# Finite Automata Sheet

Gallo Tenis

April 10, 2025

*Automata theory problems from the book "Automata Theory, Languages, and Computation" by J. Ullman.*

## DFA and NFA equivalence

1. **(Equivalence theorem.)** If $D = (Q_D, \Sigma, \delta_D, \{q_0\}, F_D)$ is the DFA constructed from NFA $N = (Q_N, \Sigma, \delta_N, q_0, F_N)$ by the subset construction, then $L(D) = L(N)$. Conversely if $L$ is a language that is accepted by some DFA then it is accepted by some NFA.
**Proof.** By definition of language of both $N$ and a generic DFA

$$L(N) := \{w \in \Sigma^* : \hat{\delta}_N(q_0, w) \cap F_N \neq \varnothing\}$$

$$L(D) := \{w \in \Sigma^* : \hat{\delta}_D(\{q_0\}, w) \in F_D\}$$

In our case, let $S = \hat{\delta}_D(\{q_0\}, w)$ (which is a state made of a set of $F_N$), then

$$S \in F_D \iff S \cap F_N \neq \varnothing.$$

In other words, the only thing to prove is that for any $w \in \Sigma^n$

$$\hat{\delta}_D(\{q_0\}, w) = \hat{\delta}_N(q_0, w).$$

We need to rely in the transition function constructed by subsets out of $\hat{\delta}_N$. By induction on $n = |w|$:

**Basis.** Assume $n = 0$ so that $w = \varepsilon$. Then the construction by subsets transition function will simply be $\delta_D(\{q_0\}, \varepsilon) = \{q_0\}$ then also we have

$$\hat{\delta}_D(\{q_0\}, \varepsilon) = q_0 = \hat{\delta}_N(q_0, \varepsilon).$$

Thus we conclude that $\delta_D = \hat{\delta}_D = \hat{\delta}_N$ (under renaming the singleton set $\{q_0\}$ as $q_0$).

**Hypothesis.** Assume that for the string $x$ of lenght $n$ it held that

$$\hat{\delta}_D(\{q_0\}, x) = \hat{\delta}_N(q_0, x) = \{p_1, \ldots, p_m\} = S.$$

**Thesis.** We have to prove that for a string $w = xa$ such that $n = |x|$ it follows that

$$\hat{\delta}_D(\{q_0\}, w) = \hat{\delta}_N(q_0, w).$$

To prove this assertion we recall the definition of $\hat{\delta}_D$ for any DFA and compare it to the one constructed by subsets. We have that

$$\hat{\delta}_D(\{q_0\}, w) \stackrel{def}{=} \delta_D(\hat{\delta}_D(\{q_0\}, x), a) \tag{1}$$

$$\delta_D(S, a) \stackrel{c.b.s.}{=} \bigcup_{p \in S} \delta_N(p, a) = \bigcup_{i=1}^{m} \delta_N(p_i, a). \tag{2}$$

Note that in (2), by construction $S$ is a subset of $Q_N$ such that

$$\hat{\delta}_D(\{q_0\}, w) = \delta_D(\hat{\delta}_D(\{q_0\}, x), a) = \delta_D(S, a) \tag{3}$$

hence, the transition function constructed by subsets is a valid transition function for $D$. Furthermore by equation (1)

$$\hat{\delta}_D(\{q_0\}, w) = \bigcup_{i=1}^{m} \delta_N(p_i, a).$$

We also note that in the case of the extended transition function of $N$

$$\hat{\delta}_N(q_0, w) \stackrel{def}{=} \bigcup_{i=1}^{m} \delta_N(p_i, a) = \delta_D(S, a) \stackrel{(3)}{=} \hat{\delta}_D(\{q_0\}, w).$$

Since Conversely, to prove that a DFA $D = (Q_D, \Sigma, \delta_D, \{q_0\}, F_D)$ induced language is the same as some NFA language, by structural induction on $n$, the lenght of the input string, we define the NFA $N$ as

**Basis.** We let the start state $q_0$ of $D$ be the start state also of $N$. So that for $w = \varepsilon$, $\hat{\delta}_D(\{q_0\}, \varepsilon) = \hat{\delta}_N(q_0, \varepsilon)$.

