Converting an NFA to a DFA

Given:
A non-deterministic finite state machine (NFA)

Goal:
Convert to an equivalent deterministic finite state machine (DFA)

Why?
Faster recognizer!

Approach:
Consider simulating a NFA.
Work with sets of states.
IDEA: Each state in the DFA will correspond to a set of NFA states.

Worst-case:
There can be an exponential number \( O(2^N) \) of sets of states.
The DFA can have exponentially many more states than the NFA
... but this is rare.

NFA to DFA

Input: A NFA 
\( S = \text{States} = \{ s_0, s_1, ..., s_N \} = S_{\text{NFA}} \)
\( \delta = \text{Move function} = \text{Move}_{\text{NFA}} \)
Move’(S, a) \( \rightarrow \) Set of states

Output: A DFA 
\( S = \text{States} = \{ ?, ?, ..., ? \} = S_{\text{DFA}} \)
\( \delta = \text{Move function} = \text{Move}_{\text{DFA}} \)
Move(s, a) \( \rightarrow \) Single state from \( S_{\text{DFA}} \)

Main Idea:
Each state in \( S_{\text{DFA}} \) will be a set of states from the NFA
\( S_{\text{DFA}} = \{ ..., \{ ... \} , ..., \{ ... \} \} \)

(The names of the states is arbitrary and can be changed later, if desired.)
Algorithm: Convert NFA to DFA

We’ll use...
- $\text{Move}_{\text{NFA}}(S, a)$: the transition function from NFA
- $\epsilon$-$\text{Closure}(s)$: where $s$ is a single state from NFA
- $\epsilon$-$\text{Closure}(S)$: where $S$ is a set of states from NFA

We’ll construct...
- $S_{\text{DFA}}$: the set of states in the DFA
  Initially, we’ll set $S_{\text{DFA}}$ to $\{\}$
- Add $X$ to $S_{\text{DFA}}$: where $X$ is some set of NFA states
  Example: “Add $\{3,5,7\}$ to $S_{\text{DFA}}”$
  We’ll “mark” some of the states in the DFA.
  - Marked = “We’ve done this one” (√)
  - Unmarked = “Still need to do this one”
- $\text{Move}_{\text{DFA}}(T, b)$: The transition function from DFA
  To add an edge to the growing DFA...
  Set $\text{Move}_{\text{DFA}}(T, b)$ to $S$

...where $S$ and $T$ are sets of NFA states

Example

Start state: $\epsilon$-$\text{Closure}(0)$ = $\{a,b\}$

Σ = $\{a,b\}$
Lexical Analysis - Part 3

Example

Start state:
\( \varepsilon \)-Closure (0) = \{0, 1, 2, 4, 7\}

\( \Sigma = \{a, b\} \)

Lexical Analysis - Part 3

Example

Start state:
\( \varepsilon \)-Closure (0) = \{0, 1, 2, 4, 7\} = A
Example

\[\Sigma = \{a, b\}\]

Start state:
\[\varepsilon\text{-Closure (0)} = \{0, 1, 2, 4, 7\} = A\]

\[\text{Move}_{\text{DFA}}(A, a) = \]

\[\text{Move}_{\text{DFA}}(A, b) = \]

Example

\[\Sigma = \{a, b\}\]

Start state:
\[\varepsilon\text{-Closure (0)} = \{0, 1, 2, 4, 7\} = A\]

\[\text{Move}_{\text{DFA}}(A, a) = \varepsilon\text{-Closure (Move}_{\text{NFA}}(A, a)) = \]

\[\text{Move}_{\text{DFA}}(A, b) = \]

Example
Example

\[\Sigma = \{a, b\}\]

Start state:
\[\varepsilon\text{-Closure (0)} = \{0, 1, 2, 4, 7\} = A\]

\[\text{Move}_{\text{DFA}}(A,a) = \varepsilon\text{-Closure (Move}_{\text{NFA}}(A,a)) = \varepsilon\text{-Closure (\{3,8\})} = \]

\[\text{Move}_{\text{DFA}}(A,b) = \]

\[\Sigma = \{a, b\}\]

