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.

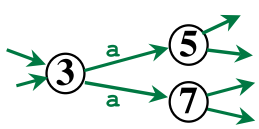
Lexical Analysis - Part 3

NFA to DFA

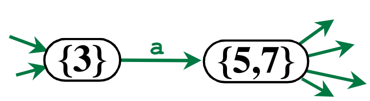
Input: A NFA
S = States = $\{ s_0, s_1, ..., s_N \} = S_{NFA}$
δ = Move function = $Move_{NFA}$
Move'(S, a) → Set of states

Output: A DFA
S = States = {?, ?, ..., ?} = 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} = { {...}, {...} , ..., {...} }

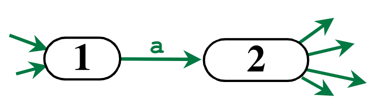


NFA



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.

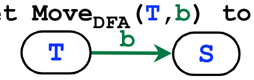
Marked = *"We've done this one"* (√)

Unmarked = *"Still need to do this one"*

`Move`$_{DFA}$`(T,b)` — The transition function from DFA

To add an edge to the growing DFA...

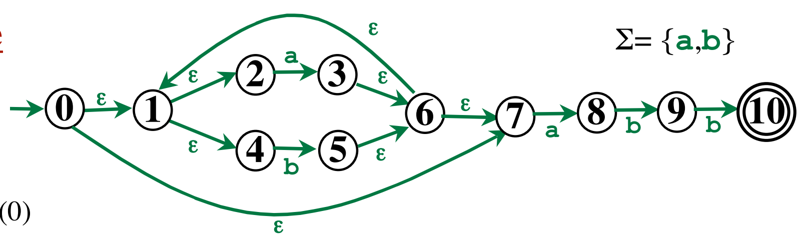
`Set Move`$_{DFA}$`(T,b) to S`



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

Lexical Analysis - Part 3

Example



Σ= {a,b}

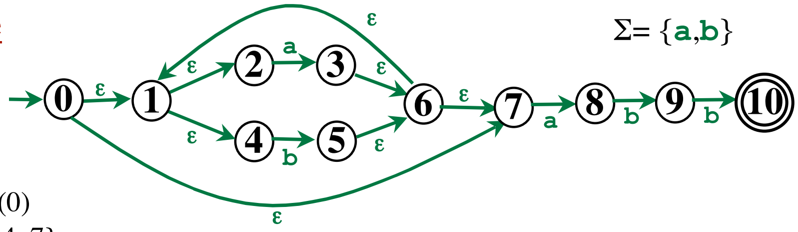
Start state:

ε-Closure (0)

=

Example

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



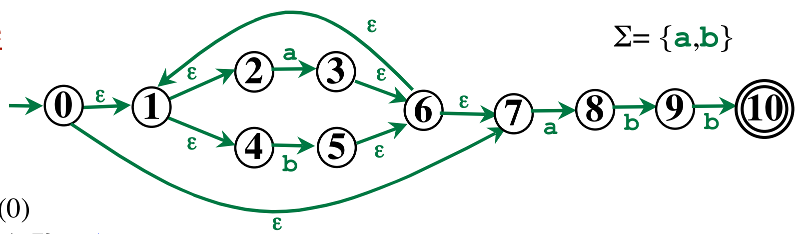
Start state:

ε-Closure (0)

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

Example

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



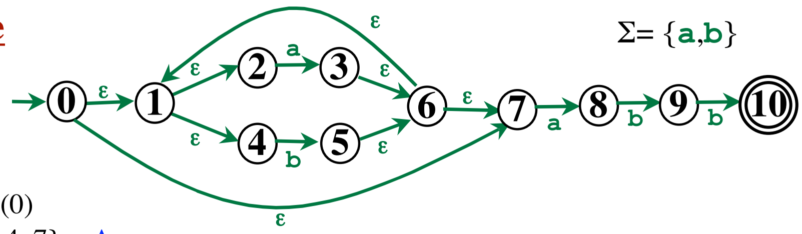
Start state:

ε-Closure (0)

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

Example

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



Start state:

ε-Closure (0)

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

$Move_{DFA}$(A,a)

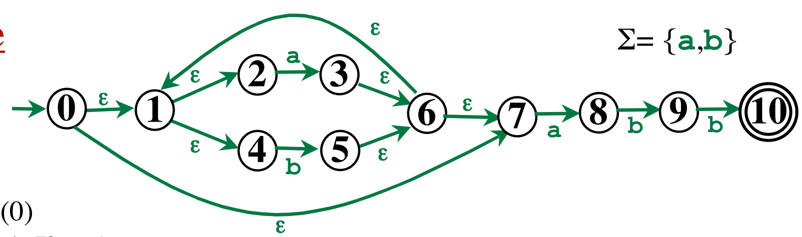
=

$Move_{DFA}$(A,b)

=

Example

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



Start state:

ε-Closure (0)

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

$Move_{DFA}$(A,a)

= ε-Closure ($Move_{NFA}$(A,a))

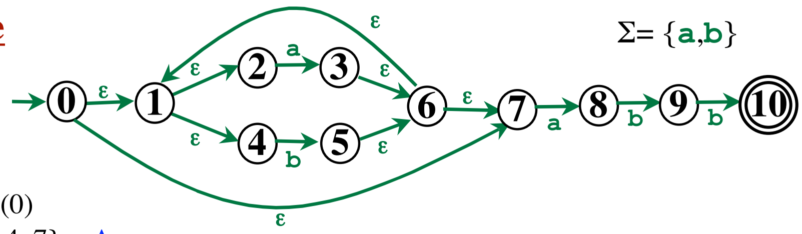
=

$Move_{DFA}$(A,b)

=

Example

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



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})

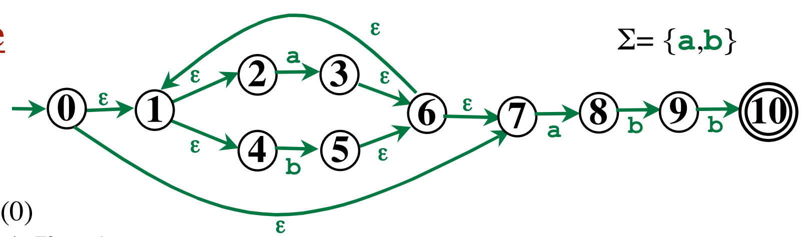
=

$Move_{DFA}(A, b)$

=

Example

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



Start state:

