

Theorem: For every NFA  $N = (Q, \Sigma, \delta, s, F)$  we can build a DFA  $M = (2^Q, \Sigma, \Delta, S, F)$  such that  $L(M) = L(N)$ . The powerset  $P(Q)$  is the set of all subsets of  $Q$  but we hope  $|2^Q| \ll |P(Q)| = 2^{|Q|}$ .

Members of  $2^Q$  are sets of states, i.e. subsets of  $Q$ . If  $N$  has no  $\epsilon$ -trans then  $S = \{s\}$ .

Proof: Take  $S = \{q \in Q : N \text{ can process } \epsilon \text{ from } s \text{ to } q\}$ .

Idea: Maintain the inductive invariant  $R_i = \{q : N \text{ can process } x_1 \dots x_i \text{ from } s \text{ to } q\}$  that for all  $i$ , on any given input  $x = x_1 \dots x_i \dots x_n$ , the current DFA state is  $R_i$ .

Note  $R_0 = \{q : N \text{ can process } \epsilon \text{ from } s \text{ to } q\}$  = our defn of  $S$ , which is the initial current state of the DFA.

If we do this for all  $i$ , then with  $i=n$  we get  $R_n = \{q : N \text{ can process } x \text{ from } s \text{ to } q\}$ . Thus  $x \in L(N) \iff$  some member of  $F$  is in  $R_n$ , i.e. iff  $R_n \cap F \neq \emptyset$ .

Hence take  $\mathcal{F} = \{R \subseteq Q : R \cap F \neq \emptyset\}$  (will be fine to restrict to  $R \in 2^Q$ ).

Example:  $N =$  to DFA  $M = (2^Q, \Sigma, \Delta, S, F)$

wlog = "without loss of generality."

Example:  $N =$  to DFA  $M = (2^Q, \Sigma, \Delta, S, F)$

$|P(Q)| = 2^{|Q|}$

If  $N$  has no  $\epsilon$ -trans then  $S = \{s\}$ .

$\Delta(S, b) = \bigcup_{p \in S} \Delta(p, b) = \{4, 3, 2\} = \{3, 2, 4\}$  so  $p \in R_{i-1}$ , and  $N$  processed  $c = x_i$  from  $p$  to  $r$ , so  $r$  is in  $\Delta(R_{i-1}, c)$ .

For any subset  $P$  of  $Q$  (which might be a DFA state we've found) and any char  $c \in \Sigma$  (NOT  $\epsilon$ ) define  $\Delta(P, c) = \bigcup_{p \in P} \{ \text{states } r \text{ such that } N \text{ can process } c \text{ from } p \text{ to } r \}$

Thus  $R_i$  given by  $\Delta(R_{i-1}, c)$  is exactly the set of states  $r$  s.t.  $N$  can process  $x_1 \dots x_i$  from  $s$  to  $r$ .

Suppose  $c = x_i$  and  $R_{i-1}$  is correct, meaning  $R_{i-1}$  is both the set of  $p$  st.  $N$  can process  $x_1 \dots x_{i-1}$  from  $s$  to  $p$ .

We need to show that  $R_i = \Delta(R_{i-1}, c)$  satisfies (\*).

① If  $r \in \Delta(P, c)$  then for some  $p \in P$ ,  $N$  can process  $x_1 \dots x_{i-1}$  from  $s$  to  $p$  and  $x_i = c$  from  $p$  to  $r$ . So  $N$  can process  $x_1 \dots x_i$  from  $s$  to  $r$ .

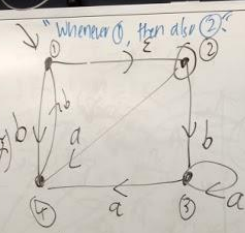
Note: the state chosen as "p" was after any  $\epsilon$ s that came after processing  $x_{i-1}$ , i.e. just before taking an arc on  $c$ . Hence it helps to define  $\Delta(p, c) = \{r : N \text{ can process } c \text{ from } p \text{ to } r \text{ by first taking an arc on } c.\}$

conversely, ② if  $N$  can process  $x_1 \dots x_i$  from  $s$  to  $r$ , then there is a  $p$  to which  $N$  processed  $x_1 \dots x_{i-1}$ , and  $N$  processed  $c = x_i$  from  $p$  to  $r$ , so  $r$  is in  $\Delta(R_{i-1}, c)$ .

Thus  $R_i$  given by  $\Delta(R_{i-1}, c)$  is exactly the set of states  $r$  s.t.  $N$  can process  $x_1 \dots x_i$  from  $s$  to  $r$ .

Example:  $N =$   
to

DFA  $M = (Q, \Sigma, \delta, q_0, F)$



$|P(q)| = 2^{|Q|}$

If  $N$  has no  $\epsilon$ -trans then  $S = \{s\}$ .

from  $s$  to  $q$

$q \neq$  our defn of  $S$ , which is the initial current we get state of the DFA at. Thus  $R_n$ , i.e. iff  $R_n \cap F \neq \emptyset$  be fine to restrict to  $R \in \mathcal{P}(Q)$

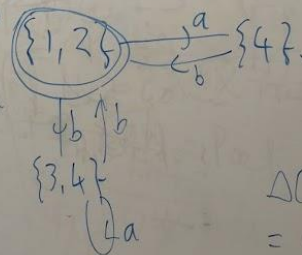
$\therefore S = \{1, 2\}$  not just  $\{1\}$

| $\delta$ | a           | b           |
|----------|-------------|-------------|
| 1        | $\emptyset$ | $\{4\}$     |
| 2        | $\{4\}$     | $\{3\}$     |
| 3        | $\{3, 4\}$  | $\emptyset$ |
| 4        | $\emptyset$ | $\{1, 2\}$  |

Build the DFA economically by breadth-first search.  $\mathcal{F} = \{\text{states that include } a\}$

Note: The computation itself may include  $\epsilon$ 's. But we may suppose that the state chosen as "p" was after any  $\epsilon$ 's that came after processing  $x_{i-1}$ , i.e. just before taking an arc on  $c$ . Hence it helps to define

$\delta(p, c) = \{r : N \text{ can process } c \text{ from } p \text{ to } r \text{ by first taking an arc on } c.\}$   
 $\Delta(p, c) = \bigcup_{p \in P} \delta(p, c)$



$\Delta(S, a) = \delta(1, a) \cup \delta(2, a) = \emptyset \cup \{4\} = \{4\}$

$\Delta(\{3, 4\}, b) = \delta(3, b) \cup \delta(4, b) = \emptyset \cup \{1, 2\} = \{1, 2\}$

No more new states.  $\therefore$  done with BFS and  $M$ .