Regular Expressions

- DFAs are finite descriptions of (finite or infinite) sets of strings

- DFAs are finite descriptions of (finite or infinite) sets of strings
  - Finite number of symbols, states, transitions

Regular Expressions

- DFAs are finite descriptions of (finite or infinite) sets of strings
  - Finite number of symbols, states, transitions
- Regular Expressions provide an algebraic expression framework to describe the same class of strings

Regular Expressions

- DFAs are finite descriptions of (finite or infinite) sets of strings
  - Finite number of symbols, states, transitions
- Regular Expressions provide an algebraic expression framework to describe the same class of strings
- Thus, DFAs and Regular Expressions are equivalent.

Regular Expressions

- For every regular expression, there is a corresponding regular set or language

Regular Expressions

- For every regular expression, there is a corresponding regular set or language
- R, R_1, R_2 are regular expressions; $L(R)$ denotes the corresponding regular set

Regular Expressions

- For every regular expression, there is a corresponding regular set or language
- R, R_1, R_2 are regular expressions; $L(R)$ denotes the corresponding regular set

Regular Expression	**Regular Set**

Regular Expressions

- For every regular expression, there is a corresponding regular set or language
- R, R_1, R_2 are regular expressions; $L(R)$ denotes the corresponding regular set

Regular Expression	**Regular Set**
ϕ	$\{\}$

Regular Expressions

- For every regular expression, there is a corresponding regular set or language
- R, R_1, R_2 are regular expressions; $L(R)$ denotes the corresponding regular set

Regular Expression	**Regular Set**
ϕ	$\{\}$
a for a $\in \Sigma$	$\{a\}$

REGULAR EXPRESSIONS

- For every regular expression, there is a corresponding regular set or language
- R, R_1, R_2 are regular expressions; $L(R)$ denotes the corresponding regular set

Regular Expression	**Regular Set**
ϕ	$\{\}$
a for a $\in \Sigma$	$\{a\}$
ϵ	$\{\epsilon\}$

Regular Expressions

- For every regular expression, there is a corresponding regular set or language
- R, R_1, R_2 are regular expressions; $L(R)$ denotes the corresponding regular set

Regular Expression	**Regular Set**
ϕ	$\{\}$
a for a $\in \Sigma$	$\{a\}$
ϵ	$\{\epsilon\}$
$(R_1 \cup R_2)$	$L(R_1) \cup L(R_2)$

Regular Expressions

- For every regular expression, there is a corresponding regular set or language
- R, R_1, R_2 are regular expressions; $L(R)$ denotes the corresponding regular set

Regular Expression	**Regular Set**
ϕ	$\{\}$
a for a $\in \Sigma$	$\{a\}$
ϵ	$\{\epsilon\}$
$(R_1 \cup R_2)$	$L(R_1) \cup L(R_2)$
$(R_1 R_2)$	$L(R_1)L(R_2)$

Regular Expressions

- For every regular expression, there is a corresponding regular set or language
- R, R_1, R_2 are regular expressions; $L(R)$ denotes the corresponding regular set

Regular Expression	**Regular Set**
ϕ	$\{\}$
a for a $\in \Sigma$	$\{a\}$
ϵ	$\{\epsilon\}$
$(R_1 \cup R_2)$	$L(R_1) \cup L(R_2)$
$(R_1 R_2)$	$L(R_1)L(R_2)$
(R^*)	$L(R)^*$

Regular Expression	Regular Set
ϕ	$\{\}$
a for $\in \Sigma$	$\{a\}$
ϵ	$\{\epsilon\}$
$(R_1 \cup R_2)$	$L(R_1) \cup L(R_2)$
$(R_1 \circ R_2)$	$L(R_1) \circ L(R_2)$
(R^*)	$L(R)^*$

- Also some books use $R_1 + R_2$ to denote union.

Regular Expressions– More Syntax

Regular Expression	Regular Set
ϕ	$\{\}$
a for $\in \Sigma$	$\{a\}$
ϵ	$\{\epsilon\}$
$(R_1 \cup R_2)$	$L(R_1) \cup L(R_2)$
$(R_1 \circ R_2)$	$L(R_1) \circ L(R_2)$
(R^*)	$L(R)^*$

- Also some books use $R_1 + R_2$ to denote union.
- In $(\ldots)$, the parenthesis can be deleted

Regular Expressions– More Syntax

Regular Expression	Regular Set
ϕ	$\{\}$
a for $\in \Sigma$	$\{a\}$
ϵ	$\{\epsilon\}$
$(R_1 \cup R_2)$	$L(R_1) \cup L(R_2)$
$(R_1 \circ R_2)$	$L(R_1) \circ L(R_2)$
(R^*)	$L(R)^*$

- Also some books use $R_1 + R_2$ to denote union.
- In $(\ldots)$, the parenthesis can be deleted
  - In which case, interpretation is done in the precedence order: star, concatenation and then union.

