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:

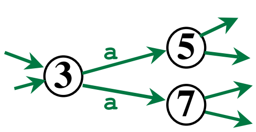
There can be an exponential number $\mathbf{O}(2^N)$ of sets of states.
The DFA can have exponentially many more states than the NFA
... but this is rare.

NFA to DFA

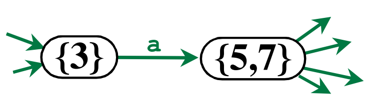
Input: A NFA — S = States = $\{ s_0, s_1, \ldots, s_N \} = S_{NFA}$
δ = Move function = $Move_{NFA}$
Move'(S, a) → Set of states

Output: A DFA — S = States = $\{?, ?, \ldots, ?\} = S_{DFA}$
δ = Move function = $Move_{DFA}$
Move(s, a) → Single state from S_{DFA}

Main Idea: Each state in S_{DFA} will be a set of states from the NFA
$S_{DFA} = \{ \{\ldots\}, \{\ldots\}, \ldots, \{\ldots\} \}$

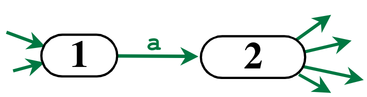


NFA



DFA

(The names of the states is arbitrary and can be changed later, if desired.)



Algorithm: Convert NFA to DFA

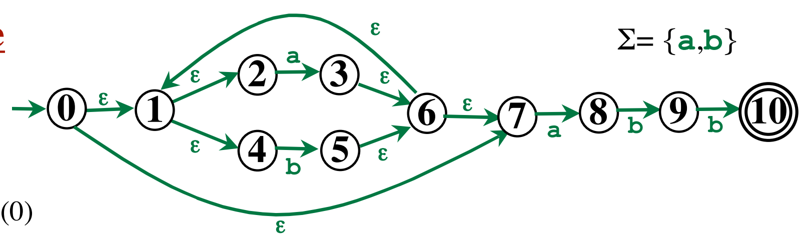
We'll use...

- `Move`$_{NFA}$`(S,a)` — the transition function from NFA
- `ε-Closure(s)` — where **s** is a single state from NFA
- `ε-Closure(S)` — where **S** is a set of states from NFA

We'll construct...

- **S**$_{DFA}$ — the set of states in the DFA
  - Initially, we'll set **S**$_{DFA}$ to {}
- Add **X** to **S**$_{DFA}$ — where **X** is some *set of* NFA states
  - *Example:* "Add {3,5,7} to **S**$_{DFA}$"
  - We'll "mark" some of the states in the DFA.
    - Marked = *"We've done this one"* (√)
    - Unmarked = *"Still need to do this one"*
- `Move`$_{DFA}$`(T,b)` — The transition function from DFA
  - To add an edge to the growing DFA...
    - `Set Move`$_{DFA}$`(T,b) to S`
    - $T \xrightarrow{b} S$
  - ...where **S** and **T** are sets of NFA states

Example



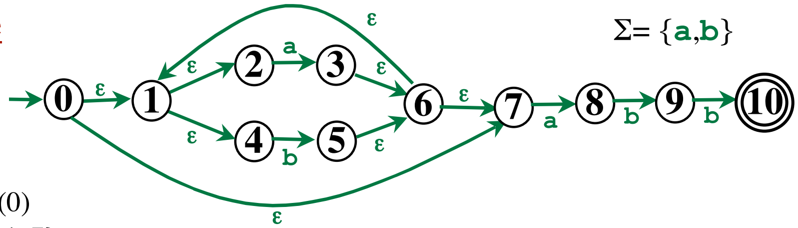
Σ= {a,b}

Start state:

ε-Closure (0)

=

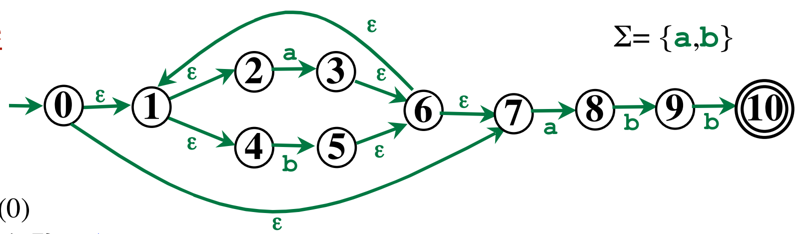
Example



Σ= {a,b}

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

Example

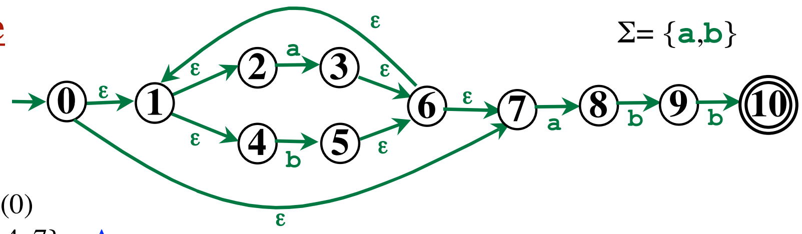


Σ= {a,b}

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

Example

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



Start state:

ε-Closure (0)

= {0, 1, 2, 4, 7} = A

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

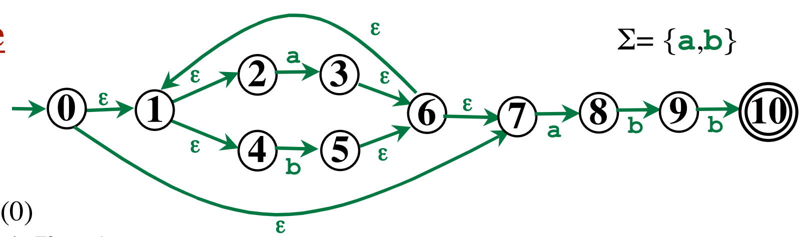
=

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

=

Example

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



Start state:

ε-Closure (0)

= {0, 1, 2, 4, 7} = A

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

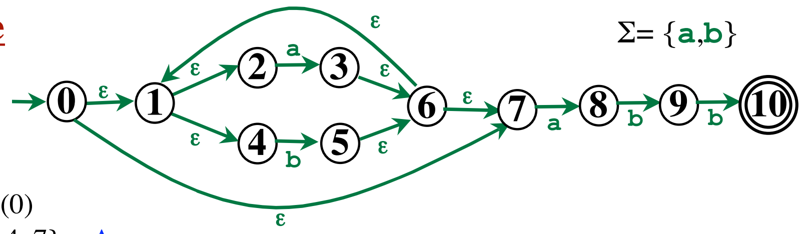
= ε-Closure ($\text{Move}_{\text{NFA}}(A, a)$)

=

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

=

Example



Σ= {a,b}

Start state:

ε-Closure (0)