Start state:
\[\varepsilon\text{-Closure (0)} = \{0, 1, 2, 4, 7\} = A\]

\[\text{Move}_{\text{DFA}}(A,a) = \varepsilon\text{-Closure (Move}_{\text{NFA}}(A,a)) = \varepsilon\text{-Closure (\{3,8\})} = \{1,2,3,4,6,7,8\}\]

\[\text{Move}_{\text{DFA}}(A,b) = \]
Example

Start state:
\( \varepsilon \)-Closure (0)
= \{0, 1, 2, 4, 7\} = A

Move\(_{\text{DFA}}\)(A,a)
= \( \varepsilon \)-Closure (Move\(_{\text{NFA}}\)(A,a))
= \( \varepsilon \)-Closure ({3,8})
= \{1,2,3,4,6,7,8\} = B

Move\(_{\text{DFA}}\)(A,b)
= 

\[ \Sigma = \{a,b\} \]
Example

Start state:
\( \varepsilon\)-Closure (0) = \{0, 1, 2, 4, 7\} = A

\text{Move}_{\text{DFA}}(A,a) = \varepsilon\)-Closure (Move_{\text{NFA}}(A,a)) = \varepsilon\)-Closure (\{3,8\}) = \{1,2,3,4,6,7,8\} = B

\text{Move}_{\text{DFA}}(A,b) = \varepsilon\)-Closure (Move_{\text{NFA}}(A,b)) = \varepsilon\)-Closure (\{5\}) = \{1,2,4,5,6,7\} = C
Example

Start state:
\( \varepsilon\text{-Closure (0)} \)
\( = \{0, 1, 2, 4, 7\} = A \)

\( \text{Move}_{\text{DFA}}(A,a) \)
\( = \varepsilon\text{-Closure (Move}_{\text{NFA}}(A,a)) \)
\( = \varepsilon\text{-Closure } (\{3,8\}) \)
\( = \{1,2,3,4,6,7,8\} = B \)

\( \text{Move}_{\text{DFA}}(A,b) \)
\( = \varepsilon\text{-Closure (Move}_{\text{NFA}}(A,b)) \)
\( = \varepsilon\text{-Closure } (\{5\}) \)
\( = \{1,2,4,5,6,7\} = C \)

So far:

A is now done; mark it!
B and C are unmarked.
Let’s do B next...
Example

Process $B = \{1,2,3,4,6,7,8\}$

Move$_{DFA}(B,a)$

© Harry H. Porter, 2005
Process $B = \{1,2,3,4,6,7,8\}$

$\text{Move}_{\text{DFA}}(B,a) = \varepsilon\text{-Closure} (\text{Move}_{\text{NFA}}(B,a)) = \{3,8\}$

Example

$\Sigma = \{a,b\}$

© Harry H. Porter, 2005
Example

\[ \Sigma = \{ a, b \} \]

Process \( B = \{ 1, 2, 3, 4, 6, 7, 8 \} \)

\[
\text{Move}_{\text{DFA}}(B, a) = \varepsilon\text{-Closure (Move}_{\text{NFA}}(B, a)) = \varepsilon\text{-Closure (\{3, 8\})} = \{1, 2, 3, 4, 6, 7, 8\} = B
\]
**Example**

\[ \Sigma = \{a, b\} \]

Process \( B = \{1,2,3,4,6,7,8\} \)

\[ \text{Move}_{\text{DFA}}(B, a) = \varepsilon\text{-Closure (Move}_{\text{NFA}}(B, a)) = \varepsilon\text{-Closure (}\{3,8\}) = \{1,2,3,4,6,7,8\} = B \]

\[ \text{Move}_{\text{DFA}}(B, b) = \text{ } \]

**Example**

\[ \Sigma = \{a, b\} \]

Process \( B = \{1,2,3,4,6,7,8\} \)

\[ \text{Move}_{\text{DFA}}(B, a) = \varepsilon\text{-Closure (Move}_{\text{NFA}}(B, a)) = \varepsilon\text{-Closure (}\{3,8\}) = \{1,2,3,4,6,7,8\} = B \]

\[ \text{Move}_{\text{DFA}}(B, b) = \text{ } \]
Example

Process $B = \{1,2,3,4,6,7,8\}$

$\text{Move}_{\text{DFA}}(B,a) = \varepsilon$-Closure ($\text{Move}_{\text{NFA}}(B,a)$)

$= \varepsilon$-Closure ($\{3,8\}$)

$= \{1,2,3,4,6,7,8\} = B$

$\text{Move}_{\text{DFA}}(B,b) = \varepsilon$-Closure ($\text{Move}_{\text{NFA}}(B,b)$)

$= \varepsilon$-Closure ($\{5,9\}$)