Regular Expressions– More Syntax

Regular Expression	Regular Set
ϕ	$\{\}$
a for $\in \Sigma$	$\{a\}$
ϵ	$\{\epsilon\}$
$(R_1 \cup R_2)$	$L(R_1) \cup L(R_2)$
$(R_1 \circ R_2)$	$L(R_1) \circ L(R_2)$
(R^*)	$L(R)^*$

- Also some books use $R_1 + R_2$ to denote union.
- In $(\ldots)$, the parenthesis can be deleted
  - In which case, interpretation is done in the precedence order: star, concatenation and then union.
- $R^+ = RR^*$ and R^k for k-fold concatenation are useful shorthands.

Regular Expression Examples

Regular Expression		Regular Language
0^*10^*	$\rightarrow$	

Regular Expression Examples

Regular Expression		**Regular Language**
$\mathbf{0}^*\mathbf{1}\mathbf{0}^*$	$\rightarrow$	$\{\omega \mid \omega$ contains a single $1\}$

Regular Expression Examples

Regular Expression		Regular Language
$\mathbf{0}^*\mathbf{1}\mathbf{0}^*$	$\rightarrow$	$\{\omega \mid \omega$ contains a single $\mathsf{1}\}$
$(\mathbf{0} \cup \mathbf{1})^*\mathbf{1}(\mathbf{0} \cup \mathbf{1})^*$	$\rightarrow$	

Regular Expression Examples

Regular Expression		Regular Language
$\mathbf{0}^*\mathbf{10}^*$	$\rightarrow$	$\{\omega \mid \omega$ contains a single 1$\}$
$(\mathbf{0} \cup \mathbf{1})^*\mathbf{1}(\mathbf{0} \cup \mathbf{1})^*$	$\rightarrow$	$\{\omega \mid \omega$ has at least one 1$\}$

Regular Expression Examples

Regular Expression		Regular Language
$\mathbf{0}^*\mathbf{1}\mathbf{0}^*$	$\rightarrow$	$\{\omega \mid \omega$ contains a single 1$\}$
$(\mathbf{0} \cup \mathbf{1})^*\mathbf{1}(\mathbf{0} \cup \mathbf{1})^*$	$\rightarrow$	$\{\omega \mid \omega$ has at least one 1$\}$
$\mathbf{0}(\mathbf{0} \cup \mathbf{1})^*\mathbf{0} \cup \mathbf{1}(\mathbf{0} \cup \mathbf{1})^*\mathbf{1} \cup \mathbf{0} \cup \mathbf{1}$	$\rightarrow$	

Regular Expression Examples

Regular Expression		Regular Language
$\mathbf{0}^*\mathbf{1}\mathbf{0}^*$	$\rightarrow$	$\{\omega \mid \omega$ contains a single $\mathsf{1}\}$
$(\mathbf{0} \cup \mathbf{1})^*\mathbf{1}(\mathbf{0} \cup \mathbf{1})^*$	$\rightarrow$	$\{\omega \mid \omega$ has at least one $\mathsf{1}\}$
$\mathbf{0}(\mathbf{0} \cup \mathbf{1})^*\mathbf{0} \cup \mathbf{1}(\mathbf{0} \cup \mathbf{1})^*\mathbf{1} \cup \mathbf{0} \cup \mathbf{1}$	$\rightarrow$	$\{\omega \mid \omega$ starts and ends with the same symbol$\}$

REGULAR EXPRESSION EXAMPLES

Regular Expression		Regular Language
$\mathbf{0}^*\mathbf{1}\mathbf{0}^*$	$\rightarrow$	$\{\omega \mid \omega$ contains a single $1\}$
$(\mathbf{0} \cup \mathbf{1})^*\mathbf{1}(\mathbf{0} \cup \mathbf{1})^*$	$\rightarrow$	$\{\omega \mid \omega$ has at least one $1\}$
$\mathbf{0}(\mathbf{0} \cup \mathbf{1})^*\mathbf{0} \cup \mathbf{1}(\mathbf{0} \cup \mathbf{1})^*\mathbf{1} \cup \mathbf{0} \cup \mathbf{1}$	$\rightarrow$	$\{\omega \mid \omega$ starts and ends with the same symbol$\}$
$(\mathbf{0}^*\mathbf{1}\mathbf{0}^*\mathbf{1})^*\mathbf{0}^*$	$\rightarrow$	