= {0, 1, 2, 4, 7} = A

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

= ε-Closure (Move$_{NFA}$(A,a))

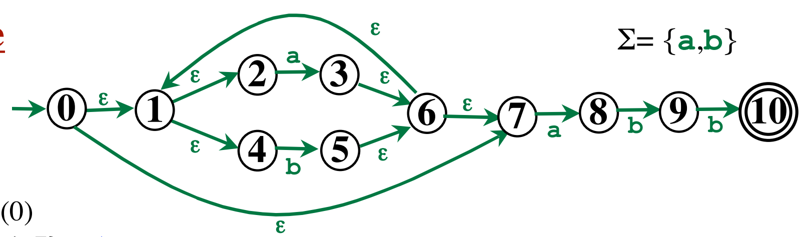
= ε-Closure ({3,8})

=

Move$_{DFA}$(A,b)

=

Example



Σ= {a,b}

Start state:

ε-Closure (0)

= {0, 1, 2, 4, 7} = A

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

= ε-Closure (Move$_{NFA}$(A,a))

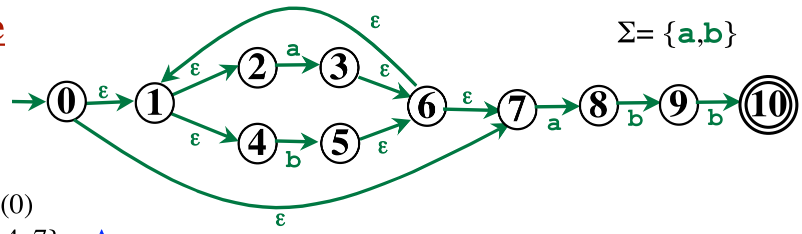
= ε-Closure ({3,8})

= {1,2,3,4,6,7,8}

Move$_{DFA}$(A,b)

=

Example



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

Start state:

ε-Closure (0)

= {0, 1, 2, 4, 7} = A

$\text{Move}_{\textbf{DFA}}$(A,**a**)

= ε-Closure ($\text{Move}_{\textbf{NFA}}$(A,**a**))

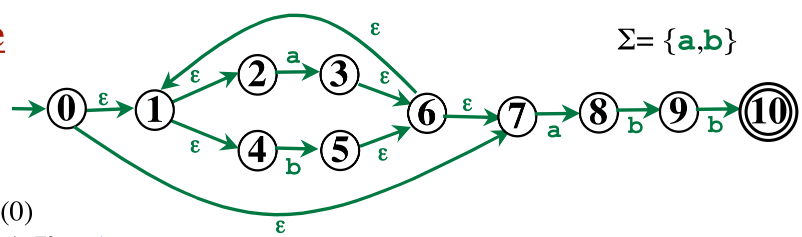
= ε-Closure ({3,8})

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

$\text{Move}_{\textbf{DFA}}$(A,**b**)

=

Example



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

Start state:

ε-Closure (0)

= {0, 1, 2, 4, 7} = A

$\text{Move}_{\textbf{DFA}}$(A,**a**)

= ε-Closure ($\text{Move}_{\textbf{NFA}}$(A,**a**))

= ε-Closure ({3,8})

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

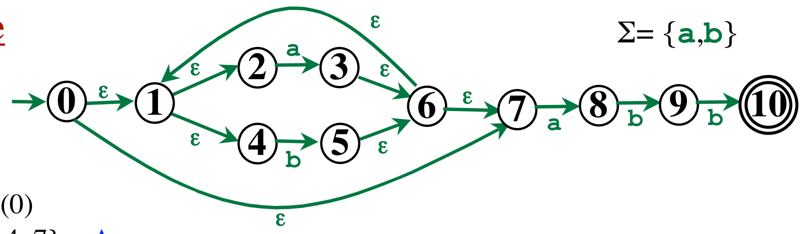
$\text{Move}_{\textbf{DFA}}$(A,**b**)

= ε-Closure ($\text{Move}_{\textbf{NFA}}$(A,**b**))

=

Example

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



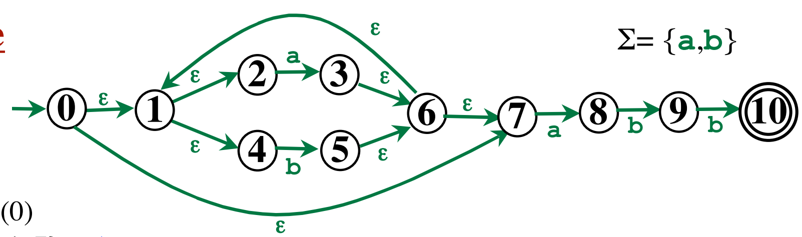
Start state:
ε-Closure (0)
= {0, 1, 2, 4, 7} = A

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

$\text{Move}_{\textbf{DFA}}$(A,**b**)
= ε-Closure ($\text{Move}_{\textbf{NFA}}$(A,**b**))
= ε-Closure ({5})
=

Example

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



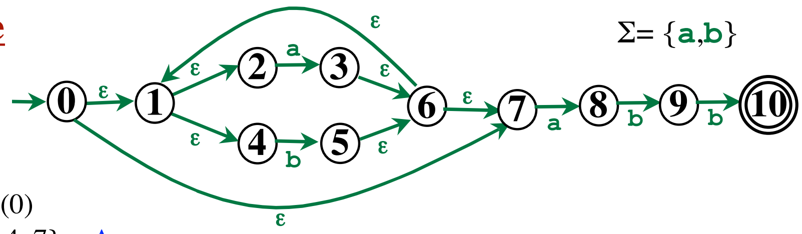
Start state:
ε-Closure (0)
= {0, 1, 2, 4, 7} = A

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

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

Example

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



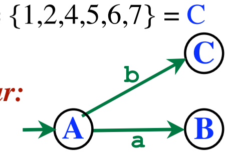
Start state:

ε-Closure (0)
= {0, 1, 2, 4, 7} = A

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

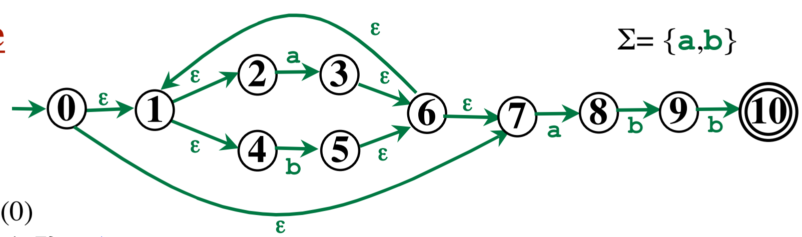
$\text{Move}_{\text{DFA}}(A,b)$
= ε-Closure ($\text{Move}_{\text{NFA}}(A,b)$)
= ε-Closure ({5})
= {1,2,4,5,6,7} = C

