Lecture 9 - June 3

Lexical Analysis

Subset Construction: Algorithm
ε-NFA: Motivation, Example, δ Function
Regular Expression to ε-NFA

Announcements/Reminders

- $\boxed{\text{Assignment 1 released}}$ → RE, ε-NFA
- Review Slides on Math posted
- Optional, if you want to pre-study for A2 and Project:
  ANTLR4 tutorial videos

Subset Construction: Algorithmic Specification

Q of input NFA : $\{q_0, q_1, q_2\}$

$\rightarrow Q$: max # of states in the output DFA $\rightarrow$ # subsets from Q_N

Given an *NFA* $N = (Q_N, \Sigma_N, \delta_N, q_0, F_N)$:

ALGORITHM : *ReachableSubsetStates* — lazy evaluation.

INPUT : $q_0 : Q_N$; **OUTPUT :** $Reachable \subseteq \mathbb{P}(Q_N)$

PROCEDURE :

$Reachable := \{ \{q_0\} \}$ ✓

$ToDiscover := \{ \{q_0\} \}$

while ($ToDiscover \neq \varnothing$) { — as soon as ToDiscover becomes ϕ, exit

* choose $S : \mathbb{P}(Q_N)$ such that $S \in ToDiscover$

remove S from $ToDiscover$

** $NotYetDiscovered :=$

$(\{ \{\delta_N(s, 0) \mid s \in S\} \} \cup \{ \{\delta_N(s, 1) \mid s \in S\} \}) \setminus Reachable$

(q_0: $\{q_0, q_1\}$; $\{q_0\}$: q_0) — set difference

$Reachable := Reachable \cup NotYetDiscovered$

$ToDiscover := ToDiscover \cup NotYetDiscovered$

} $\hookrightarrow \phi \rightarrow \{\{q_0, q_1\}\}$

$\{\{q_0, q_1\}, \{q_0\}\}$

return $Reachable$

state \ input	0	1
$\{q_0\}$	$\{q_0, q_1\}$	$\{q_0\}$
$\{q_0, q_1\}$	$\{q_0, q_1\}$	$\{q_0, q_2\}$
$\{q_0, q_2\}$	$\{q_0, q_1\}$	$\{q_0\}$

* 1st iteration : choose $S = \{q_0\}$, ToDiscover becomes ϕ

** NotYetDiscovered : $\{\{q_0, q_1\}\} \cup \{\{q_0\}\} \setminus \{\{q_0\}\} = \{\{q_0, q_1\}\}$

$$\underbrace{\{\zeta_0, \zeta_1\}}_{\text{proc. 0}} \cup \underbrace{\{\zeta_0\}}_{\text{proc. 1}} = \{\zeta_0, \zeta_1\}$$

vs.

$$\{\underline{\{\zeta_0, \zeta_1\}}\} \cup \{\underline{\{\zeta_0\}}\}$$

$$= \{\{\zeta_0, \zeta_1\}, \{\zeta_0\}\}$$

$\boxed{Q_{NFA}} = \{q_0, q_1, q_2\}$

$|Q_{DFA}| \le 2^{|Q_N|}$

theoretical upper bound

practically, may not reach there

Instead, use lazy evaluation.

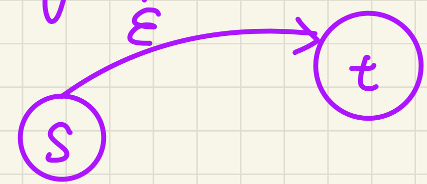
	q_0	q_1	q_2	
0	0	0	0	$\emptyset$
1	0	0	1	$\{q_2\}$
	0	1	0	$\{q_1\}$
⋮		⋮		⋮ reachable
6	1	1	0	$\boxed{\{q_0, q_1\}}$
7	1	1	1	$\boxed{\{q_0, q_1, q_2\}}$

not reachable.