$= \{1,2,4,5,6,7,9\} = D$

Example

Process $B = \{1,2,3,4,6,7,8\}$

$\text{Move}_{\text{DFA}}(B,a) = \varepsilon$-Closure ($\text{Move}_{\text{NFA}}(B,a)$)

$= \varepsilon$-Closure ($\{3,8\}$)

$= \{1,2,3,4,6,7,8\} = B$

$\text{Move}_{\text{DFA}}(B,b) = \varepsilon$-Closure ($\text{Move}_{\text{NFA}}(B,b)$)

$= \varepsilon$-Closure ($\{5,9\}$)

$= \{1,2,4,5,6,7,9\} = D$
Example

Process $B = \{1,2,3,4,6,7,8\}$

$\text{Move}_{\text{DFA}}(B,a) = \varepsilon$-Closure ($\text{Move}_{\text{NFA}}(B,a)$)

$= \varepsilon$-Closure ($\{3,8\}$)

$= \{1,2,3,4,6,7,8\} = B$

$\text{Move}_{\text{DFA}}(B,b) = \varepsilon$-Closure ($\text{Move}_{\text{NFA}}(B,b)$)

$= \varepsilon$-Closure ($\{5,9\}$)

$= \{1,2,4,5,6,7,9\} = D$

Example

Process $B = \{1,2,3,4,6,7,8\}$

$\text{Move}_{\text{DFA}}(B,a) = \varepsilon$-Closure ($\text{Move}_{\text{NFA}}(B,a)$)

$= \varepsilon$-Closure ($\{3,8\}$)

$= \{1,2,3,4,6,7,8\} = B$

$\text{Move}_{\text{DFA}}(B,b) = \varepsilon$-Closure ($\text{Move}_{\text{NFA}}(B,b)$)

$= \varepsilon$-Closure ($\{5,9\}$)

$= \{1,2,4,5,6,7,9\} = D$
Example

Process $C = \{1,2,4,5,6,7\}$

$\Sigma = \{a,b\}$

Move$_{DFA}(C,a) =$
Move$_{DFA}(C,b) =$
Example

Process $C = \{1,2,4,5,6,7\}$

$\text{Move}_{DFA}(C,a) = \{1,2,3,4,6,7,8\} = B$

$\text{Move}_{DFA}(C,b) =$

© Harry H. Porter, 2005
Example

Process $C = \{1,2,4,5,6,7\}$

$\text{Move}_{\text{DFA}}(C,a) = \{1,2,3,4,6,7,8\} = B$

$\text{Move}_{\text{DFA}}(C,b) = \{1,2,4,5,6,7\} = C$

Example

Process $D = \{1,2,4,5,6,7,9\}$

$\text{Move}_{\text{DFA}}(D,a) =$

$\text{Move}_{\text{DFA}}(D,b) =$
Example

\[\Sigma = \{a, b\}\]

Process \(C = \{1, 2, 4, 5, 6, 7\}\)

- \(\text{Move}_{\text{DFA}}(C,a) = \{1, 2, 3, 4, 6, 7, 8\} = B\)
- \(\text{Move}_{\text{DFA}}(C,b) = \{1, 2, 4, 5, 6, 7\} = C\)

Process \(D = \{1, 2, 4, 5, 6, 7, 9\}\)

- \(\text{Move}_{\text{DFA}}(D,a) = \{1, 2, 3, 4, 6, 7, 8\} = B\)
- \(\text{Move}_{\text{DFA}}(D,b) = \{1, 2, 4, 5, 6, 7, 10\} = E\)
Example

Process $C = \{1,2,4,5,6,7\}$

$\text{Move}_{DFA}(C,a) = \{1,2,3,4,6,7,8\} = B$

$\text{Move}_{DFA}(C,b) = \{1,2,4,5,6,7\} = C$

Process $D = \{1,2,4,5,6,7,9\}$

$\text{Move}_{DFA}(D,a) = \{1,2,3,4,6,7,8\} = B$

$\text{Move}_{DFA}(D,b) = \{1,2,4,5,6,7,10\} = E$

Example

Process $E = \{1,2,4,5,6,7,10\}$

$\text{Move}_{DFA}(E,a) = \{1,2,3,4,6,7,8\} = B$

$\text{Move}_{DFA}(E,b) = \{1,2,4,5,6,7,10\} = E$
Example

Process $C = \{1,2,4,5,6,7\}$
\[
\text{Move}_{\text{DFA}}(C,a) = \{1,2,3,4,6,7,8\} = B \\
\text{Move}_{\text{DFA}}(C,b) = \{1,2,4,5,6,7\} = C
\]