So far:



Example

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



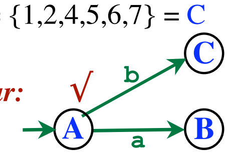
Start state:

ε-Closure (0)
= {0, 1, 2, 4, 7} = A

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

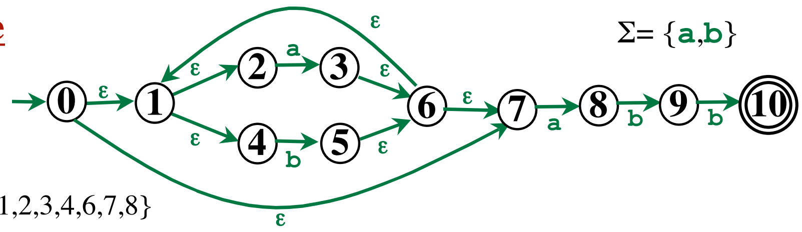
$\text{Move}_{\text{DFA}}(A,b)$
= ε-Closure ($\text{Move}_{\text{NFA}}(A,b)$)
= ε-Closure ({5})
= {1,2,4,5,6,7} = C

So far:



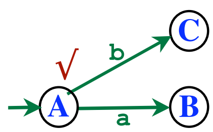
A is now done; mark it!
B and C are unmarked.
Let's do B next...

Example

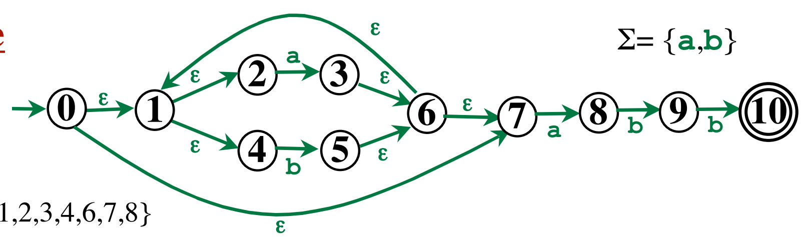


Σ= {a,b}

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



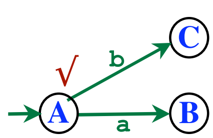
Example



Σ= {a,b}

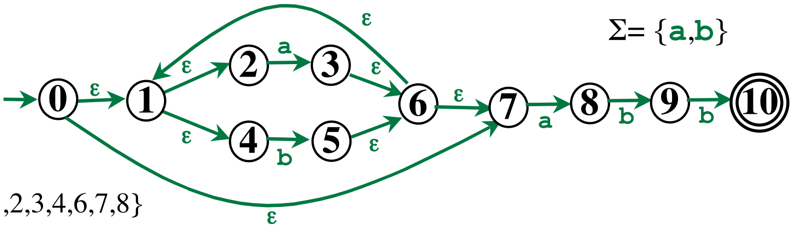
Process B = {1,2,3,4,6,7,8}

Move_{DFA}(B,a)
=



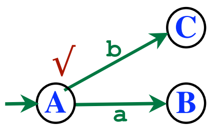
Example

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



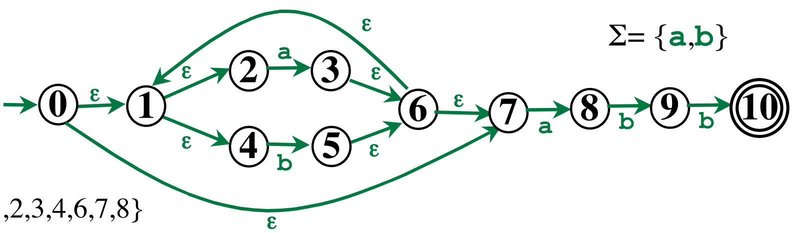
Process B = {1,2,3,4,6,7,8}

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



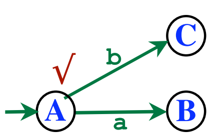
Example

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



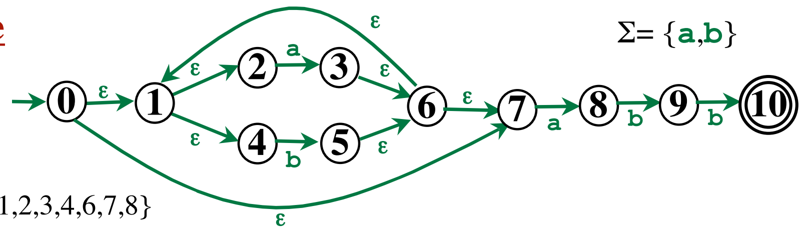
Process B = {1,2,3,4,6,7,8}

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



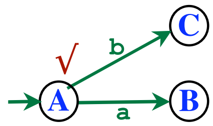
Example

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



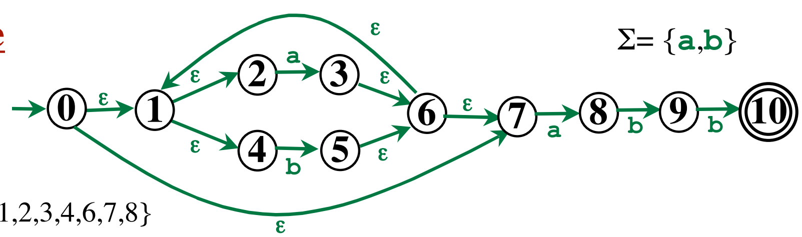
Process B = {1,2,3,4,6,7,8}

$\text{Move}_{\textbf{DFA}}(B,\mathtt{a})$
= ε-Closure ($\text{Move}_{\textbf{NFA}}(B,\mathtt{a})$)
= ε-Closure ({3,8})
= {1,2,3,4,6,7,8} = B



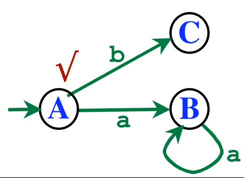
Example

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



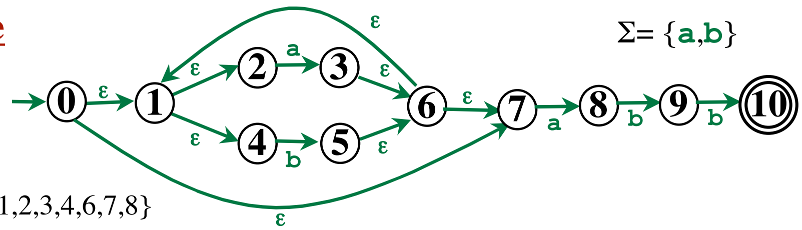
Process B = {1,2,3,4,6,7,8}

$\text{Move}_{\textbf{DFA}}(B,\mathtt{a})$
= ε-Closure ($\text{Move}_{\textbf{NFA}}(B,\mathtt{a})$)
= ε-Closure ({3,8})
= {1,2,3,4,6,7,8} = B



