some} q \in Q\}$.

4. Example cont'd:

The fundamental theorem

---

2 => 3:

- Given: a DFA $M = (K, \Sigma, s, \delta, F)$.
- Need to show that $L(M)$ is regular.

1. Suppose $K = \{q_1, \ldots, q_n\}$, $s = q_1$.
2. $R(i, j, k) = \{$all strings in Σ^* that drive M from q_i to q_j without passing through any state numbered k or greater$\}$
   $= \{\sigma_1\sigma_2\ldots\sigma_l : q_i = q_{r_1} \rightarrow^{\sigma_1} q_{r_2} \rightarrow^{\sigma_2} \ldots \rightarrow^{\sigma_{l-1}} q_{l-1} \rightarrow^{\sigma_l} q_{r_l} = q_l$ s.t. $r_2, \ldots, r_{l-1} < k\}$.
3. $R(i, j, n+1)$ = all the strings that drive M from q_i to q_j.
4. $\cup\{R(1, j, n+1) : q_j \in F\}$
   = all the strings that drive M
   from $q_1 (= s)$ to a final state
   $= L(M)$.
5. If each $R(1, j, n+1)$ is regular, than $L(M)$ is regular.

The fundamental theorem

Given any i and j, prove by induction that $R(i, j, k)$ is regular for all k.

1. *Basis step*: When $k = 1$,

   $$R(i, j, 1) = \begin{cases} \{\sigma{\in}\Sigma : \delta(q_i, \sigma) = q_j\} & \text{if } i \neq j \\ \{e\} \cup \{\sigma{\in}\Sigma : \delta(q_i, \sigma) = q_j\} & \text{if } i = j. \end{cases}$$

   Hence $R(i, j, 1)$ is regular.

2. *Induction hypothesis*: Suppose $R(i, j, k)$ is regular for some k.

3. *Inductive step*: Show that $R(i, j, k + 1)$ is regular.

- $R(i, j, k+1) = \{$all strings that drive M from q_i to q_j without passing through any state numbered $k+1$ or greater$\}$

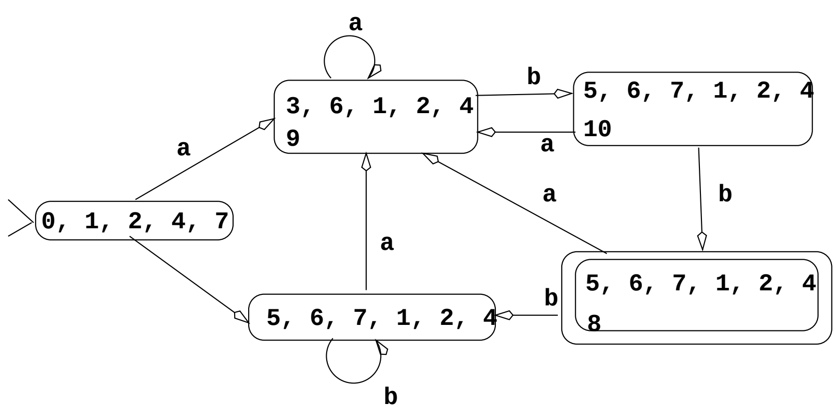
  $= \{$ all strings that drive M from q_i to q_j without passing through any state numbered k or greater$\}$

  $\cup\{$all the strings that drive M from q_i to q_k without passing any state numbered k or $>$,
  from q_k to q_k several times without passing any state numbered k or $>$,
  from q_k to q_j without passing any state numbered k or $>\}$

  $= R(i, j, k) \cup R(i, k, k) R(k, k, k)^* R(k, j, k)$.

- By the induction hypothesis, $R(i, j, k)$, $R(i, k, k)$, $R(k, k, k)$, and $R(k, j, k)$ are regular.
- Hence $R(i, j, k+1)$ is regular. The proof is completed.

Example cont'd: Get regular expression from DFA.


a
b
3, 6, 1, 2, 4
9
5, 6, 7, 1, 2, 4
10
a
a
a
b
0, 1, 2, 4, 7
a
5, 6, 7, 1, 2, 4
8
b
5, 6, 7, 1, 2, 4
b