ε-Closure (0)

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

$Move_{DFA}(A, a)$

= ε-Closure ($Move_{NFA}(A, a)$)

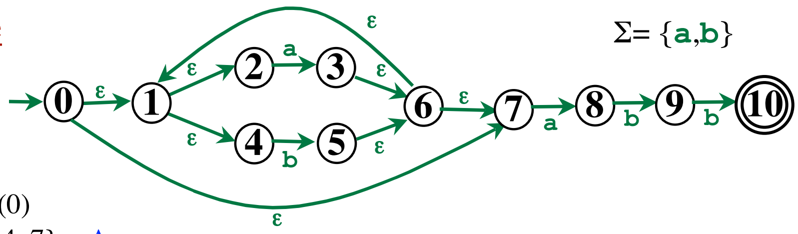
= ε-Closure ({3,8})

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

$Move_{DFA}(A, b)$

=

Example



Σ= {**a**,**b**}

Start state:

ε-Closure (0)

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

Move_{DFA}(A,**a**)

= ε-Closure (Move_{NFA}(A,**a**))

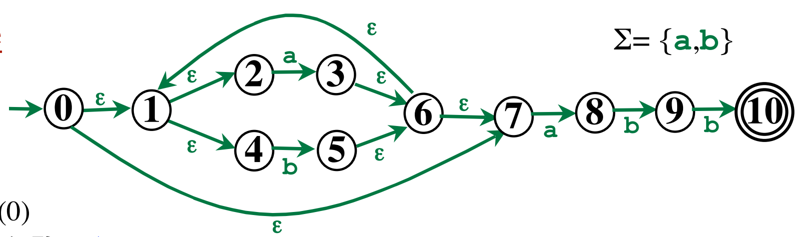
= ε-Closure ({3,8})

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

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})

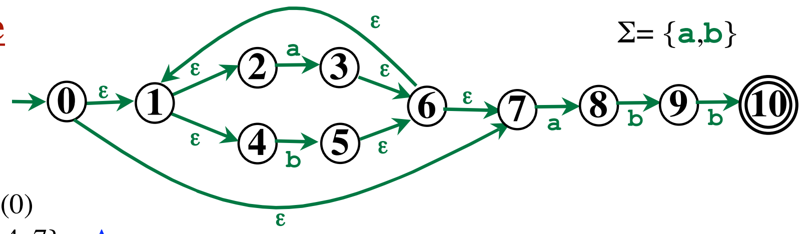
= {1,2,3,4,6,7,8} = B

Move_{DFA}(A,**b**)

= ε-Closure (Move_{NFA}(A,**b**))

=

Example



Σ= {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

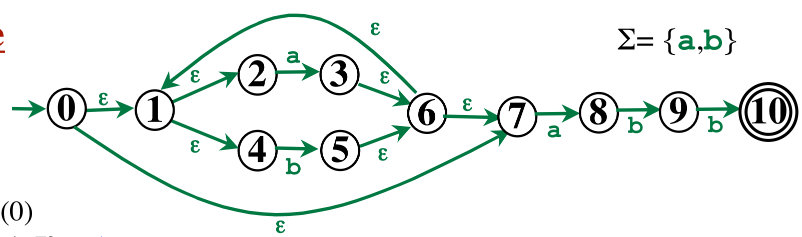
$\text{Move}_{\textbf{DFA}}$(A,b)

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

= ε-Closure ({5})

=

Example



Σ= {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)

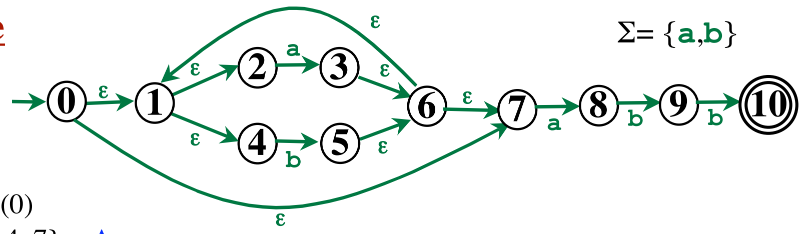
= ε-Closure ($\text{Move}_{\textbf{NFA}}$(A,b))

= ε-Closure ({5})

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

Example

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

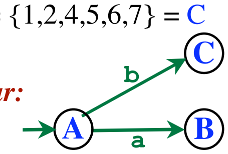


Start state:
ε-Closure (0)
= {0, 1, 2, 4, 7} = A

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

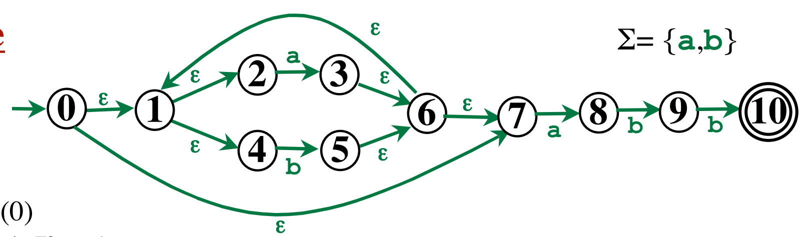
Move_{DFA}(A,b)
= ε-Closure (Move_{NFA}(A,b))
= ε-Closure ({5})
= {1,2,4,5,6,7} = C

So far:



Example

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

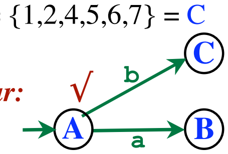


Start state:
ε-Closure (0)
= {0, 1, 2, 4, 7} = A

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

Move_{DFA}(A,b)
= ε-Closure (Move_{NFA}(A,b))
= ε-Closure ({5})
= {1,2,4,5,6,7} = C

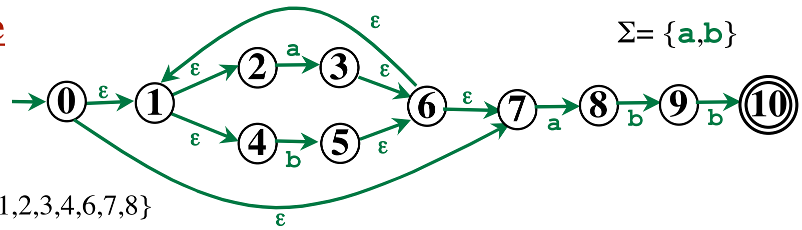
So far:



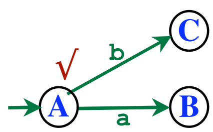
A is now done; mark it!
B and C are unmarked.
Let's do B next...