Example

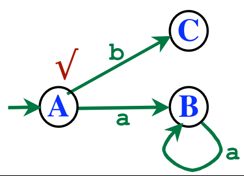
$\Sigma = \{a,b\}$



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

$Move_{DFA}$(B,a)
= ε-Closure ($Move_{NFA}$(B,a))
= ε-Closure ({3,8})
= {1,2,3,4,6,7,8} = B

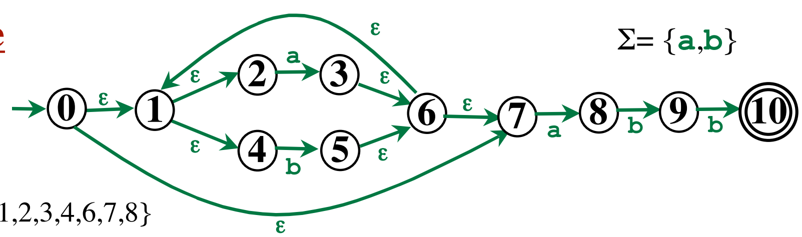
$Move_{DFA}$(B,b)
=



Lexical Analysis - Part 3

Example

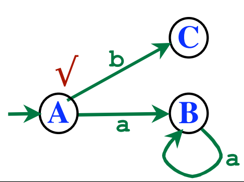
$\Sigma = \{a,b\}$



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

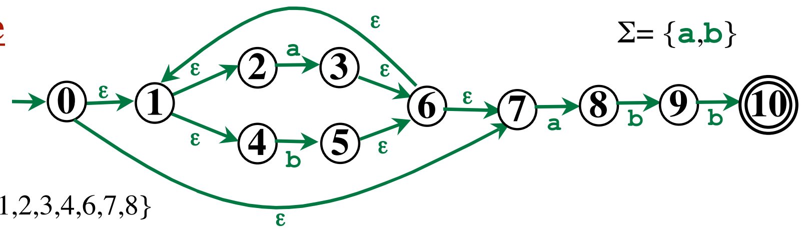
$Move_{DFA}$(B,a)
= ε-Closure ($Move_{NFA}$(B,a))
= ε-Closure ({3,8})
= {1,2,3,4,6,7,8} = B

$Move_{DFA}$(B,b)
= ε-Closure ($Move_{NFA}$(B,b))
=



Example

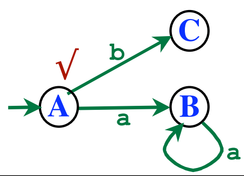
$\Sigma = \{a,b\}$



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

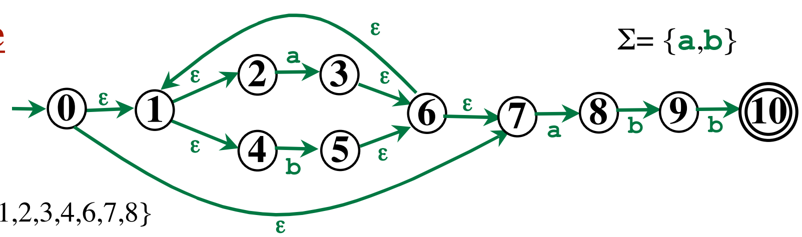
$\text{Move}_{\text{DFA}}(B,a)$
= ε-Closure ($\text{Move}_{\text{NFA}}(B,a)$)
= ε-Closure ({3,8})
= {1,2,3,4,6,7,8} = B

$\text{Move}_{\text{DFA}}(B,b)$
= ε-Closure ($\text{Move}_{\text{NFA}}(B,b)$)
= ε-Closure ({5,9})
=



Example

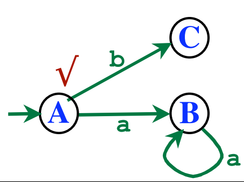
$\Sigma = \{a,b\}$



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

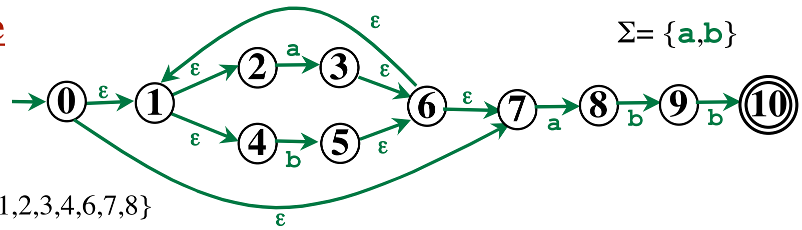
$\text{Move}_{\text{DFA}}(B,a)$
= ε-Closure ($\text{Move}_{\text{NFA}}(B,a)$)
= ε-Closure ({3,8})
= {1,2,3,4,6,7,8} = B

$\text{Move}_{\text{DFA}}(B,b)$
= ε-Closure ($\text{Move}_{\text{NFA}}(B,b)$)
= ε-Closure ({5,9})
= {1,2,4,5,6,7,9} = D



Example

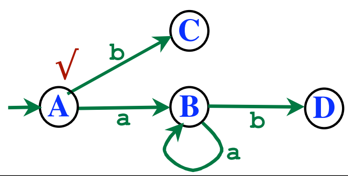
$\Sigma = \{\mathtt{a},\mathtt{b}\}$



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

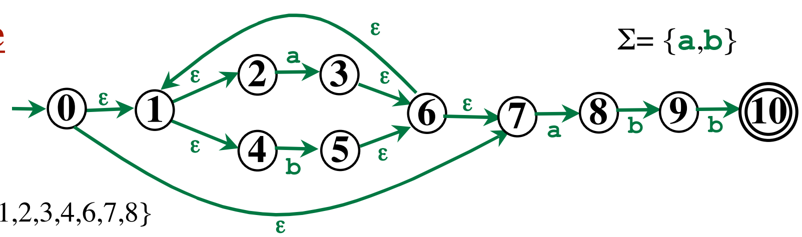
$\text{Move}_{\text{DFA}}(B,\mathtt{a})$
$= \varepsilon\text{-Closure}(\text{Move}_{\text{NFA}}(B,\mathtt{a}))$
$= \varepsilon\text{-Closure}(\{3,8\})$
$= \{1,2,3,4,6,7,8\} = B$

$\text{Move}_{\text{DFA}}(B,\mathtt{b})$
$= \varepsilon\text{-Closure}(\text{Move}_{\text{NFA}}(B,\mathtt{b}))$
$= \varepsilon\text{-Closure}(\{5,9\})$
$= \{1,2,4,5,6,7,9\} = D$