Regular Expression Examples

Regular Expression		Regular Language
0^*10^*	$\rightarrow$	$\{\omega \mid \omega$ contains a single $1\}$
$(0 \cup 1)^*1(0 \cup 1)^*$	$\rightarrow$	$\{\omega \mid \omega$ has at least one $1\}$
$0(0 \cup 1)^*0 \cup 1(0 \cup 1)^*1 \cup 0 \cup 1$	$\rightarrow$	$\{\omega \mid \omega$ starts and ends with the same symbol$\}$
$(0^*10^*1)^*0^*$	$\rightarrow$	$\{\omega \mid n_1(\omega)$ is even$\}$

- All strings with at least one pair of consecutive 0s

Writing Regular Expressions

- All strings with at least one pair of consecutive 0s
  - $(\mathbf{0} \cup \mathbf{1})^*\mathbf{00}(\mathbf{0} \cup \mathbf{1})^*$

Writing Regular Expressions

- All strings with at least one pair of consecutive 0s
  - $(\mathbf{0} \cup \mathbf{1})^*\mathbf{00}(\mathbf{0} \cup \mathbf{1})^*$
- All strings such that fourth symbol from the end is a 1

Writing Regular Expressions

- All strings with at least one pair of consecutive 0s
  - $(\mathbf{0} \cup \mathbf{1})^* \mathbf{00} (\mathbf{0} \cup \mathbf{1})^*$
- All strings such that fourth symbol from the end is a 1
  - $(\mathbf{0} \cup \mathbf{1})^* \mathbf{1} (\mathbf{0} \cup \mathbf{1})(\mathbf{0} \cup \mathbf{1})(\mathbf{0} \cup \mathbf{1})$

Writing Regular Expressions

- All strings with at least one pair of consecutive 0s
  - $(\mathbf{0} \cup \mathbf{1})^*\mathbf{00}(\mathbf{0} \cup \mathbf{1})^*$
- All strings such that fourth symbol from the end is a 1
  - $(\mathbf{0} \cup \mathbf{1})^*\mathbf{1}(\mathbf{0} \cup \mathbf{1})(\mathbf{0} \cup \mathbf{1})(\mathbf{0} \cup \mathbf{1})$
- All strings with no pair of consecutive 0s

Writing Regular Expressions

- All strings with at least one pair of consecutive 0s
  - $(\mathbf{0} \cup \mathbf{1})^*\mathbf{00}(\mathbf{0} \cup \mathbf{1})^*$
- All strings such that fourth symbol from the end is a 1
  - $(\mathbf{0} \cup \mathbf{1})^*\mathbf{1}(\mathbf{0} \cup \mathbf{1})(\mathbf{0} \cup \mathbf{1})(\mathbf{0} \cup \mathbf{1})$
- All strings with no pair of consecutive 0s
  - $(\mathbf{1}^*\mathbf{011}^*)^*(\mathbf{0} \cup \epsilon) \cup \mathbf{1}*$

Writing Regular Expressions

- All strings with at least one pair of consecutive 0s
  - $(\mathbf{0} \cup \mathbf{1})^*\mathbf{00}(\mathbf{0} \cup \mathbf{1})^*$
- All strings such that fourth symbol from the end is a 1
  - $(\mathbf{0} \cup \mathbf{1})^*\mathbf{1}(\mathbf{0} \cup \mathbf{1})(\mathbf{0} \cup \mathbf{1})(\mathbf{0} \cup \mathbf{1})$
- All strings with no pair of consecutive 0s
  - $(\mathbf{1}^*\mathbf{011}^*)^*(\mathbf{0} \cup \epsilon) \cup \mathbf{1}*$
  - Strings consist of repetitions of 1 or 01 or two boundary cases: $(\mathbf{1} \cup \mathbf{01})^*(\mathbf{0} \cup \epsilon)$

Writing Regular Expressions

- All strings with at least one pair of consecutive 0s
  - $(\mathbf{0} \cup \mathbf{1})^*\mathbf{00}(\mathbf{0} \cup \mathbf{1})^*$
- All strings such that fourth symbol from the end is a 1
  - $(\mathbf{0} \cup \mathbf{1})^*\mathbf{1}(\mathbf{0} \cup \mathbf{1})(\mathbf{0} \cup \mathbf{1})(\mathbf{0} \cup \mathbf{1})$
- All strings with no pair of consecutive 0s
  - $(\mathbf{1}^*\mathbf{011}^*)^*(\mathbf{0} \cup \epsilon) \cup \mathbf{1}*$
  - Strings consist of repetitions of 1 or 01 or two boundary cases: $(\mathbf{1} \cup \mathbf{01})^*(\mathbf{0} \cup \epsilon)$
- All strings that do not end in 01.