Example

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

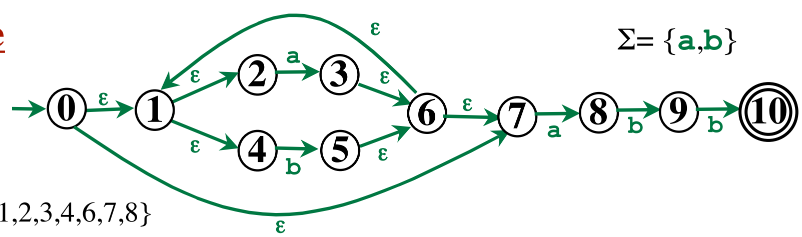


Process B = {1,2,3,4,6,7,8}



Example

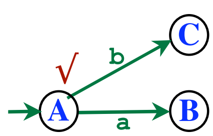
$\Sigma = \{a,b\}$



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

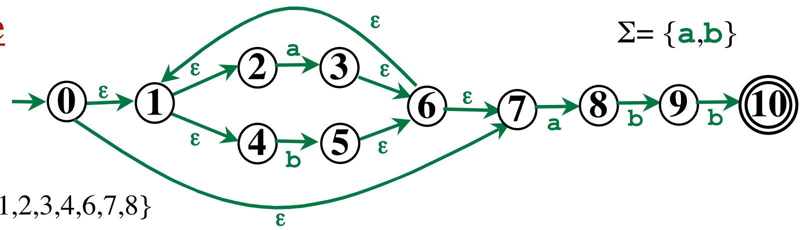
$Move_{DFA}(B,a)$

=



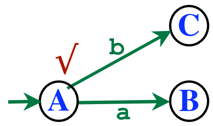
Example

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



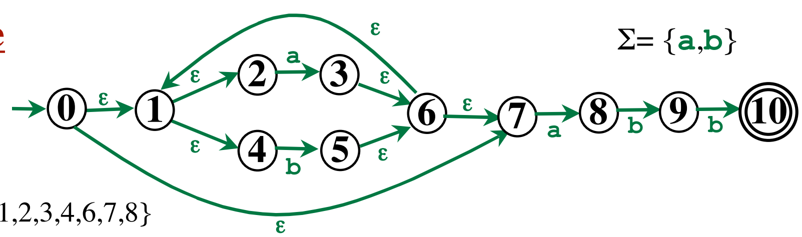
Process B = {1,2,3,4,6,7,8}

Move_{DFA}(B,**a**)
= ε-Closure (Move_{NFA}(B,**a**))
=



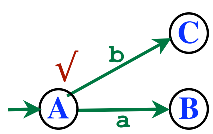
Example

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



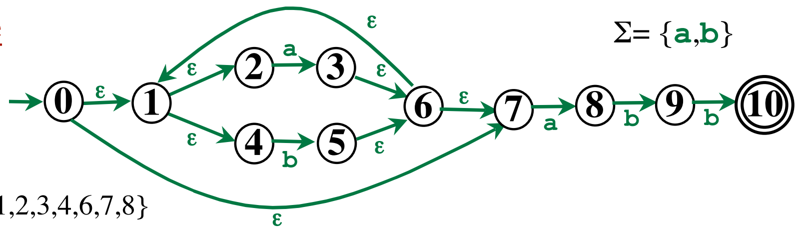
Process B = {1,2,3,4,6,7,8}

Move_{DFA}(B,**a**)
= ε-Closure (Move_{NFA}(B,**a**))
= ε-Closure ({3,8})
=



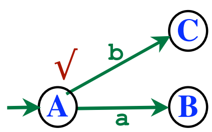
Example

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



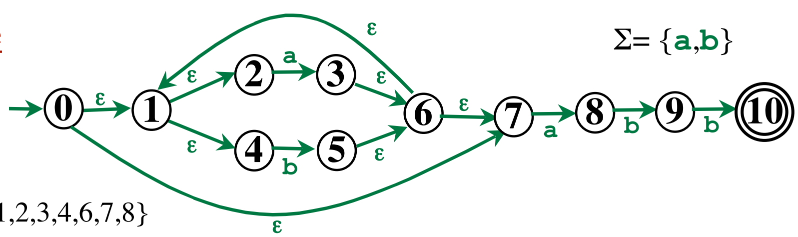
Process B = {1,2,3,4,6,7,8}

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



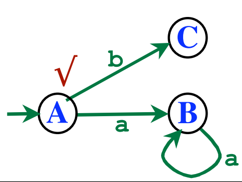
Example

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



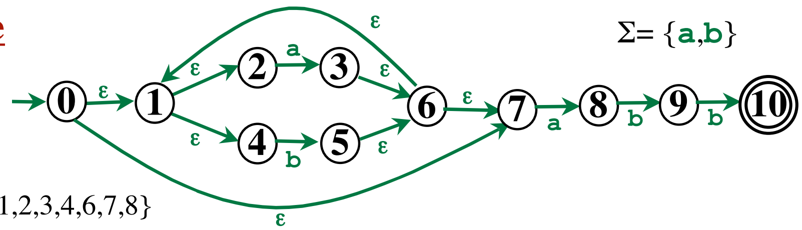
Process B = {1,2,3,4,6,7,8}

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



Example

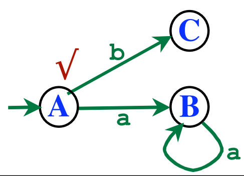
$\Sigma = \{a,b\}$



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

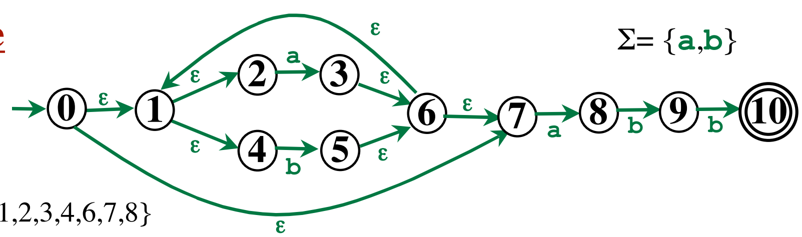
$\text{Move}_{DFA}(B,a)$
= ε-Closure ($\text{Move}_{NFA}(B,a)$)
= ε-Closure ({3,8})
= {1,2,3,4,6,7,8} = B

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