Example

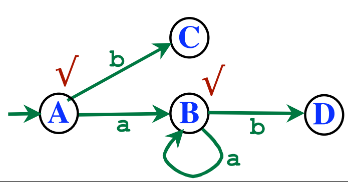
$\Sigma = \{\mathtt{a},\mathtt{b}\}$



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

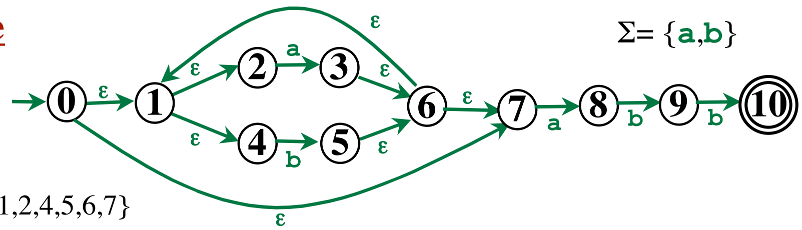
$\text{Move}_{\text{DFA}}(B,\mathtt{a})$
$= \varepsilon\text{-Closure}(\text{Move}_{\text{NFA}}(B,\mathtt{a}))$
$= \varepsilon\text{-Closure}(\{3,8\})$
$= \{1,2,3,4,6,7,8\} = B$

$\text{Move}_{\text{DFA}}(B,\mathtt{b})$
$= \varepsilon\text{-Closure}(\text{Move}_{\text{NFA}}(B,\mathtt{b}))$
$= \varepsilon\text{-Closure}(\{5,9\})$
$= \{1,2,4,5,6,7,9\} = D$

