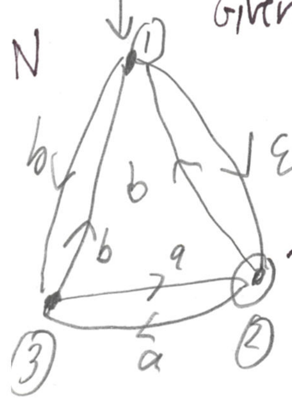
CSE396 Lecture Tue 2/20 Spring 2018 (1)

Top Hat# 1009

Theorem: For every NFA $N=(Q,\Sigma,\delta,s,F)$ we can build a DFA $M=(\mathcal{Q},\Sigma,\Delta,S,\mathcal{F})$ st. $L(M)=L(N)$.

Key Def^n: A NFA N can process a (sub-)string y of length n from state p to state r if there is a trace $\vec{c}=(q_0,u_1,q_1,u_2,q_2,\ldots,q_{m-1},u_m,q_m)$ such that: • for each j, $1\le j\le m$, $(q_{j-1},u_j,q_j)\in\delta$. (valid trace)
and • $q_0=p$, • $u_1\cdot u_2\cdots u_m=y$, and • $q_m=r$. (Some u_j can be ε. Later: other strings)



Given p,r, denote the set of such y by L_{pr}. $\boxed{\therefore L(N)=\bigcup_{q\in F} L_{sq}.}$

Our proof will maintain the following Inductive Invariant as $i=0$ to n.

INV. $\boxed{\text{The current state } R \text{ of the DFA equals the set of states } r\in Q \text{ st. } N \text{ can process } x_1\cdots x_i \text{ from } s \text{ to } r.}$

Initially, $R=R_0$ should be

Proof: Take $S=\{r: N \text{ can process } \varepsilon \text{ from } s \text{ to } r\}$ ($\equiv E(\{s\})$ in text). This makes INV hold with $i=0$ as the base case. Now let $i>0$, assume INV holds for $i-1$ as the induction hypothesis. Why will $i=n$ imply $L(M)=L(N)$?

- By INV for $i=n$, we want the final state R_n of the DFA M to equal the set of r st N can process x from s to r.
- R_n must be an accepting state exactly when some $q\in R_n$ belongs to F.

$\therefore S=R_0$, $\mathcal{F}=\{R\subseteq Q: R\cap F\neq\emptyset\}$. * We may be able to economize on $\mathcal{Q}$ and $\mathcal{F}$.

Thus if we define Δ so that INV always holds, we will get $L(M)=L(N)$.

Note: We may assume R_{i-1} comprises all states q st. N can process $x_1\cdots x_{i-1}$ from s to q, including "all trailing εs": Text: R_{i-1} is already ε-closed.

Helpful auxiliary fn (when doing problems): $\underline{\delta}(p,c) = \{r : N$ can process c from p to r by **first** using an arc (p,c,q) and then taking 0 or more ϵs to $r.\}$ ($c \in \Sigma$, not ϵ)

Finally define, for all $P \subseteq Q$ and $c \in \Sigma$:

$$\boxed{\Delta(P,c) = \bigcup_{p \in P} \underline{\delta}(p,c).}$$

To prove INV holds for $x_1 \cdots x_{i-1} x_i$, take $c = x_i$ and instantiate $P = R_{i-1}$. By Ind-Hyp, INV holds for R_{i-1}. Let any $r \in Q$ be given. We need: $r \in R_i \Leftrightarrow r \in \Delta(P,c)$

First suppose $r \in R_i$, meaning N can process $x_1 \cdots x_i$ from s to r. Where was N after processing the last ϵ before it did $c = x_i$? It was in some state p. By IH, $p \in R_{i-1}$

Picture the trace $x_1 \cdots x_{i-1}\ \epsilon \cdots \epsilon\ \overset{c}{x_i}\ \epsilon \cdots \epsilon$ with states p, q, r. Hence the DFA M includes r.

