Regular Expressions and Regular Languages

Theorem 1.54 (page 66):
A language is regular iff
some regular expression describes it.

“If” Direction

Proof (⇐) the “if” direction:
Given a regular expression R, show that the language described by R is regular.

Since R is a regular expression it was built according to the rules are the following slide.

Definition of Regular Expression

Basis:

a where a $\in \Sigma$ is a regular expression

ε is a regular expression

Φ is a regular expression

Recursion: Where R1 and R2 are regular expressions

$R1 \cup R2$ is a regular expression

$R1 \circ R2$ is a regular expression

$R1^*$ is a regular expression

"If" Direction continued

Show that an NFA can be defined for each of the basis regular expressions.

Basis

a where $a \in \Sigma$ is a regular expression
can be described by the NFA
$M = (\{q_0, q_1\}, \Sigma, \{((q_0, a), q_1)\}, q_0, \{q_1\})$

ε is a regular expression
can be described by the NFA
$M = (\{q_0\}, \Sigma, \{\}, q_0, \{q_0\})$

Φ is a regular expression
can be described by the NFA
$M = (\{q_0\}, \Sigma, \{\}, q_0, \{\})$

“If” Direction continued

Show that given NFAs for regular expressions R_1 and R_2, an NFA can be defined for each of the recursive steps.

Induction

Given NFAs for R_1 and R_2 as follows:

$M_1 = (Q_1, \Sigma, \delta_1, q_{1,0}, F_1)$ and

$M_2 = (Q_2, \Sigma, \delta_2, q_{2.0}, F_2)$,

respectively.

Induction, union

$R_1 \cup R_2$ is a regular expression can be described by the NFA

$M_{union} =$

$(Q_1 \cup Q_2 \cup \{q_{new}\}, \Sigma, \delta, q_{new}, F_1 \cup F_2)$

where

δ (q,x) is equal to the following:

δ_1 (q,x) for $q \in Q_1$

δ_2 (q,x) for $q \in Q_2$

$\{q_{1,0}, q_{2.0}\}$ for $q = \{q_{new}\}$ and $x = \varepsilon$

Induction, concatenation

$R_1 \circ R_2$ is a regular expression can be described by the NFA

$$M_{concat} = (Q_1 \cup Q_2, \Sigma, \delta, q_{1,0}, F_2)$$

where

$\delta(q,x)$ is equal to the following:

$\delta_1(q,x)$ for $q \in Q_1 - F_1$ or $x \in \Sigma$

$\delta_1(q,x) \cup \{q_{2,0}\}$ for $q \in F_1$ and $x = \varepsilon$

$\delta_2(q,x)$ for $q \in Q_2$

Induction, Kleene closure

R_1* is a regular expression can be described by the NFA

$M_{Kleene} = (Q_1 \cup \{q_{new}\}, \Sigma, \delta, q_{new}, \{q_{new}\})$

where
δ (q,x) is equal to the following:
δ_1 (q,x) for $q \in Q_1 - F_1$ or $x \in \Sigma$
$\delta_1 (q,x) \cup \{q_{new}\}$ for $q \in F_1$ and $x = \varepsilon$

Continued

Thus we have shown that if the inductive hypothesis holds, there are NFAs which recognize the languages described by the three new regular expressions above.

Conclusion

Thus by the Principle of Strong Mathematical Induction, given any regular expression R, there is an NFA which recognizes the language described by R.

Conclusion - continued

Since, from a previous theorem, any NFA can be converted into a DFA and since any language which can be recognized by a DFA is regular, we have shown that the language described by any regular expression is regular.
☺