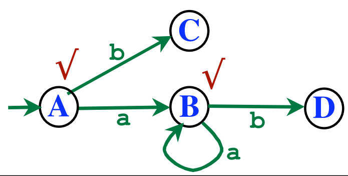


Example

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

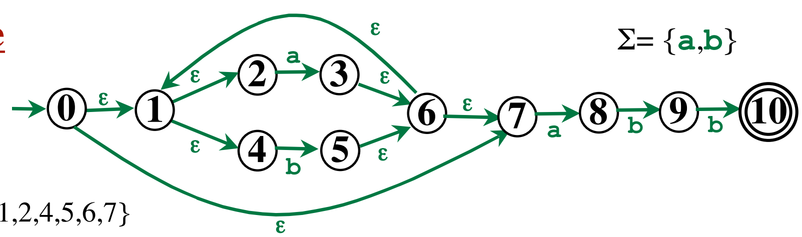


Process C = {1,2,4,5,6,7}



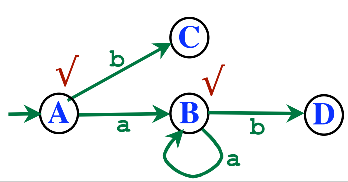
Example

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



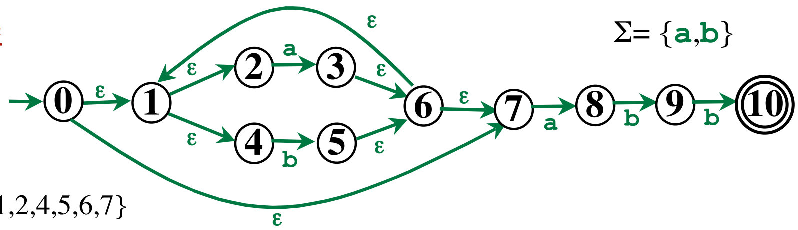
Process C = {1,2,4,5,6,7}

$Move_{DFA}(C,a) =$
$Move_{DFA}(C,b) =$



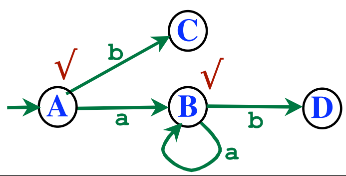
Example

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



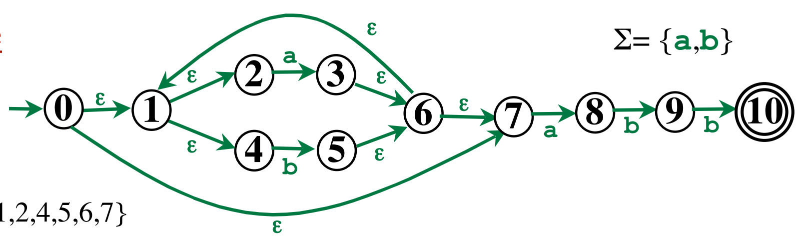
Process C = {1,2,4,5,6,7}

$Move_{DFA}$(C,a) = {1,2,3,4,6,7,8} = B
$Move_{DFA}$(C,b) =



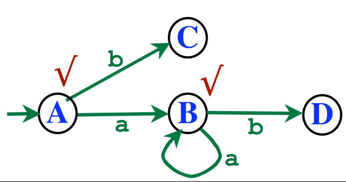
Example

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



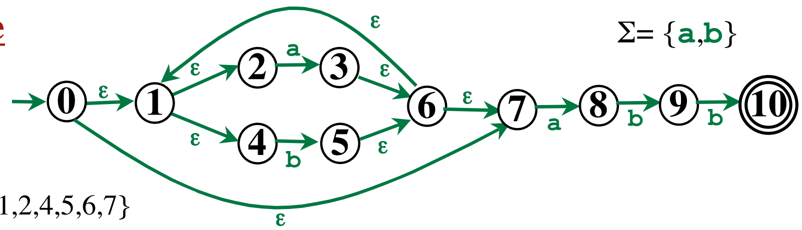
Process C = {1,2,4,5,6,7}

$Move_{DFA}$(C,a) = {1,2,3,4,6,7,8} = B
$Move_{DFA}$(C,b) = {1,2,4,5,6,7} = C



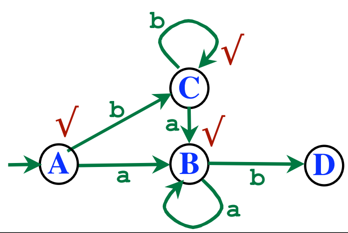
Example

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



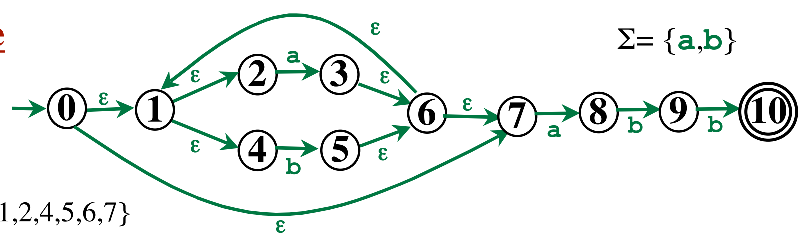
Process C = {1,2,4,5,6,7}

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



Example

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

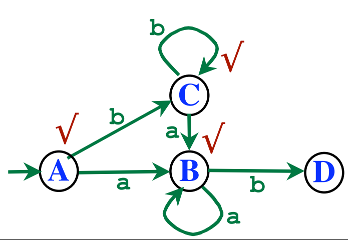


Process C = {1,2,4,5,6,7}

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

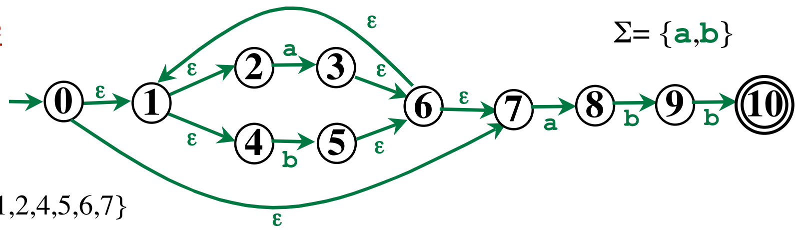
Process D = {1,2,4,5,6,7,9}

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



Example

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

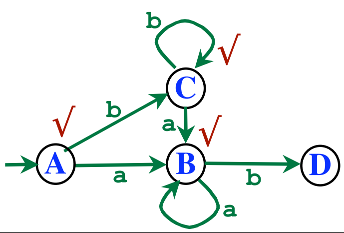


Process C = {1,2,4,5,6,7}

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

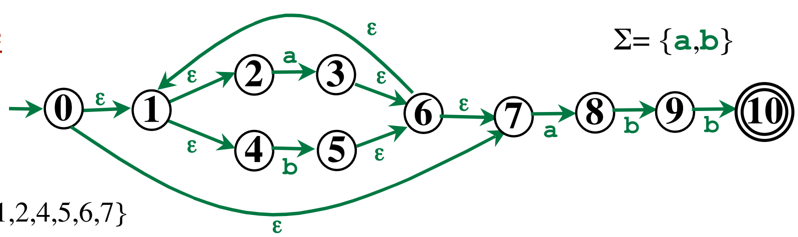
Process D = {1,2,4,5,6,7,9}

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



Example

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

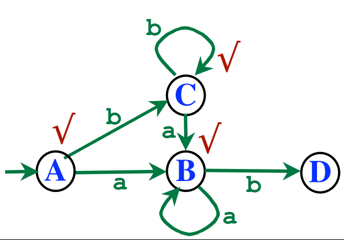


Process C = {1,2,4,5,6,7}

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

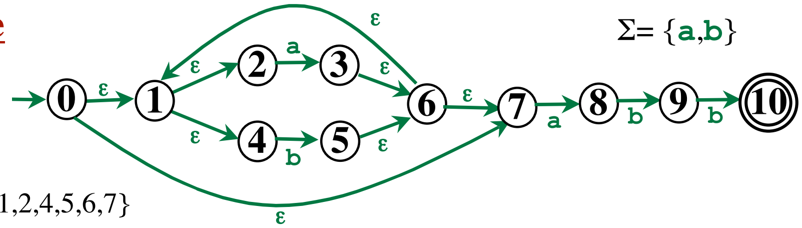
Process D = {1,2,4,5,6,7,9}

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



Example

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

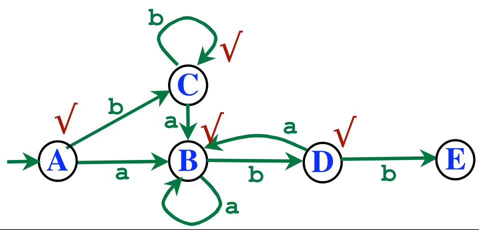


Process C = {1,2,4,5,6,7}

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

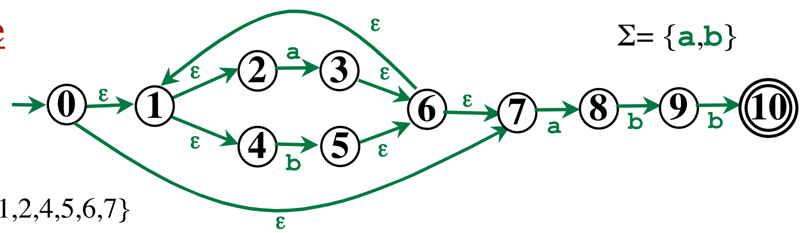
Process D = {1,2,4,5,6,7,9}

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



Example

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



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

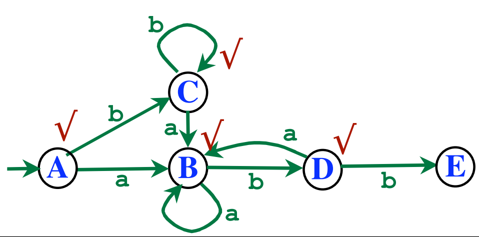
$\text{Move}_{\text{DFA}}(C,a) = \{1,2,3,4,6,7,8\} = B$
$\text{Move}_{\text{DFA}}(C,b) = \{1,2,4,5,6,7\} = C$

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

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

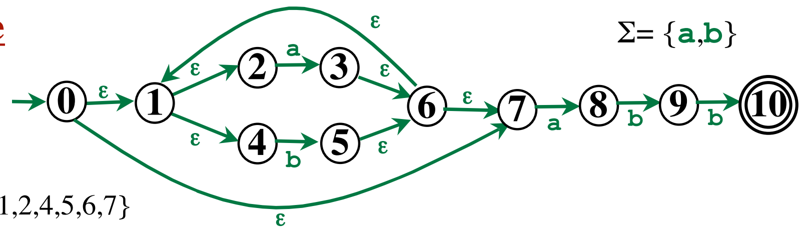
Process E = {1,2,4,5,6,7,10}

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



Example

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



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

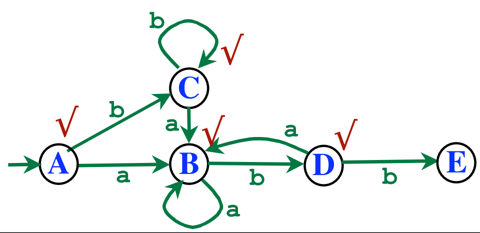
$Move_{DFA}$(C,**a**) = {1,2,3,4,6,7,8} = B
$Move_{DFA}$(C,**b**) = {1,2,4,5,6,7} = C

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

$Move_{DFA}$(D,**a**) = {1,2,3,4,6,7,8} = B
$Move_{DFA}$(D,**b**) = {1,2,4,5,6,7,10} = E