Example

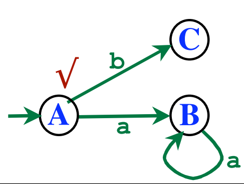
$\Sigma = \{a,b\}$



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

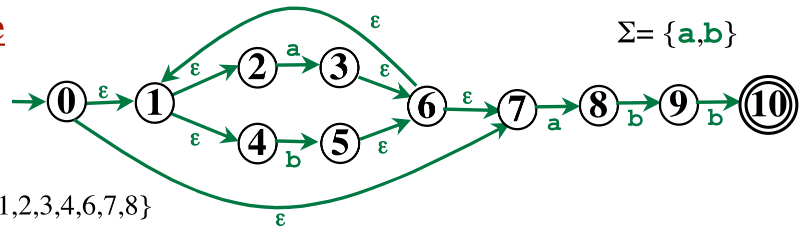
$\text{Move}_{DFA}(B,a)$
= ε-Closure ($\text{Move}_{NFA}(B,a)$)
= ε-Closure ({3,8})
= {1,2,3,4,6,7,8} = B

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



Example

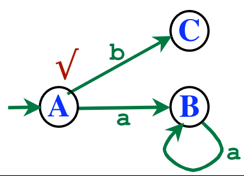
Σ= {a,b}



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

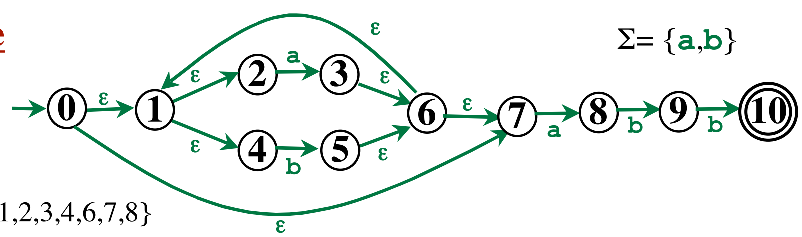
$\text{Move}_{\textbf{DFA}}$(B,a)
= ε-Closure ($\text{Move}_{\textbf{NFA}}$(B,a))
= ε-Closure ({3,8})
= {1,2,3,4,6,7,8} = B

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



Example

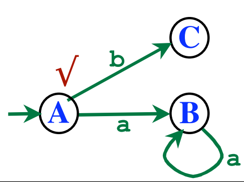
Σ= {a,b}



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

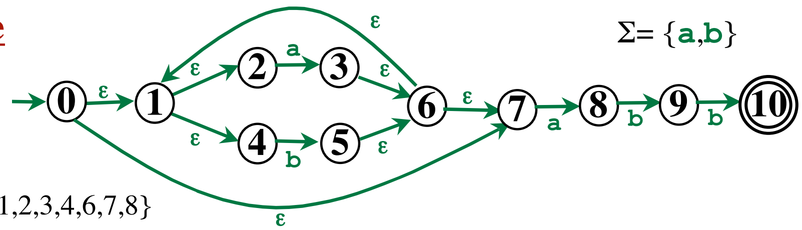
$\text{Move}_{\textbf{DFA}}$(B,a)
= ε-Closure ($\text{Move}_{\textbf{NFA}}$(B,a))
= ε-Closure ({3,8})
= {1,2,3,4,6,7,8} = B

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



Example

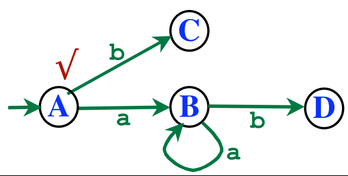
Σ= {a,b}



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

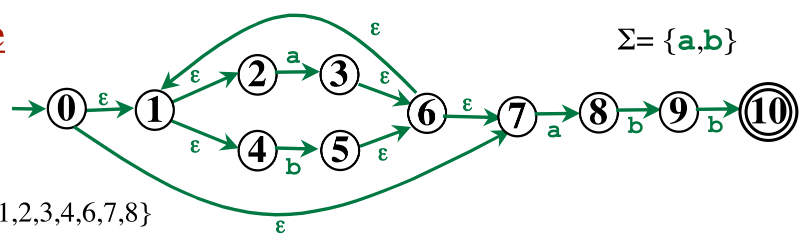
$Move_{DFA}$(B,a)
= ε-Closure ($Move_{NFA}$(B,a))
= ε-Closure ({3,8})
= {1,2,3,4,6,7,8} = B

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



Example

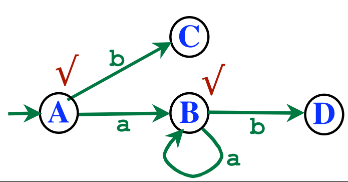
Σ= {a,b}



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

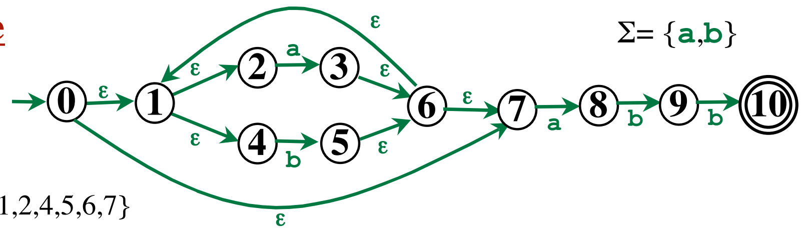
$Move_{DFA}$(B,a)
= ε-Closure ($Move_{NFA}$(B,a))
= ε-Closure ({3,8})
= {1,2,3,4,6,7,8} = B