Writing Regular Expressions

- All strings with at least one pair of consecutive 0s
  - $(\mathbf{0} \cup \mathbf{1})^*\mathbf{00}(\mathbf{0} \cup \mathbf{1})^*$
- All strings such that fourth symbol from the end is a 1
  - $(\mathbf{0} \cup \mathbf{1})^*\mathbf{1}(\mathbf{0} \cup \mathbf{1})(\mathbf{0} \cup \mathbf{1})(\mathbf{0} \cup \mathbf{1})$
- All strings with no pair of consecutive 0s
  - $(\mathbf{1}^*\mathbf{011}^*)^*(\mathbf{0} \cup \epsilon) \cup \mathbf{1}*$
  - Strings consist of repetitions of 1 or 01 or two boundary cases: $(\mathbf{1} \cup \mathbf{01})^*(\mathbf{0} \cup \epsilon)$
- All strings that do not end in 01.
  - $(\mathbf{0} \cup \mathbf{1})^*(\mathbf{00} \cup \mathbf{10} \cup \mathbf{11}) \cup \mathbf{0} \cup \mathbf{1} \cup \epsilon$

RE IDENTITIES

- $\mathbf{R} \cup \phi = \mathbf{R}$

RE Identities

- $\mathbf{R} \cup \phi = \mathbf{R}$
- $\mathbf{R}\epsilon = \epsilon\mathbf{R} = \mathbf{R}$

RE IDENTITIES

- $\mathbf{R} \cup \phi = \mathbf{R}$
- $\mathbf{R}\epsilon = \epsilon\mathbf{R} = \mathbf{R}$
- $\phi^* = \epsilon$

RE Identities

- $\mathbf{R} \cup \phi = \mathbf{R}$
- $\mathbf{R}\epsilon = \epsilon\mathbf{R} = \mathbf{R}$
- $\phi^* = \epsilon$
- Note that we do not have explicit operators for intersection or complementation!

Converting REs to NFAs: Basis Cases

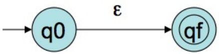
Regular Expression **Corresponding NFA**

ϕ

Regular Expression **Corresponding NFA**

ϕ

ϵ



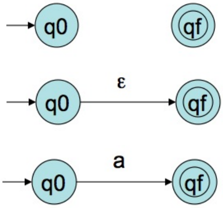
Converting REs to NFAs: Basis Cases

Regular Expression **Corresponding NFA**

ϕ

ϵ

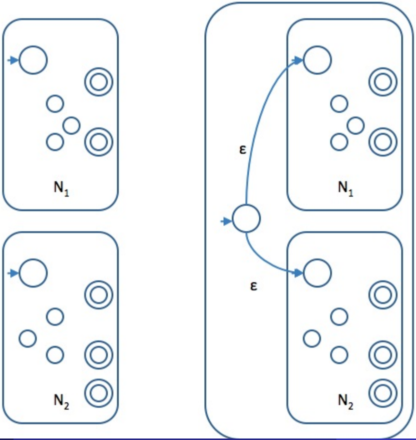
a for $a \in \Sigma$



Converting REs to NFAs

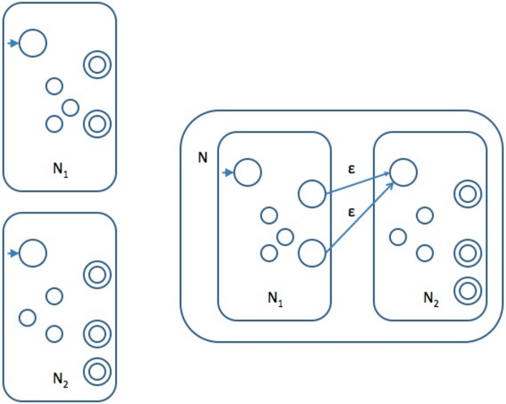
Union

- Let N_1 and N_2 be NFAs for R_1 and R_2 respectively. Then the NFA for $\mathbf{R_1 \cup R_2}$ is



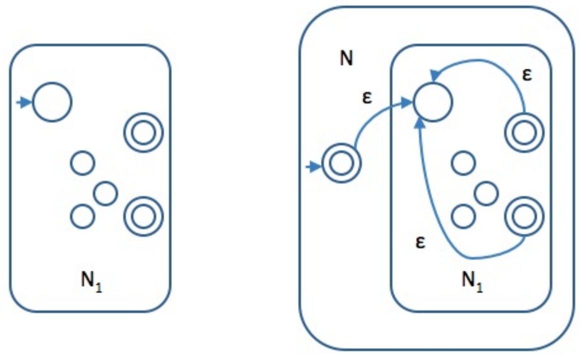
Concatenation

- Let N_1 and N_2 be NFAs for R_1 and R_2 respectively. Then the NFA for $\mathbf{R_1 R_2}$ is



