

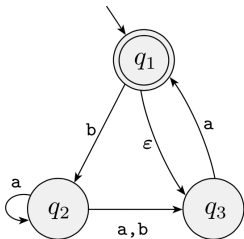
Lec 03. Equivalence of DFA and NFA & Closure under Regular Operations

Eunjung Kim

FORMAL DEFINITION OF NFA

NONDETERMINISTIC FA IS A 5-TUPLE $(Q, \Sigma, \delta, q_0, F)$

- Q a finite set called the states,
- Σ a finite set called the alphabet,
- δ a function from $Q \times \Sigma_{\epsilon}$ to 2^Q called the transition function,
- $q_0 \in Q$ the start state,
- $F \subseteq Q$ the set of accept states.

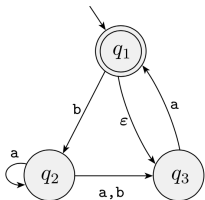


Write a formal description of this NFA

LANGUAGE RECOGNIZED BY NFA

NFA N ACCEPTS w IF

- 1 w can be written as y_1, \dots, y_m with $y_i \in \Sigma_\epsilon = \Sigma \cup \{\epsilon\}$,
- 2 there **exists** a sequence of states r_0, \dots, r_m s.t.
 - $r_0 = q_0$,
 - $r_{i+1} \quad (??) \quad \delta(q_i, y_i)$,
 - $r_m \in F$.



Write a computation tree for $w = baabaaa$. How many accepting paths?

EQUIVALENCE OF NFA AND DFA

NFA AND DFA OWN THE SAME COMPUTATIONAL POWER

For every NFA, there exists a deterministic finite automata which recognizes the same language (a.k.a. equivalent DFA).

Proof outline.

- Let $N = (Q, \Sigma, \delta, q_0, F)$ be an NFA.
- We want to construct a DFA $M = (Q', \Sigma, \delta', q'_0, F')$ such that $L(N) = L(M)$.
- Define $Q' := 2^Q$, i.e. the collection of all subsets of Q .
- Let us define $\delta', q'_0 \in 2^Q$ and $F' \subseteq 2^Q$,

PROOF: CONSTRUCTING DFA M , WHEN NO ϵ -TRANSITION

- $q'_0 = q_0$.
- transition function δ' from $2^Q \times \Sigma$ to 2^Q :
for every $R \in 2^Q$ (R is a subset of Q) and every symbol $a \in \Sigma$,

$$\delta'(R, a) := \bigcup_{r \in R} \delta(r, a)$$

- Define $F' \subseteq 2^Q$ as the collection of all subsets of Q containing at least one accept state of N .

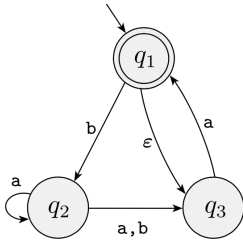
PROOF: CONSTRUCTING DFA M WITH ϵ -TRANSITION

- How to define the initial state for DFA: from a state $q \in Q$ of NFA N , any other state q' that can be reached by reading a string ϵ , can be aggregated with q to form a single state in DFA.
- Define $ext(q) \subseteq Q$ as the set of all states q' of N such that there is a directed path from q to q' in (the state diagram of) N each of whose arcs carries the label ϵ . Extend the definition $ext(X) := \bigcup_{q \in X} ext(q)$.
- transition function δ' from $2^Q \times \Sigma_\epsilon$ to 2^Q :
for every $R \in 2^Q$ (R is a subset of Q) and every symbol $a \in \Sigma$,

$$\delta'(R, a) := ext\left(\bigcup_{r \in R} \delta(r, a)\right)$$

- Define $q'_0 := ext(q_0) \in 2^Q$. Note that q'_0 corresponds to a subset of Q .
- Define $F' \subseteq 2^Q$ as the family of all subsets of Q containing at least one accept state of N .

CONSTRUCTING DFA M FROM NFA N , EXAMPLE



PROOF: $L(N) \subseteq L(M)$

- Let $\pi = (q_0, w = w_0), \dots, (q_i, w_i), \dots, (q_s, w_s = \epsilon)$ be an accepting computation history of N for w , that is

$$q_i \in \delta(q_{i-1}, y_i) \text{ for every } i \in [s] \text{ and } q_s \in F,$$

where $w_{i-1} = y_i \circ w_i$ with $y_i \in \Sigma \cup \{\epsilon\}$ for every $i \in [s]$.

- Let $0 \leq i_0 < \dots < i_{t-1} < s$ be the indices in π such that the leading symbol of w_{i_j} in (q_{i_j}, w_{i_j}) is in Σ (not ϵ) for each $j \in \{0, \dots, t-1\}$. Let $i_t = s$.
- Note that w is the form

$$\epsilon^* \circ y_{i_0+1} \circ \epsilon^* \circ y_{i_1+1} \circ \dots \circ \epsilon^* \circ y_{i_{t-1}+1} \circ \epsilon^*$$

with $y_{i_j+1} \in \Sigma$.

