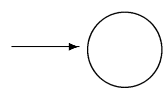
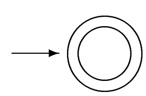
Regular Expression into Finite Automata

Let E **be a regular expression over** Σ**. Then we can construct a** λ**-NFA** M **such that** $L(M) = L(E)$**, using the following rules:**

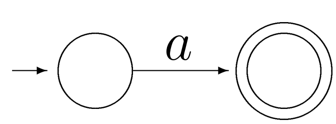
(i) $E = \emptyset$**.**
Construct M **such that** $L(M) = \emptyset$



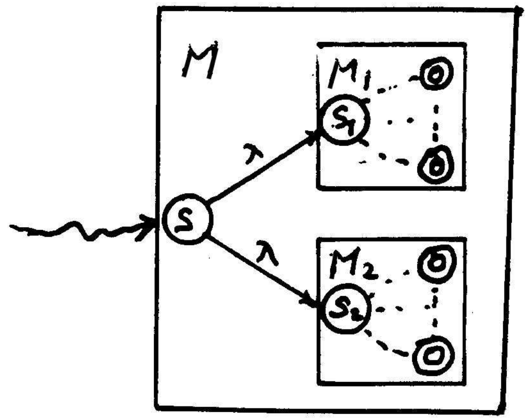
(ii) $E = \lambda$**.**
Construct M **such that** $L(M) = \{\lambda\}$



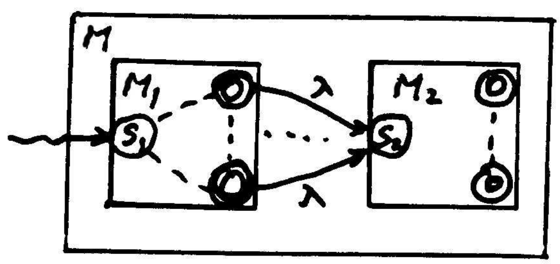
(iii) $E = a$**,** $a \in \Sigma$**.**
Construct M **such that** $L(M) = \{a\}$



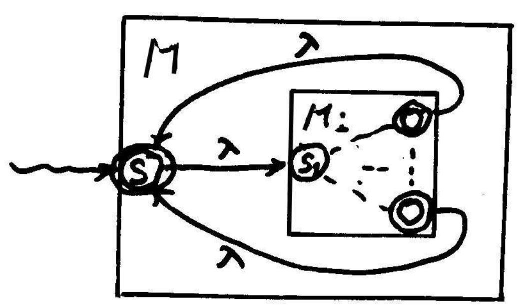
(iv) $E = [E_1 \cup E_2]$.
Construct M **such that** $L(M) = L(M_1) \cup L(M_2)$.



(v) $E = [E_1 E_2]$.
Construct M **such that** $L(M) = L(M_1)L(M_2)$.



(vi)$E = E_1^*$.
Construct M **such that** $L(M) = L(M_1)^*$.



Example

$E = [c^*[a \cup [bc^*]]^*]$

Construct a FA M such that $L(M) = L(E)$.

$\triangle$ $a, \quad b, \quad c$ **by (iii)**

$\triangle$ c^* **by (vi)**

$\triangle$ $[bc^*]$ **by (v)**

$\triangle$ $[a \cup bc^*]$ **by (iv)**

$\triangle$ $[a \cup bc^*]^*$ **by (vi)**

$\triangle$ $[c^*[a \cup bc^*]^*]$ **by (v)**

Theorem For E, an arbitrary regular expression over Σ, the λ-NFA, M, constructed as above satisfies $L(M) = L(E)$.

Proof: Let $Op(E)$ be the total number of $\cup$, $\cdot$, and * operations in E. We prove this theorem by induction on $Op(E)$.

Basis: $Op(E) = 0$. Then $E = \emptyset$, λ, or $a \in \Sigma$. Then clearly we have $L(M) = L(E)$.

Induction Hypothesis:
Assume the claim holds for all E with $Op(E) \leq k$, for some $k \geq 0$.

Induction Step:
Consider an arbitrary regular expression E with $Op(E) = k+1$. Since $k+1 \geq 1$, E contains at least one operator $\cup$, $\cdot$, or *.

Case I: $E = E_1 \cup E_2$. Then $Op(E_1) \leq k$ and $Op(E_2) \leq k$. So, $L(M_1) = L(E_1)$ and $L(M_2) = L(E_2)$ by I.H.. We know the construction of $M = M_1 \cup M_2$ satisfies $L(M) = L(M_1) \cup L(M_2)$, and $L(E) = L(E_1) \cup L(E_2)$

Therefore, $L(M) = L(E)$**.**

<u>Case II</u>: $E = E_1 E_2$

. . . . .

<u>Case III</u>: $E = E_1^*$

. . . . .

In each of the three cases, we have shown that $L(M) = L(E)$**. Therefore this holds for all regular expressions by the principle of induction.** *q.e.d.*