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

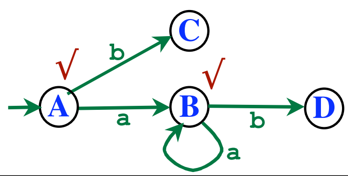


Example

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

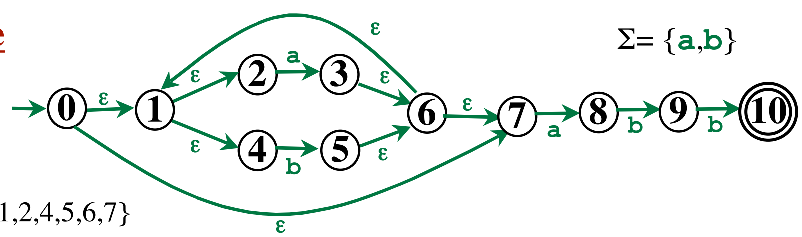


Process C = {1,2,4,5,6,7}



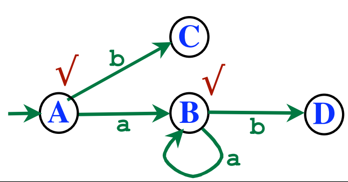
Example

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



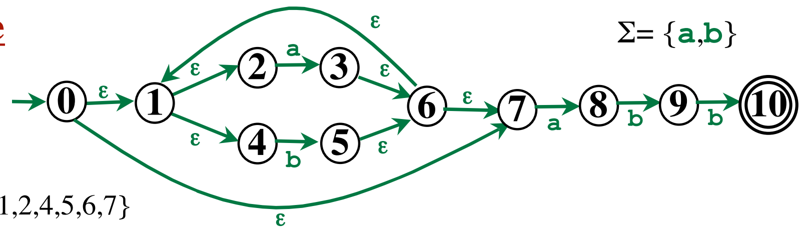
Process C = {1,2,4,5,6,7}

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



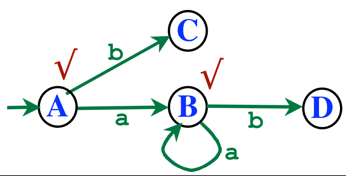
Example

Σ= {a,b}



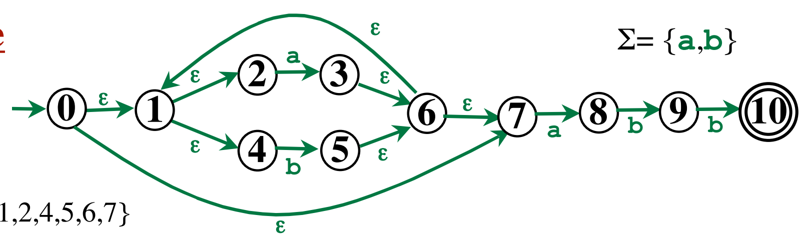
Process C = {1,2,4,5,6,7}

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



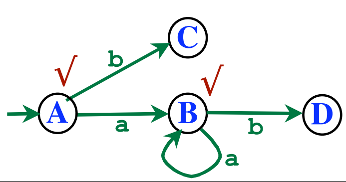
Example

Σ= {a,b}



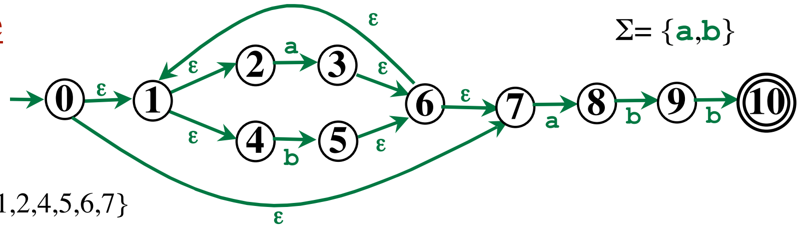
Process C = {1,2,4,5,6,7}

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



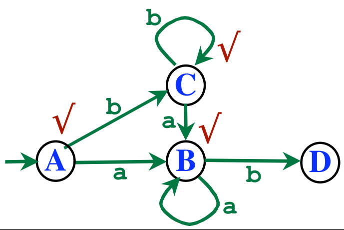
Example

Σ= {a,b}



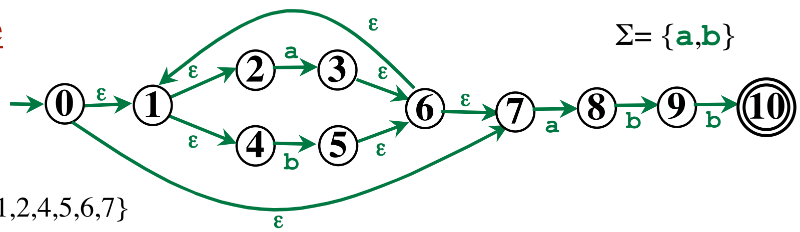
Process C = {1,2,4,5,6,7}

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



Example

Σ= {a,b}

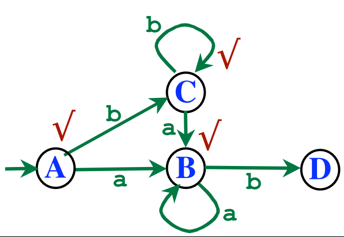


Process C = {1,2,4,5,6,7}

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

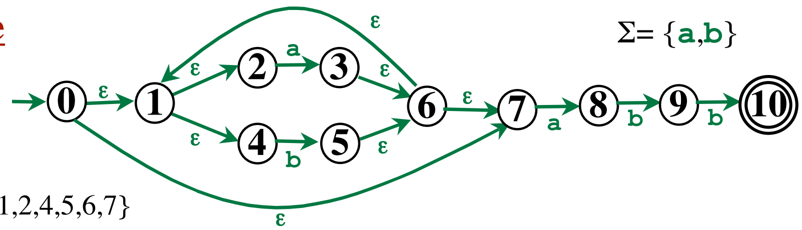
Process D = {1,2,4,5,6,7,9}

Move_{DFA}(D,a) =
Move_{DFA}(D,b) =



Example

Σ= {a,b}

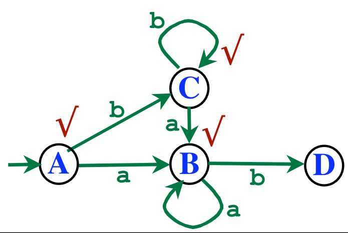


