Phases of a compiler

Figure 1.6, page 5 of text
Process of building lexical analyzer

5) The minimal DFA is our lexical analyzer.
Focus last time
focus today

NFA ➔ DFA
first we construct an NFA from this regular expression

(alb)*abb
$(ab)^*abb$
(a|b)*abb
$(a|b)^{*}abb$
$(a|b)^*abb$
(a|b)*abb
\((ab)^*abb\)
(a|b)*abb
$(a|b)^*abb$
$(a|b)^*abb$
Operations

- \( \varepsilon\)-closure\( (t) \) is the set of states reachable from state \( t \) using only \( \varepsilon \)-transitions.

- \( \varepsilon\)-closure\( (T) \) is the set of states reachable from any state \( t \in T \) using only \( \varepsilon \)-transitions.

- move\( (T,a) \) is the set of states reachable from any state \( t \in T \) following a transition on symbol \( a \in \Sigma \).
NFA -> DFA algorithm
(set of states construction - page 153 of text)

- **INPUT:** An NFA $N = (S, \Sigma, \delta, s_0, F)$
- **OUTPUT:** A DFA $D = (S', \Sigma, \delta', s_0', F')$ such that $\mathcal{L}(D) = \mathcal{L}(N)$

**ALGORITHM:**

Compute $s_0' = \varepsilon$-closure($s_0$), an unmarked set of states
Set $S' = \{ s_0' \}$
while there is an unmarked $T \in S'$
  mark $T$
  for each symbol $a \in \Sigma$
    let $U = \varepsilon$-closure(move($T, a$))
    if $U \not\in S'$, add unmarked $U$ to $S'$
    add transition: $\delta'(T, a) = U$

$F'$ is the subset of $S'$ all of whose members contain a state in $F$. 
NFA -> DFA algorithm
(set of states construction - page 153 of text)

\[ S_0' = \{ A = \{0,1,2,4,7\} \} \]

Pick an unmarked set from \( S_0' \), A, mark it, and \( \forall x \in \Sigma \) let \( U = \varepsilon\text{-closure}(\text{move}(A,x)) \), if \( U \notin S' \), add unmarked \( U \) to \( S' \) and add transition: \( \delta'(A,x) = U \)

\[ S_1' = \{ A' , B = \{1,2,3,4,6,7,8\} , C = \{1,2,4,5,6,7\} \} \]
\[ \delta'(A,a) = B \]
\[ \delta'(A,b) = C \]

Pick an unmarked set from \( S_1' \), B, mark it, and \( \forall x \in \Sigma \) let \( U = \varepsilon\text{-closure}(\text{move}(B,x)) \), if \( U \notin S' \), add unmarked \( U \) to \( S' \) and add transition: \( \delta'(B,x) = U \)

\[ S_2' = \{ A' , B' , C , D = \{1,2,4,5,6,7,9\} \} \]
\[ \delta'(B,a) = B \]
\[ \delta'(B,b) = D \]

Pick an unmarked set from \( S_2' \), C, mark it, and \( \forall x \in \Sigma \) let \( U = \varepsilon\text{-closure}(\text{move}(C,x)) \), if \( U \notin S' \), add unmarked \( U \) to \( S' \) and add transition: \( \delta'(C,x) = U \)

\[ S_3' = \{ A' , B' , C' , D \} \]
\[ \delta'(C,a) = B \]
\[ \delta'(C,b) = C \]
NFA → DFA algorithm
(set of states construction - page 153 of text)

Pick an unmarked set from $S_3'$, $D$, mark it, and $\forall x \in \Sigma$ let $U = \varepsilon$-closure(move($D$, $x$)), if $U \notin S'$, add unmarked U to $S'$ and add transition: $\delta'(D, x) = U$

$S_4' = \{ A' \, ✔, \, B' \, ✔, \, C' \, ✔, \, D' \, ✔, \, E = \{1,2,4,5,6,7,10\} \} $
$\delta'(D, a) = B$
$\delta'(D, b) = E$

Pick an unmarked set from $S_4'$, $E$, mark it, and $\forall a \in \Sigma$ let $U = \varepsilon$-closure(move($E$, $a$)), if $U \notin S'$, add unmarked U to $S'$ and add transition: $\delta'(E, a) = U$

$S_5' = \{ A' \, ✔, \, B' \, ✔, \, C' \, ✔, \, D' \, ✔, \, E' \} $
$\delta'(E, a) = B$
$\delta'(E, b) = C$

Since there are no unmarked sets $S_5'$ the algorithm has reached a fixed point. STOP. $F'$ is the subset of $S'$ all of whose members contain a state in $F$: $\{E\}$
The resulting DFA

DFA = ( \{A, B, C, D, E\}, \{a, b\}, A, \delta', \{E\} ), where

\[\begin{align*}
\delta'(A,a) &= B \\
\delta'(A,b) &= C \\
\delta'(B,a) &= B \\
\delta'(B,b) &= D \\
\delta'(C,a) &= B \\
\delta'(C,b) &= C \\
\delta'(D,a) &= B \\
\delta'(D,b) &= E \\
\delta'(E,a) &= B \\
\delta'(E,b) &= C \\
\end{align*}\]