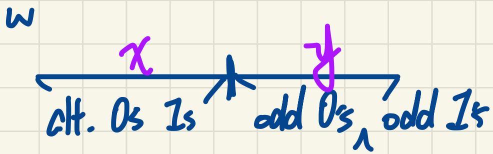
epsilon-NFA: Motivation

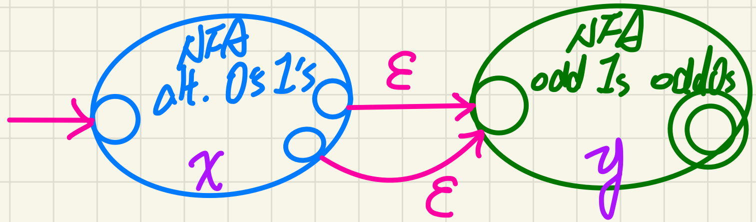
allows an ε transition to another state without reading any input

Draw NFA

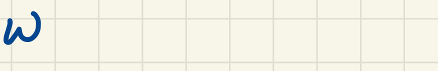


$$\left\{ xy \;\middle|\; \begin{array}{ll} & x \in \{0,1\}^* \\ \wedge & y \in \{0,1\}^* \\ \wedge & x \text{ has alternating } 0\text{'s and } 1\text{'s} \\ \wedge & y \text{ has an odd \# } 0\text{'s and an odd \# } 1\text{'s} \end{array} \right\}$$

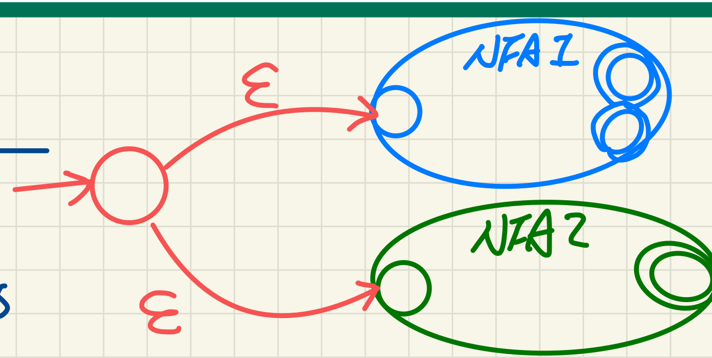




$$\left\{ w : \{0,1\}^* \;\middle|\; \begin{array}{ll} & w \text{ has alternating } 0\text{'s and } 1\text{'s} \\ \vee & w \text{ has an odd \# } 0\text{'s and an odd \# } 1\text{'s} \end{array} \right\}$$