**Hypothesis.** For $|x| = n$ suppose $\hat{\delta}_D(\{q_0\}, x) = p_n$ then we let $\hat{\delta}_N(q_0, x) = \{p_n\}$.

**Thesis.** For $w = xa$ with $|w| = n + 1$, by hypothesis

$$\hat{\delta}_D(\{q_0\}, w) = \delta_D(\hat{\delta}_D(\{q_0\}, x), a) = \delta_D(p_n, a) = p_{n+1}.$$

We let $\hat{\delta}_N(q_0, w) = \{p_{n+1}\}$.

□

2. **(Worst case.)** Show that there exists a NFA with $n$ states such that equivalent DFA has $2^n$ states.
**Solution.** We shall build an NFA that accepts univocally the language

$$\{w \in \{0,1\}^* : \text{ the } n\text{-th symbol from the end of } w \text{ is } 1\}.$$

We first construct an NFA of this language having $n$ states. By structural induction on $|w|$

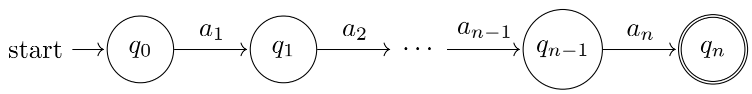
**Basis.**

**Hypothesis.**

**Thesis.**

3. Prove that if $N$ is an NFA that has at most one choice of state for any state and input symbol (i.e., $\delta(q, a)$ never has size greater than 1), then the DFA $D$ constructed from $N$ by the subset construction has exactly the states and transitions of $N$ plus transitions to a new dead state whenever $N$ is missing a transition for a given state and input symbol.

   **Solution.** Let $Q_N = q_0, q_1, \ldots, q_n$. Consider the maximal case in which each state except possibly the last one has exactly one transition. Therefore the automaton is isomorphic to

   It is clear that the NFA described must accept only strings $w$ of lenght $n$. Else, for the sake of contradiction, suppose that the lenght were $k > n$ then, by pigeonhole principle there must exist at least one state $q_i$ such that for a symbol $a_i$; $\delta_N(q_i, a_i) = \{p, r\}$, contradicting the fact that $\delta_N(q, a)$ is a singleton (supposing that $k < n$ yields the same contradiction, but we get an unused state instead).

   We claim that $\hat{\delta}_N(q_0, w) = \{q_n\}$ this is, $\hat{\delta}_N(q_0, w)$ is a singleton. To prove this claim, by induction on $n$ the amount of states (not counting the starting one) and the lenght of the accepted strings

**Basis.** If $n = 1$, then by problem hypothesis

$$\hat{\delta}_N(q_0, a_1) = \delta_N(q_0, a_1) = \{q_1\}.$$

**Hypothesis.** Assume that for $n$ states and for a string $x = a_1 \ldots a_n$ we had

$$\hat{\delta}_N(q_0, x) = \{q_n\}.$$

**Thesis.** We have to prove that for $n + 1$ states and for a string $xa_{n+1}$

$$\hat{\delta}_N(q_0, xa_{n+1}) = \{q_{n+1}\}.$$

By hypothesis,

$$\hat{\delta}_N(q_0, xa_{n+1}) = \bigcup_{p \in \hat{\delta}_N(q_0, x)} \delta_N(p, a_{n+1}) = \bigcup_{p \in \{q_n\}} \delta_N(p, a_{n+1}) = \delta_N(q_n, a_{n+1})$$

so by construction,

$$\delta_N(q_n, a_{n+1}) = \{q_{n+1}\}.$$

By induction on $k$ the lenght of any accepted string (and the amount of states distincts from the starting one), we claim that the equivalent DFA yields the same states as the NFA plus a dead state.