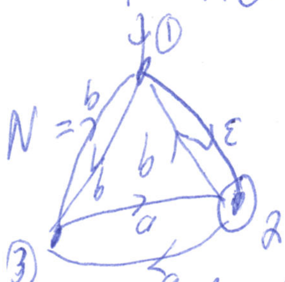
Conversely, suppose $r \in \Delta(P,c)$ where $c = x_i$ and $P = R_{i-1} = \{p : N$ can process $x_1 \cdots x_{i-1}$ from s to $p\}$

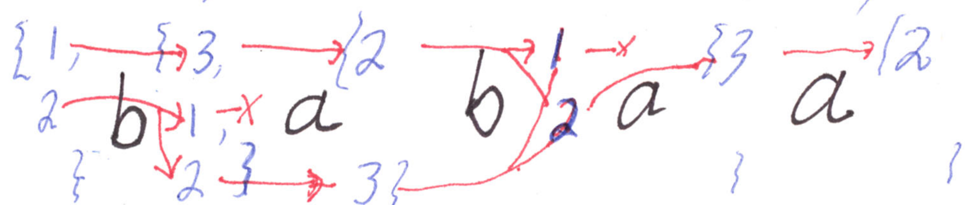
$r \in \bigcup_{p \in P} \underline{\delta}(p,c)$ so $r \in \underline{\delta}(p,c)$ for some p in R_{i-1}.

$\therefore$ N can process $x_1 \cdots x_{i-1}$ from s to this p, then process c and any trailing ϵs from p to r. $\therefore$ N can process $x_1 \cdots x_i$ from s to r. $\therefore$ INV holds for all $i = 1$ to n by induction

---

Added: An example of tracing the states of an NFA. Repeated for convenience:





Notice that individual paths can die and flow together. But the DFA only needs to maintain which NFA states are "active" at any one time. (Two paths process all of x; two others die.)

$R_0 = S = \{1,2\}$ $R_1 = \{1,2,3\}$ $R_2 = \{2,3\}$ $R_3 = \{1,2\}$ $R_4 = \{3\}$ $R_5 = \{2\}$

$x = babaa$

Say: "All light bulbs are lit after one step on b."

Down to one light bulb lit. Is there a string x' that turns all three lights off?

Example for N. By the ε-arc write "Whenever ① then also ②"

$S = \{1, \underline{2}\}$ from before.

Use "Breadth-First Search" (BFS)

$\underline{\delta}(1,a) = \underline{\underline{\emptyset}}$ Discipline: Applying ε's only $S=\{1,2\}$ will catch this.

$\underline{\delta}(1,b) = \{3\}$ Set 3, not just 3.

$\underline{\delta}(2,a) = \{3\}$

$\underline{\delta}(2,b) = \{1, \underline{\underline{2}}\}$

$\underline{\delta}(3,a) = \{2\}$ Not back on ε!

$\underline{\delta}(3,b) = \{1, 2\}$ again

$\Delta(S,a) = \underline{\delta}(1,a) \cup \underline{\delta}(2,a) = \emptyset \cup \{3\} = \{3\}$

$\Delta(S,b) = \underline{\delta}(1,b) \cup \underline{\delta}(2,b) = \{3\} \cup \{1,2\} = \{1,2,3\}$.

$\Delta(\{3\},a) = \underline{\delta}(3,a) = \{2\}$ new state

$\Delta(\{3\},b) = \underline{\delta}(3,b) = \{1,2\}$

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

done! Since $\{1,2\}$ has already been seen, we say that "the BFS closed."

accepting since ② ∈ F and ② ∈ S

DFA M =

$Q = \{\{1,2\}, \{1,2,3\}, \{2,3\}, \{2\}, \{3\}\}$ not all of $\mathcal{P}(\{1,2,3\})$!

$F = \{R \in Q : 2 \in R\} = Q \setminus \{\{3\}\}$.

$\boxed{\text{Added}}$: Let us trace the input $x = babaa$ on the DFA now:

(oops, not accepting)

$\vec{c} = (\{1,2\}, b, \{1,2,3\}, a, \{2,3\}, b, \{1,2\}, a, \{3\}, a, \{2\})$

Thus the DFA trace is the same as the "possible states trace" at each step of the NFA. This is by design — it exemplifies the proof.

To answer the question of whether there is a string x' that "turns off all lights," that would be true if and only if M has a reachable dead state. M has no dead state, so no. If M has a reachable dead state, it is when you get $\underline{\emptyset}$. Some other remarks:

- The states $\{1\}$ and $\{1,3\}$ are impossible by the ε-arc because they have ① but not ②.
- The state $\emptyset$ was possible but didn't come up in the "Breadth First Search" — which here simply means "expanding" only those states you get from previous stages. Can save lots of work.
- Note that sometimes with Cartesian Product you can economize because not all "pair" states can happen. When the Boolean operation is $\underline{\cup}$ then it works the same as building $N' =$ [s_3, s_1, M_1, s_2, M_2] and then doing NFA-to-DFA on N'.
- This particular DFA M is minimal, but sometimes what you get isn't.