COS 487, Fall 2003
September 24, 2003

Eliminating ϵ Arrows in NFA

Let $N = (Q, \Sigma, \delta, q_0, F)$ be an NFA with possible ϵ arrows. We transform it into an NFA without ϵ arrows.

Step 1: Add an arrow q_i to q_j with label $a \in \Sigma$, whenever there exists a sequence of states $q_{c_0}, q_{c_1}, \cdots, q_{c_m}$ and a $1 \leq k \leq m$ such that

$$\begin{aligned} q_{c_0} &= q_i, \\ q_{c_t} &\in \delta(q_{c_{t-1}}, \epsilon) \text{ for } 1 \leq t < k, \\ q_{c_k} &\in \delta(q_{c_{k-1}}, a), \\ q_{c_t} &\in \delta(q_{c_{t-1}}, \epsilon) \text{ for } k < t \leq m \\ q_{c_m} &= q_j. \end{aligned}$$

Step 2: If ϵ is in $L(N)$, make q_0 an accept state (if it is not already so).

Step 3: Delete all the ϵ arrows.

First note that Steps 1 and 2 do not cause any new string to be accepted. Call the NFA obtained after Steps 1 and 2 N', then $L(N') = L(N)$.

Call the NFA obtained after Step 3 N''. Clearly $L(N'') \subseteq L(N') = L(N)$. We now show $L(N'') \supseteq L(N)$.

Let $w \in L(N)$. If $w = \epsilon$, then since q_0 is an accept state of N'' (because of Step 2), it is accepted by N''. If $w = a_1 a_2 \cdots a_m$ where $m > 0$, then there exists a sequence of states $q_{i_0}, q_{i_1}, \cdots, q_{i_m}$ such that (1) $q_{i_0} = q_0$, (2) for each $1 \leq k \leq m$, there exists a sequence of moves in N that takes $q_{i_{k-1}}$ to q_{i_k} using ϵ arrows, exactly one arrow with label a_k, and then again only ϵ arrows, and (3) q_{i_m} is an accept state in N. Now, in N'', for each k there is an arrow with label a_k from $q_{i_{k-1}}$ to q_{i_k} (because of Step 1), and therefore we can conclude that w is also accepted by N''.

This completes the proof that $L(N'') = L(N)$.