Process $D = \{1,2,4,5,6,7,9\}$
\[
\text{Move}_{\text{DFA}}(D,a) = \{1,2,3,4,6,7,8\} = B \\
\text{Move}_{\text{DFA}}(D,b) = \{1,2,4,5,6,7,10\} = E
\]

Process $E = \{1,2,4,5,6,7,10\}$
\[
\text{Move}_{\text{DFA}}(E,a) = \{1,2,3,4,6,7,8\} = B \\
\text{Move}_{\text{DFA}}(E,b) = \{1,2,4,5,6,7\} = C
\]
**Example**

Process C = \{1,2,4,5,6,7\}

- \(\text{Move}_{DFA}(C,a) = \{1,2,3,4,6,7,8\} = B\)
- \(\text{Move}_{DFA}(C,b) = \{1,2,4,5,6,7\} = C\)

Process D = \{1,2,4,5,6,7,9\}

- \(\text{Move}_{DFA}(D,a) = \{1,2,3,4,6,7,8\} = B\)
- \(\text{Move}_{DFA}(D,b) = \{1,2,4,5,6,7,10\} = E\)

Process E = \{1,2,4,5,6,7,10\}

- \(\text{Move}_{DFA}(E,a) = \{1,2,3,4,6,7,8\} = B\)
- \(\text{Move}_{DFA}(E,b) = \{1,2,4,5,6,7\} = C\)

---

**Final States in DFA?**

...which state(s) contain 10?
Example

Process C = \{1,2,4,5,6,7\}

\(\text{Move}_{\text{DFA}}(C, a) = \{1,2,3,4,6,7,8\} = B\)

\(\text{Move}_{\text{DFA}}(C, b) = \{1,2,4,5,6,7\} = C\)

Process D = \{1,2,4,5,6,7,9\}

\(\text{Move}_{\text{DFA}}(D, a) = \{1,2,3,4,6,7,8\} = B\)

\(\text{Move}_{\text{DFA}}(D, b) = \{1,2,4,5,6,7,10\} = E\)

Process E = \{1,2,4,5,6,7,10\}

\(\text{Move}_{\text{DFA}}(E, a) = \{1,2,3,4,6,7,8\} = B\)

\(\text{Move}_{\text{DFA}}(E, b) = \{1,2,4,5,6,7\} = C\)

Final Result:

\[A\] a b a
\[B\] a b a
\[C\] b a b
\[D\] b a b
\[E\] b a b

\(\Sigma = \{a, b\}\)

Algorithm: Convert NFA to DFA

\(S_{\text{DFA}} = \{}\)

Add \(\varepsilon\text{-Closure}(s_0)\) to \(S_{\text{DFA}}\) as the start state

Set the only state in \(S_{\text{DFA}}\) to “unmarked”

while \(S_{\text{DFA}}\) contains an unmarked state do

Let \(T\) be that unmarked state

Mark \(T\) for each \(a\) in \(\Sigma\) do

\(S = \varepsilon\text{-Closure}(\text{Move}_{\text{NFA}}(T, a))\)

if \(S\) is not in \(S_{\text{DFA}}\) already then

Add \(S\) to \(S_{\text{DFA}}\) (as an “unmarked” state)

endIf

Set \(\text{Move}_{\text{DFA}}(T, a)\) to \(S\)

endFor

endWhile

for each \(S\) in \(S_{\text{DFA}}\) do

if any \(s \in S\) is a final state in the NFA then

Mark \(S\) as a final state in the DFA

endIf

endFor
Resulting DFA for \((a \mid b)^*abb\)

Is it minimal?

© Harry H. Porter, 2005
Resulting DFA for \((a \mid b)^*abb\)

Is it minimal?

\[
\begin{align*}
A & \xrightarrow{a} B \xrightarrow{b} D \xrightarrow{b} E \\
C & \xrightarrow{a} B \xrightarrow{b} D \xrightarrow{b} E \\
& \xrightarrow{b} A
\end{align*}
\]
Resulting DFA for \((a|b)*abb\)

Is it minimal?

Every Regular Set is recognized by a minimal DFA!
Resulting DFA for \((a|b)*abb\)

Is it minimal?

Every Regular Set is recognized by a minimal DFA!

And it is unique, up to renaming of states