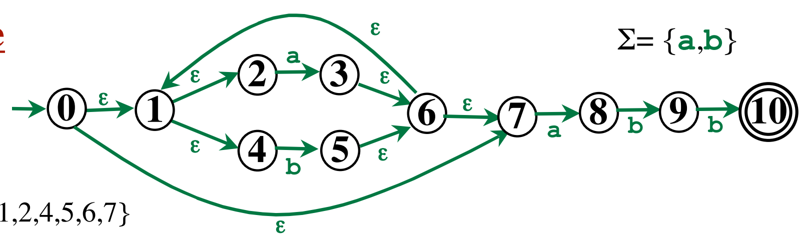
Process E = {1,2,4,5,6,7,10}

$Move_{DFA}$(E,**a**) = {1,2,3,4,6,7,8} = B
$Move_{DFA}$(E,**b**) =



Example

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



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

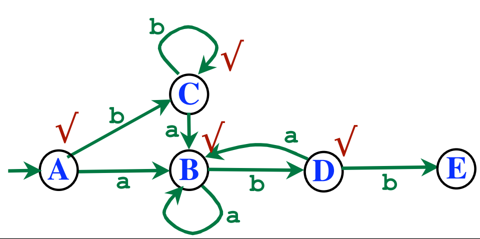
$Move_{DFA}$(C,**a**) = {1,2,3,4,6,7,8} = B
$Move_{DFA}$(C,**b**) = {1,2,4,5,6,7} = C

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

$Move_{DFA}$(D,**a**) = {1,2,3,4,6,7,8} = B
$Move_{DFA}$(D,**b**) = {1,2,4,5,6,7,10} = E

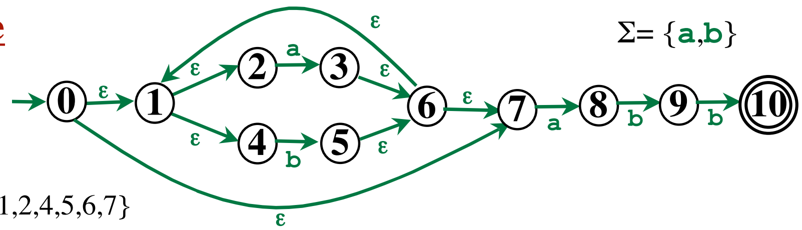
Process E = {1,2,4,5,6,7,10}

$Move_{DFA}$(E,**a**) = {1,2,3,4,6,7,8} = B
$Move_{DFA}$(E,**b**) = {1,2,4,5,6,7} = C



Example

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



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

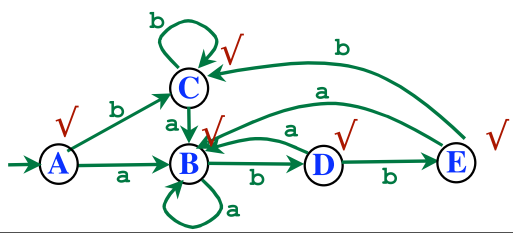
Move$_{DFA}$(C,a) = {1,2,3,4,6,7,8} = B
Move$_{DFA}$(C,b) = {1,2,4,5,6,7} = C

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

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

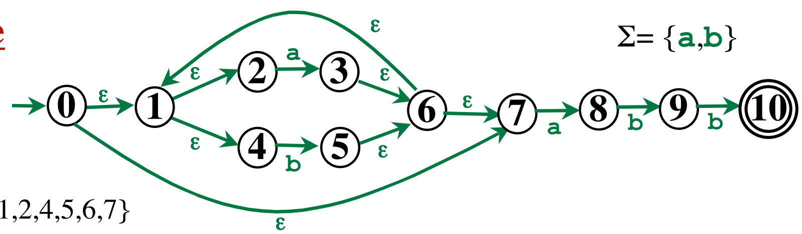
Process E = {1,2,4,5,6,7,10}

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



Example

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



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

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

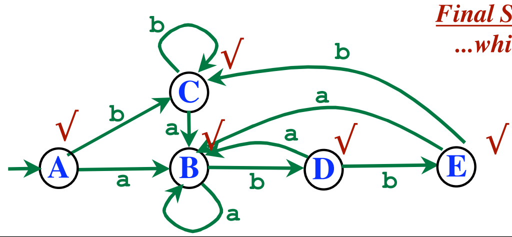
Process D = {1,2,4,5,6,7,9}

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

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

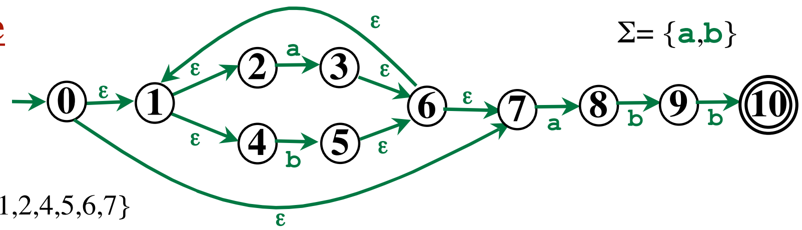
Move$_{DFA}$(E,a) = {1,2,3,4,6,7,8} = B
Move$_{DFA}$(E,b) = {1,2,4,5,6,7} = C

Final States in DFA?
...which state(s) contain 10?



Example

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



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

$Move_{DFA}(C,a) = \{1,2,3,4,6,7,8\} = B$
$Move_{DFA}(C,b) = \{1,2,4,5,6,7\} = C$

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

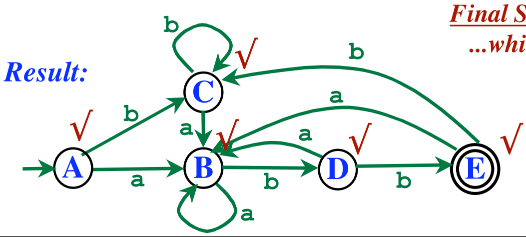
$Move_{DFA}(D,a) = \{1,2,3,4,6,7,8\} = B$
$Move_{DFA}(D,b) = \{1,2,4,5,6,7,10\} = E$

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

$Move_{DFA}(E,a) = \{1,2,3,4,6,7,8\} = B$
$Move_{DFA}(E,b) = \{1,2,4,5,6,7\} = C$

Final States in DFA?
...which state(s) contain 10?