alt 0s 1's
∨
odd 0s ∧ odd 1s

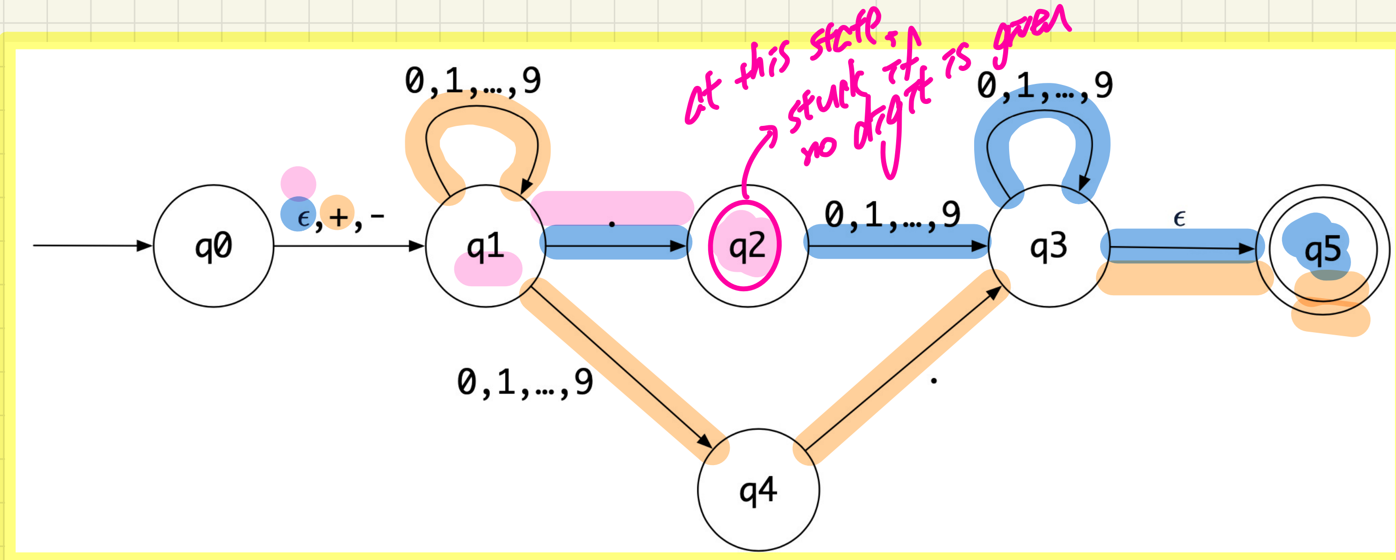


epsilon-NFA: Example

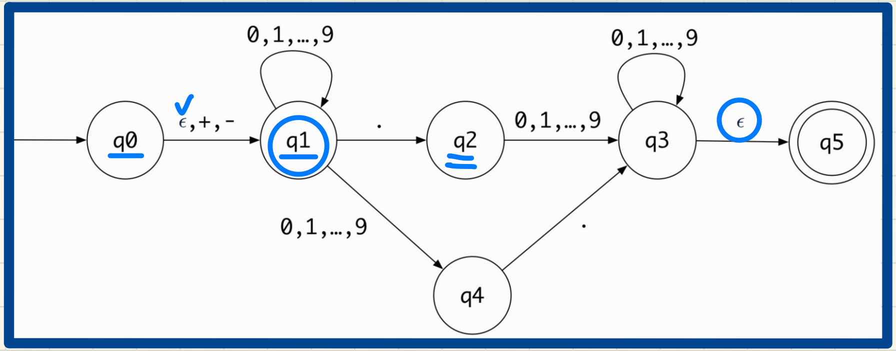
$$\left\{ sx.y \;\middle|\; \begin{array}{ll} & s \in \{+, -, \epsilon\} \\ \wedge & x \in \Sigma_{dec}^{*} \\ \wedge & y \in \Sigma_{dec}^{*} \\ \wedge & \neg(x = \epsilon \wedge y = \epsilon) \end{array} \right\}$$

Is this a DFA? N.
Is this an NFA? N.
Is this an ε-NFA? Y

.23 ✓
+46. ✓
. ✗



epsilon-NFA: Formulation (1)



An ϵ-*NFA* is a 5-tuple

$$M = (Q, \Sigma, \boxed{\delta}, q_0, F)$$

Draw the transition table.

$$(Q \times \Sigma \cup \{\epsilon\}) \to \mathbb{P}Q$$

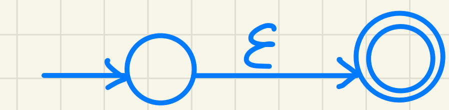
$$\epsilon \notin \Sigma$$

	ϵ	+, -	.	0 .. 9
q_0	$\{q_1\}$	$\{q_1\}$	$\varnothing$	$\varnothing$
q_1	$\varnothing$	$\varnothing$	$\{q_2\}$	$\{q_1, q_4\}$
q_2	$\varnothing$	$\varnothing$	$\varnothing$	$\{q_3\}$
q_3	$\{q_5\}$	$\varnothing$	$\varnothing$	$\{q_3\}$
q_4	$\varnothing$	$\varnothing$	$\{q_3\}$	$\varnothing$
q_5	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$