Star

- Let N be NFAs for R. Then the NFA for $\mathbf{R}^*$ is

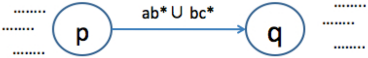


RE to NFA Conversion Example

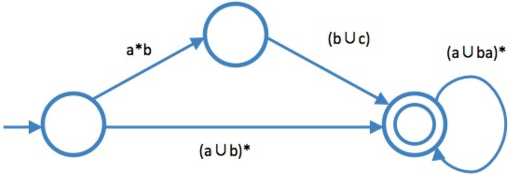
- Let's convert $(\mathbf{a} \cup \mathbf{b})^*\mathbf{aba}$ to an NFA.

Generalized Transitions and NFA

- A generalized transition is a transition whose label is a regular expression



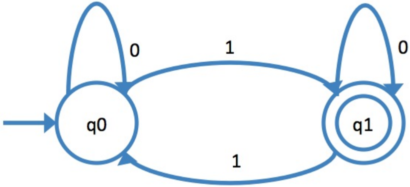
- A Generalized NFA is an NFA with generalized transitions.



  - In fact, all standard DFA transitions are generalized transitions with regular expressions of a single symbol!

GENERALIZED TRANSITIONS

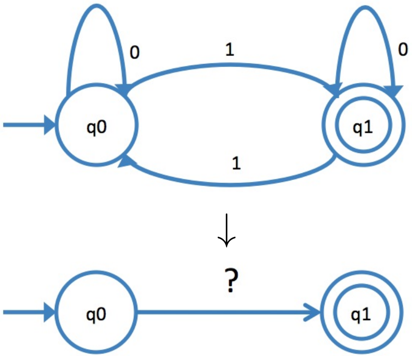
- Consider the 2-state DFA



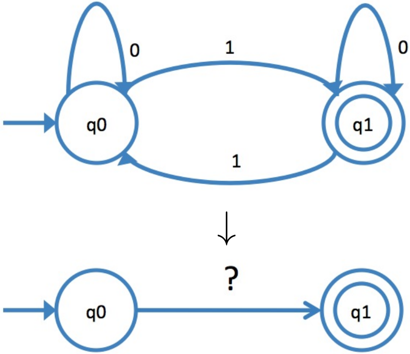
$\downarrow$

Generalized Transitions

- Consider the 2-state DFA



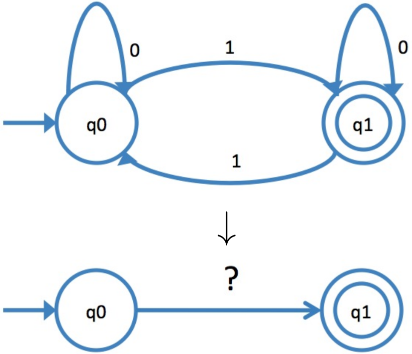
- Consider the 2-state DFA



- **0*****1** takes the DFA from state q_0 to q_1

GENERALIZED TRANSITIONS

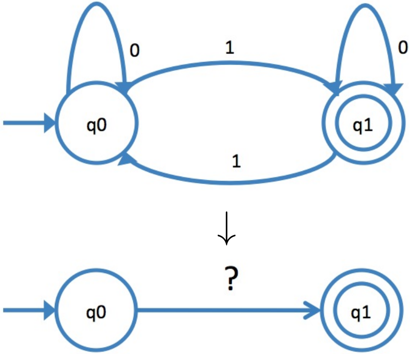
- Consider the 2-state DFA



- $\mathbf{0}^*\mathbf{1}$ takes the DFA from state q_0 to q_1
- $(\mathbf{0} \cup \mathbf{1}\mathbf{0}^*\mathbf{1})^*$ takes the machine from q_1 back to q_1

Generalized Transitions

- Consider the 2-state DFA



- $\mathbf{0}^*\mathbf{1}$ takes the DFA from state q_0 to q_1
- $(\mathbf{0} \cup \mathbf{10}^*\mathbf{1})^*$ takes the machine from q_1 back to q_1
- So $\mathbf{?} = \mathbf{0}^*\mathbf{1}(\mathbf{0} \cup \mathbf{10}^*\mathbf{1})^*$ represents all strings that take the DFA from state q_0 to q_1