Final Result:



Algorithm: Convert NFA to DFA

```
SDFA = {}
Add ε-Closure(s0) to SDFA as the start state
Set the only state in SDFA to "unmarked"
while SDFA contains an unmarked state do
  Let T be that unmarked state
  Mark T
  for each a in Σ do
    S = ε-Closure(MoveNFA(T,a))
    if S is not in SDFA already then
      Add S to SDFA (as an "unmarked" state)
    endIf
    Set MoveDFA(T,a) to S


  endFor
endWhile
for each S in SDFA do
  if any s∈S is a final state in the NFA then
    Mark S an a final state in the DFA
  endIf
endFor
```

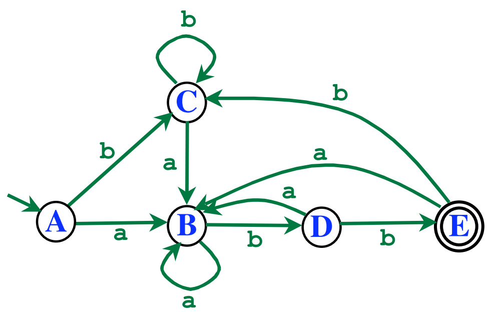
A set of NFA states

Everywhere you could possibly get to on an **a**

i.e, add an edge to the DFA...

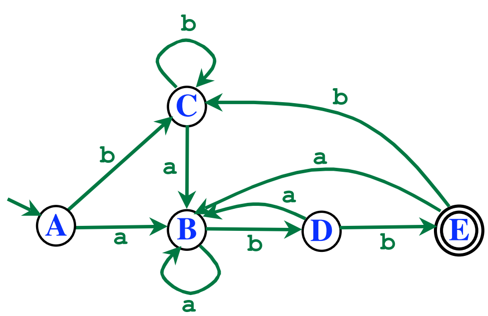
$\underset{T}{\{t_1,t_2,...\}} \xrightarrow{a} \underset{S}{\{s_1,s_2,...\}}$

Resulting DFA for **`(a|b)*abb`**

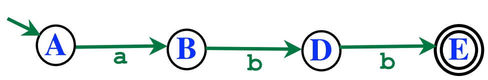


Is it minimal?

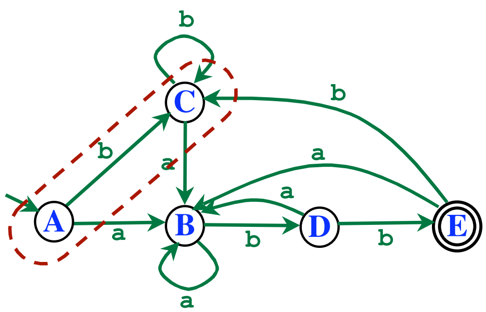
Resulting DFA for **`(a|b)*abb`**



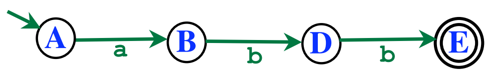
Is it minimal?



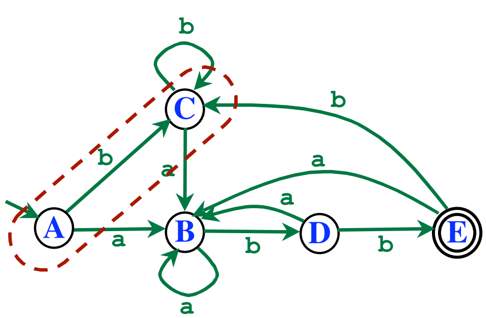
Resulting DFA for `(a|b)*abb`



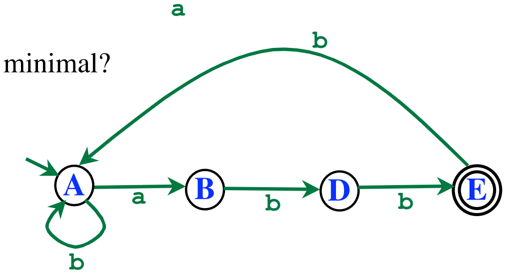
Is it minimal?



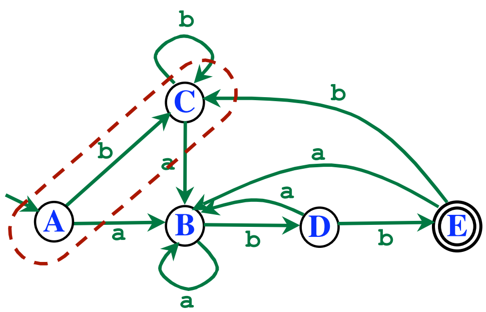
Resulting DFA for `(a|b)*abb`



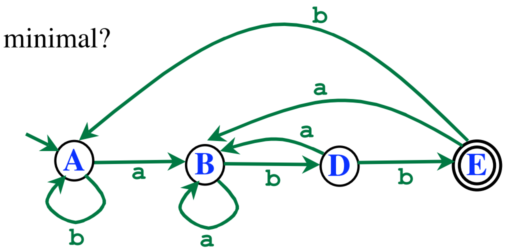
Is it minimal?



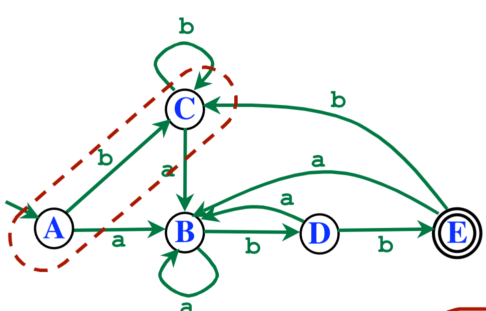
Resulting DFA for **`(a|b)*abb`**



Is it minimal?

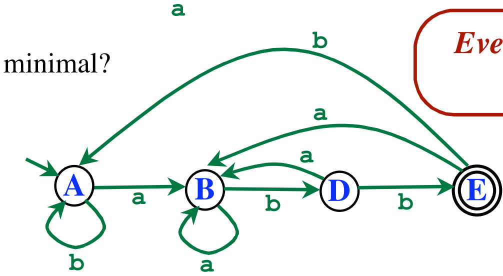


Resulting DFA for **`(a|b)*abb`**

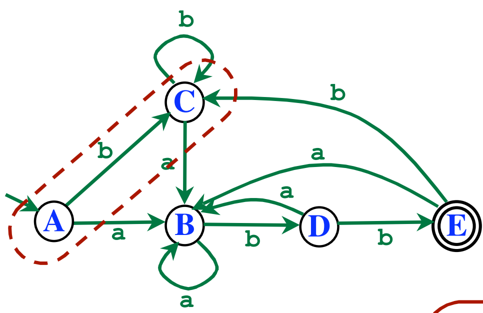


Is it minimal?

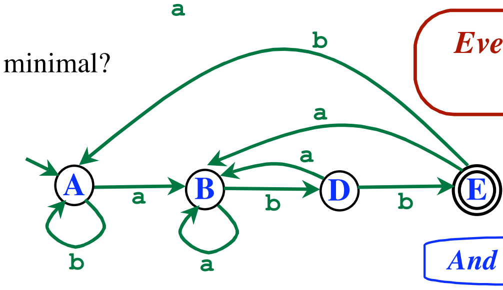
Every Regular Set is recognized by a minimal DFA!



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