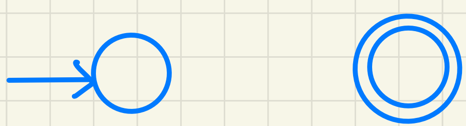
Regular Expression to epsilon-NFA

Base Cases

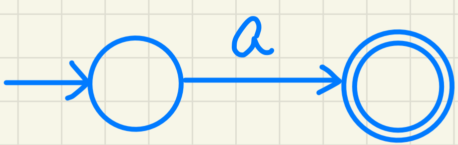
$\boxed{\varepsilon}$ $L(\varepsilon) = \{\varepsilon\}$



$\emptyset$ $L(\emptyset) = \emptyset$

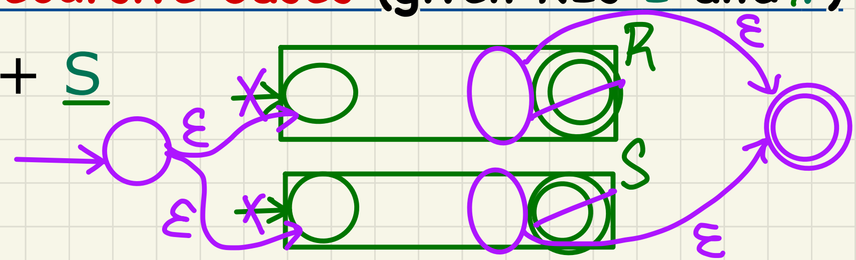


a $\quad (a \in \Sigma)$ $\quad 0 \in \{0, 1\}$

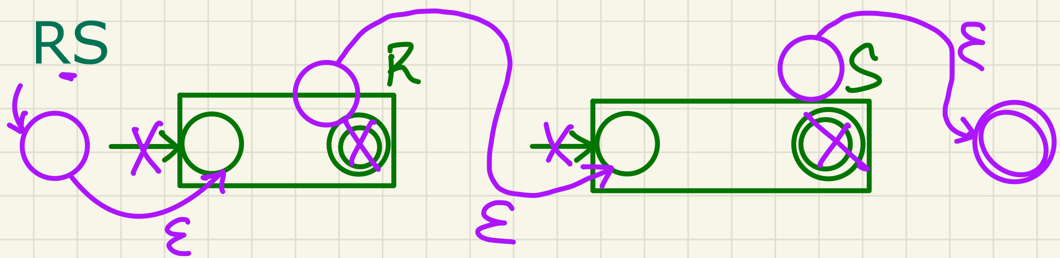


Recursive Cases (given REs ~~E~~ R and ~~F~~ S)

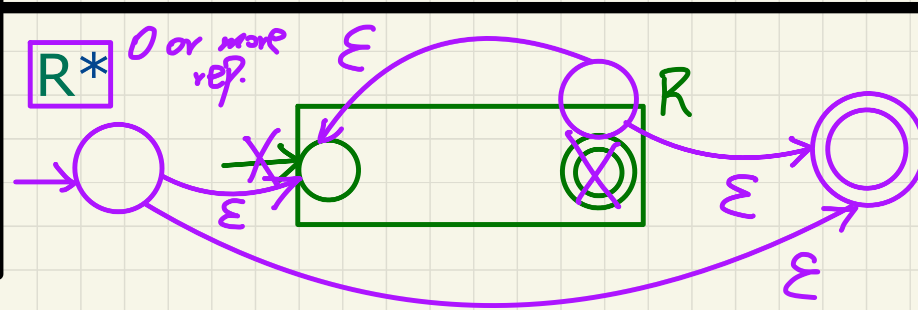
R + S



RS



$\boxed{R^*}$ 0 or more rep.



Regular Expression to epsilon-NFA: Example

$(0+1)^*1(0+1)$