- Let $Q_0 = q'_0$. Inductively for each $j \in [t]$, let

$$Q_j = \delta'(Q_{j-1}, y_{i_{j-1}+1}).$$

- Now we have a computation history for $w = y_{i_0+1} \dots y_{i_{t-1}+1}$ in DFA M

$$\pi' = (Q_0, w), (Q_1, y_{i_1+1} \dots y_{i_{t-1}+1}), \dots (Q_t, \epsilon).$$

PROOF: $L(N) \subseteq L(M)$

- It remains to see Q_t is an accept state of M , i.e. the subset $Q_t \subseteq Q$ contains at least one accept state of N .
- It suffices to prove $q_{i_j} \in Q_j$ for each $0 \leq j \leq t$ (this implies the above).
- By induction. Base case holds because $Q_0 = q'_0 = \text{ext}(q_0)$ and q_{i_0} is reachable from q_0 by reading a number of ϵ 's.
- From the partition using i_0, \dots, i_t , q_{i_j} is reachable from $q_{i_{j-1}}$ by reading a single symbol $y_{i_{j-1}+1} \in \Sigma$ and additional ϵ 's. Hence $q_{i_j} \in \text{ext}(\delta(q_{i_{j-1}}, y_{i_{j-1}+1})) \subseteq \text{ext}(\delta(Q_{j-1}, y_{i_{j-1}+1})) = \delta'(Q_{j-1}, y_{i_{j-1}+1}) = Q_j$.

PROOF: $L(M) \subseteq L(N)$

- Let $\pi' = (Q_0, w' = w'_0), \dots, (Q'_i, w'_i), \dots, (Q'_t, w'_t = \epsilon)$ be an accepting computation history of M for w . By definition of computation history $Q'_i = \delta'(Q'_{i-1}, x_i)$, where $x_i \in \Sigma$ is the leading symbol of w'_{i-1} .
- We construct an accepting computation history of N by following the sequence π' backwardly.
- Observe: for each $q \in Q'_i \subseteq Q$, there exists a state $q' \in Q'_{i-1}$ such that from q' to q there is a computation history consisting of reading a string consisting of x_i followed by ϵ 's.
- Now starting from q_f , we concatenate computation histories witnessed by the previous observation.

CLOSURE UNDER REGULAR OPERATION

UNION OPERATION

Let A_1 and A_2 be two languages recognized by NFA N_1 and N_2 respectively. Then $A_1 \cup A_2$ is recognized by some NFA.

CLOSURE UNDER REGULAR OPERATION

CONCATENATION OPERATION

Let A_1 and A_2 be two languages recognized by NFA N_1 and N_2 respectively. Then $A_1 \circ A_2$ is recognized by some NFA.

CLOSURE UNDER REGULAR OPERATION

KLEENE STAR OPERATION

Let A a languages recognized by NFA N . Then A^* is recognized by some NFA.

CLOSURE UNDER COMPLEMENTATION

COMPLEMENTATION OPERATION

Let A a languages recognized by NFA N . Then \bar{A} , that is, $\Sigma^* - A$ is recognized by some NFA.

- For a regular language L , we can obtain a DFA recognizing the complement of L .
- ...using the trick...
- Can we use the same trick for NFA in general?

CLOSURE UNDER INTERSECTION

INTERSECTION OPERATION

Let A_1 and A_2 be two languages recognized by NFA N_1 and N_2 respectively. Then $A_1 \cap A_2$ is recognized by some NFA.

- Use the expression that $A_1 \cap A_2 = \text{??????}$.
- Combine the above (which ones?) operations on NFAs...
- Direct way with two DFAs M_1 and M_2 by simulating both automata simultaneously.

CLOSURE UNDER INTERSECTION

- Direct way with two DFAs M_1 and M_2 by simulating both automata simultaneously.

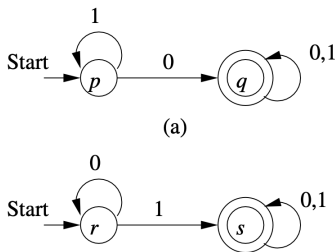


Figure 4.4 (a)-(b), Hopcroft et al. 2014.

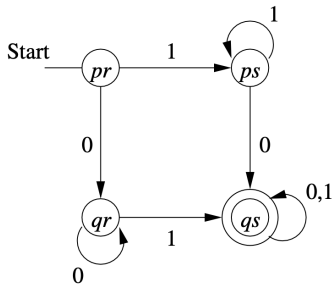


Figure 4.4 (c), Hopcroft et al. 2014.