**Basis.** For $k = 2$ consider the transition table for the NFA,

| | $a_1$ | $a_2$ |
|---|---|---|
| $\rightarrow q_0$ | $\{q_1\}$ | $\varnothing$ |
| $q_1$ | $\varnothing$ | $\{q_2\}$ |
| $*q_2$ | $\varnothing$ | $\varnothing$ |

Then, the transition table for the equivalent DFA will be

| | $a_1$ | $a_2$ |
|---|---|---|
| $\varnothing$ | $\varnothing$ | $\varnothing$ |
| $\rightarrow q_0$ | $\{q_1\}$ | $\varnothing$ |
| $q_1$ | $\varnothing$ | $\{q_2\}$ |
| $*q_2$ | $\varnothing$ | $\varnothing$ |
| $\{q_0, q_1\}$ | $\{q_1\}$ | $\{q_2\}$ |
| $*\{q_0, q_2\}$ | $\{q_1\}$ | $\varnothing$ |
| $*\{q_1, q_2\}$ | $\varnothing$ | $\{q_2\}$ |
| $*\{q_0, q_1, q_2\}$ | $\{q_1\}$ | $\{q_2\}$ |

Changing names of states we see that

| | $a_1$ | $a_2$ |
|---|---|---|
| $\varnothing$ | $\varnothing$ | $\varnothing$ |
| $\rightarrow q_0$ | $\{q_1\}$ | $\varnothing$ |
| $q_1$ | $\varnothing$ | $\{q_2\}$ |
| $*q_2$ | $\varnothing$ | $\varnothing$ |
| $A$ | $\{q_1\}$ | $\{q_2\}$ |
| $*B$ | $\{q_1\}$ | $\varnothing$ |
| $*C$ | $\varnothing$ | $\{q_2\}$ |
| $*D$ | $\{q_1\}$ | $\{q_2\}$ |

so that in particular, $A$ and $D$ are the same state, $B$ is the same state as $q_0$ and $C$ is the same state as $q_1$. Furthermore, we see that $A, D$ are unreachable. We conclude that for $k = 2$ the transition table for the NFA is

| | $a_1$ | $a_2$ |
|---|---|---|
| $\varnothing$ | $\varnothing$ | $\varnothing$ |
| $\rightarrow q_0$ | $\{q_1\}$ | $\varnothing$ |
| $q_1$ | $\varnothing$ | $\{q_2\}$ |
| $*q_2$ | $\varnothing$ | $\varnothing$ |

.

**Hypothesis.** Assume that for a string $x = a_1 \ldots a_k$ of lenght $k$ it held that for an NFA $N$ with states $Q_N = \{q_0, \ldots, q_k\}$ the equivalent DFA had same $k$ states as the NFA plus one dead state: $Q_D = \{q_0, \ldots, q_k, \varnothing\}$.

**Thesis.** Consider the string $a_1 \dots a_k a_{k+1} = x a_{k+1}$, this case can be seen by adding a column $a_{k+1}$, hence by the first claim,

$$\hat{\delta}_N(q_0, x a_{k+1}) = \bigcup_{i=1}^{k} \delta_N(q_i, a_{k+1}) = \{q_{k+1}\}$$

so that the DFA also transitions to the singleton $\{q_{k+1}\}$ only in the state $q_k$ by subset construction. Then, necessarily $q_{k+1}$ transitions to $\varnothing$ since there are not further states.

We see that by hypothesis all the subsets of $\{q_0, \dots, q_k\}$ that are not singletons are unreachable, then since the only reachable state starting from $q_k$ is $q_{k+1}$ and $q_{k+1}$ transitions to $\varnothing$, it follows that any subset state of $\{q_1, \dots, q_k, q_{k+1}\}$ is unreachable.

$\square$

# $\varepsilon$-NFA

# Regular expressions