Process C = {1,2,4,5,6,7}

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

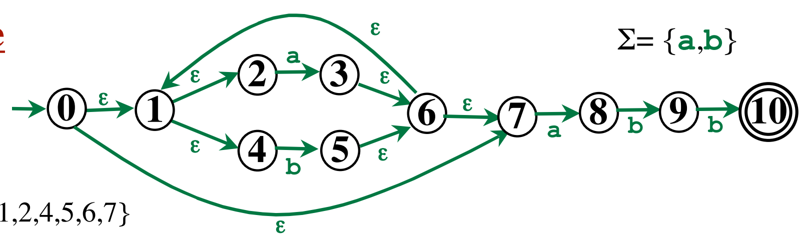
Process D = {1,2,4,5,6,7,9}

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



Example

Σ= {a,b}

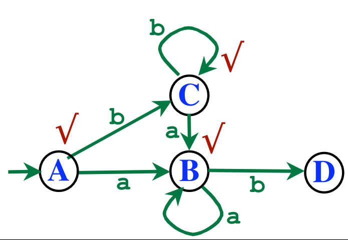


Process C = {1,2,4,5,6,7}

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

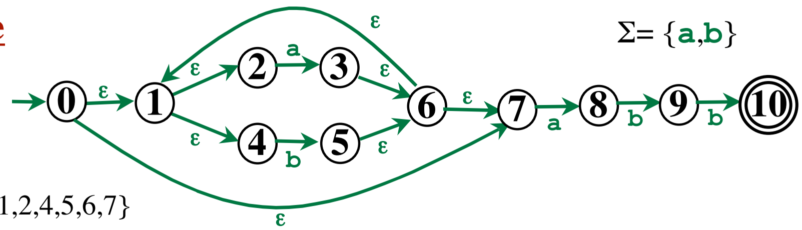
Process D = {1,2,4,5,6,7,9}

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



Example

Σ= {a,b}

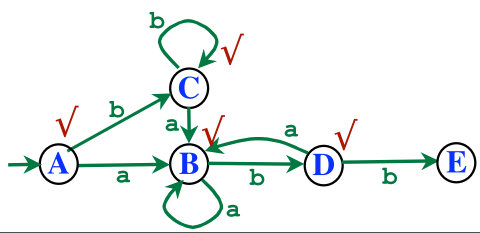


Process C = {1,2,4,5,6,7}

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

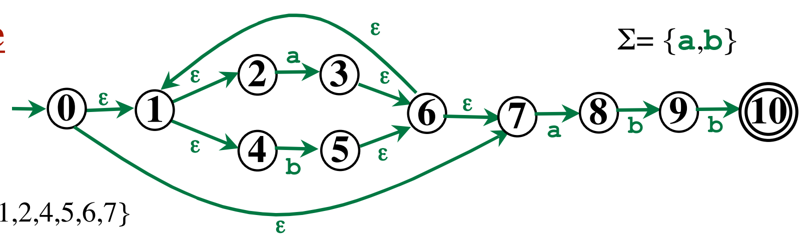
Process D = {1,2,4,5,6,7,9}

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



Example

Σ= {a,b}



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

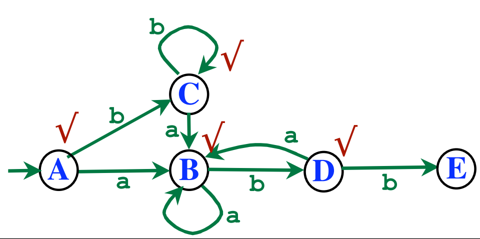
$Move_{DFA}$(C,a) = {1,2,3,4,6,7,8} = B
$Move_{DFA}$(C,b) = {1,2,4,5,6,7} = C

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

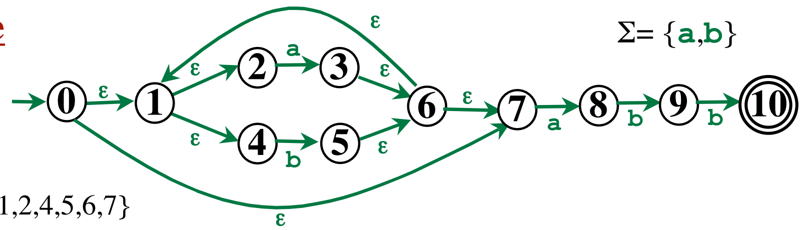
Process E = {1,2,4,5,6,7,10}

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



Example

Σ= {a,b}



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

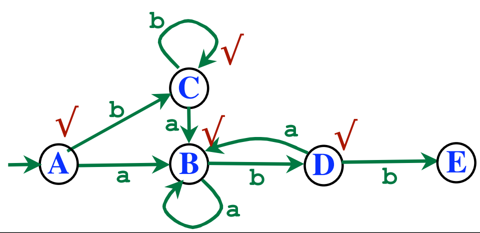
Move$_{DFA}$(C,a) = {1,2,3,4,6,7,8} = B
Move$_{DFA}$(C,b) = {1,2,4,5,6,7} = C

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

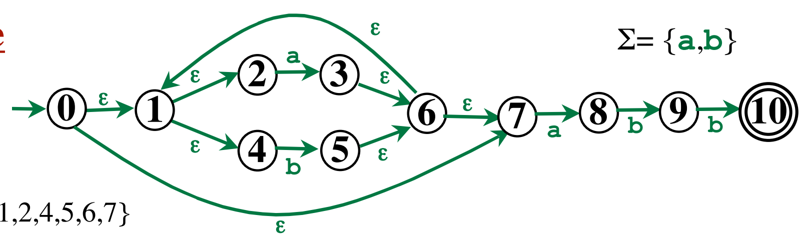
Process E = {1,2,4,5,6,7,10}

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



Example

Σ= {a,b}



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

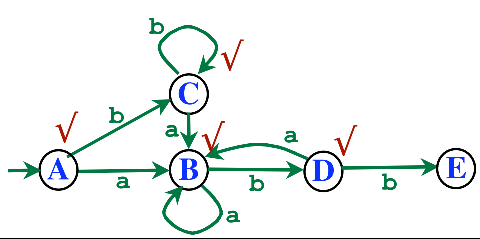
Move$_{DFA}$(C,a) = {1,2,3,4,6,7,8} = B
Move$_{DFA}$(C,b) = {1,2,4,5,6,7} = C

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

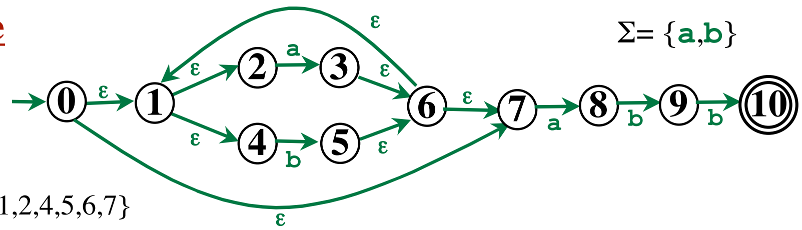
Process E = {1,2,4,5,6,7,10}

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



Example

Σ= {a,b}



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

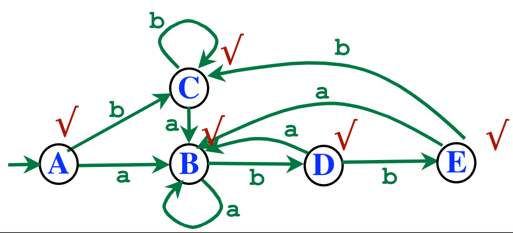
Move$_{DFA}$(C,a) = {1,2,3,4,6,7,8} = B
Move$_{DFA}$(C,b) = {1,2,4,5,6,7} = C

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}

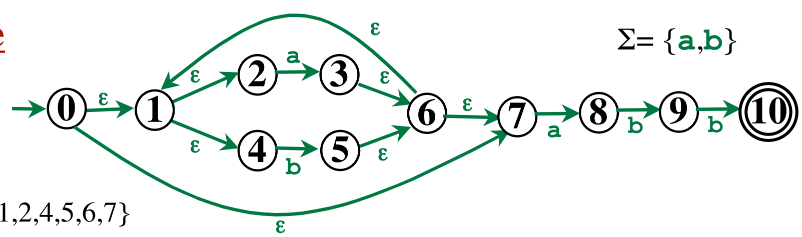
Move$_{DFA}$(E,a) = {1,2,3,4,6,7,8} = B
Move$_{DFA}$(E,b) = {1,2,4,5,6,7} = C



Lexical Analysis - Part 3

Example

Σ= {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

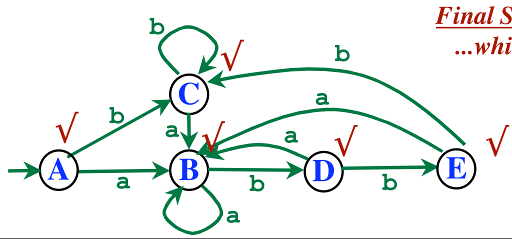
Process D = {1,2,4,5,6,7,9}

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}

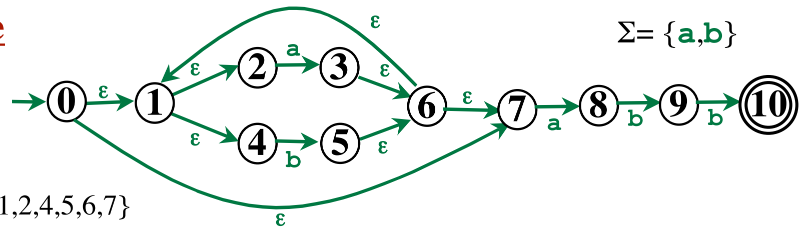
Move$_{DFA}$(E,a) = {1,2,3,4,6,7,8} = B
Move$_{DFA}$(E,b) = {1,2,4,5,6,7} = C

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

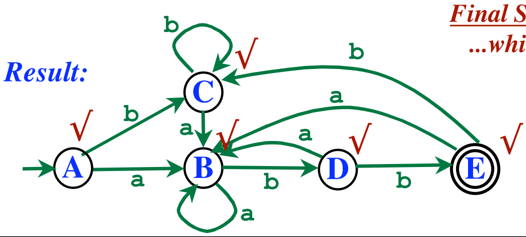
Process D = {1,2,4,5,6,7,9}

$Move_{DFA}$(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 Result:



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

Algorithm: Convert NFA to DFA

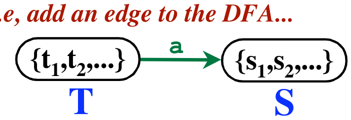
```
S_DFA = {}
Add ε-Closure(s_0) to S_DFA as the start state
Set the only state in S_DFA to "unmarked"
while S_DFA contains an unmarked state do
  Let T be that unmarked state
  Mark T
  for each a in Σ do
    S = ε-Closure(Move_NFA(T,a))
    if S is not in S_DFA already then
      Add S to S_DFA (as an "unmarked" state)
    endIf
    Set Move_DFA(T,a) to S


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

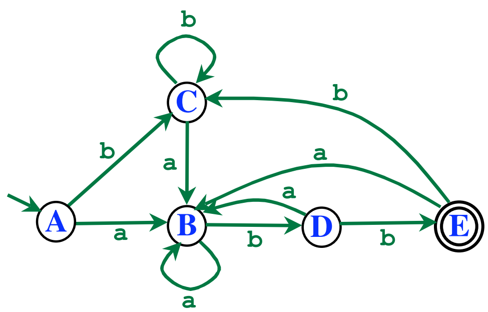
A set of NFA states

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

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

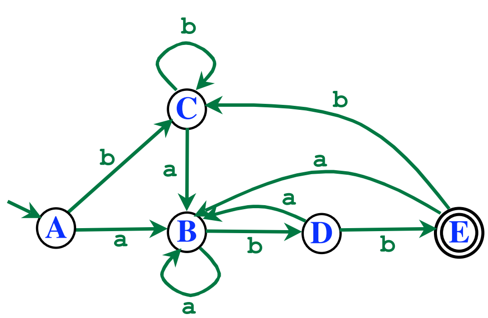


Resulting DFA for **(a|b)*abb**

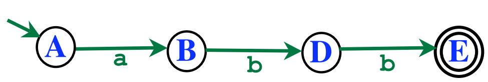


Is it minimal?

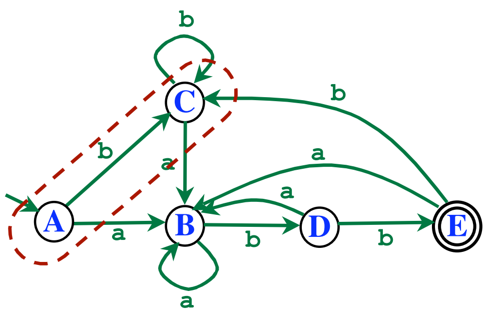
Resulting DFA for **(a|b)*abb**



Is it minimal?



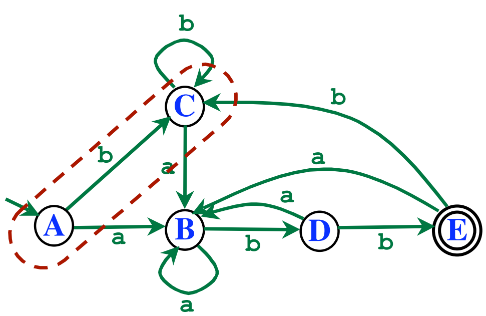
Resulting DFA for **(a|b)*abb**



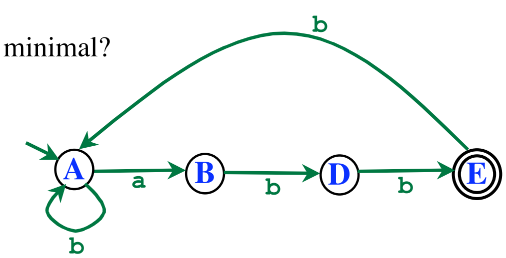
Is it minimal?

Lexical Analysis - Part 3

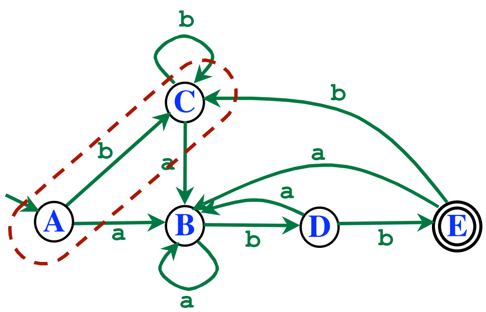
Resulting DFA for **(a|b)*abb**



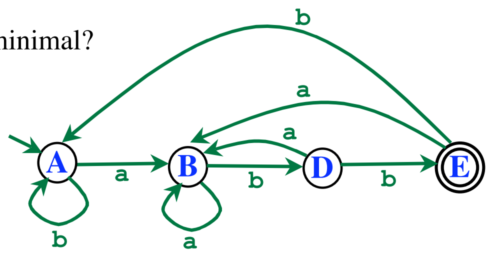
Is it minimal?



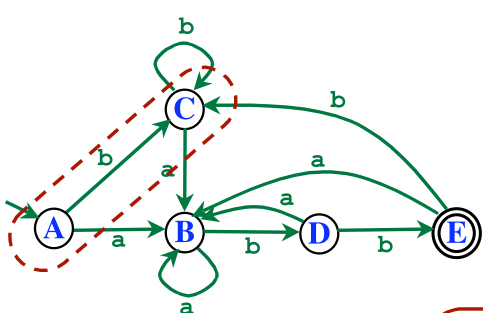
Resulting DFA for **`(a|b)*abb`**



Is it minimal?

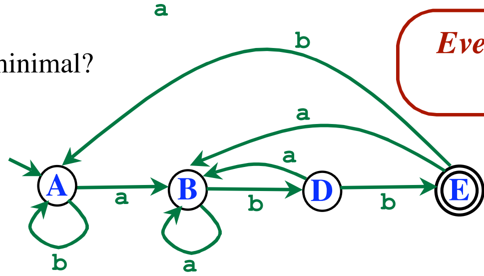


Resulting DFA for **`(a|b)*abb`**

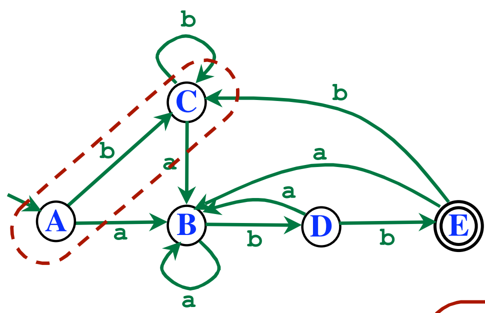


Is it minimal?

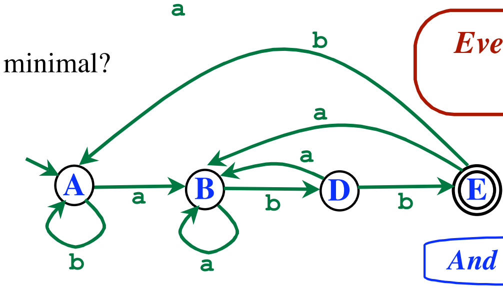
Every Regular Set is recognized by a minimal DFA!



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