Regular Expressions

DFA $\quad \delta : Q \times \Sigma \rightarrow Q$

NFA $\quad \delta : Q \times \Sigma \rightarrow 2^{Q} \quad P(Q)$

$\Sigma = \{0, 1\}$

$L(M_1) = \{ w \mid w \text{ ends in } 1 \} = \{ 1, 01, 11, 001 \}$

DFA

$L(M_2) = \{ w \mid w \text{ start with } 0 \}$

$= \{ 0, 00, 010, 001 \}$

NFA

$(2+3) * 5$

The language accepted by DFA or NFA can be expressed by a regular expression.

Def^n Let Σ be an alphabet

$L, L_1, L_2 \subseteq \Sigma^*$

$L_1 L_2 = \{ xy \mid x \in L_1, y \in L_2 \}$ $\Sigma = \{a, b\}$

$L_1 = \{a, ab\}$, $L_2 = \{b, acb\}$

$L_1 L_2 = \{ab, aacb, abb, abacb\}$

$L_1 \cup L_2 = \{ x \mid x \in L_1 \text{ or } x \in L_2 \}$

Kleene closure

Define $L^0 = \{\epsilon\}$

$L^i = L L^{i-1}$ for $i \geq 1$.

$L^1 = L L^0$

$L^2 = L L^1$

$L^3 = L L^2$

$L^* = \bigcup_{i=0}^{\infty} L^i$

$= L^0 \cup L^1 \cup L^2 \cup L^3 \cup \cdots$

$\in$ Kleene closure of L

$L = \{01, 10\}$

$L^2 = \{01, 10\} * \{01, 10\}$

$L^3 = L^2 * L$

$L^+ = \bigcup_{i \geq 1}^{\infty} L^i$

Positive closure

Regular Expression

Defn Let Σ be an alphabet

$$(0+1)^* = \bigcup_{i=0}^{\infty} (0+1)^i$$

$+ = \cup$

$L = 0+1$

$$L^* = L^0 \cup L^1 \cup L^2 \cup L^3 \ldots$$

$$= \epsilon \cup$$

$\underline{Th^m}$ The class of regular languages is closed under union operation

$\underline{L_1}$ $\underline{L_2}$ regular $\underline{L_1 \cup L_2}$

NFA M_1 NFA M_2 NFA M

$\underline{Th^m}$ The class of regular languages is closed under concatenation operation

L_1 L_2 regular $L_1 L_2$ regular

Thm The class of regular language is closed under ~~on~~ ~~*~~ star operation

L_1 regular $\qquad$ L_1^* regular

M_1 $\qquad$ M_1^*

$2+3=5$

$\sim\{0\cup 1\}^*$ 01^*0 $\Sigma=\{0,1\}$

Regular expressions

$=\{0, 01, 011, 0111, \ldots\}$

Let Σ be an alphabet

The regular expressions over Σ are defined recursively as

1. For each $a\in\Sigma$, a is a regular expression and denotes the set $\{a\}$

$\{0\}$ 01^* $\{0\}$ $\{1\}$ $1^*=\epsilon\cup\{1\}\cup\{1\}^2\cup\{1\}^3$

$=\{\epsilon, 1, 11, 111, \ldots\}$

01

$0\cup 0^*=\{0\}\cup\{\epsilon, 0, 00, 000, \ldots\}$ $01^*=0, 01, 011, 0111, 01111, \ldots$

$0+0^*=\{\epsilon, 0, 00, 000, \ldots\}$

$\Sigma = \{0, 1, 2\}$ $\quad 0, 1, 2$ re ϵ $\{0\}$ $\{1\}$ $\{2\}$

Defn Let Σ be an alphabet. The regular expression over Σ is defined recursively as

$(0+1)^* - \{0,1\}^* = \epsilon, \cup \{0,1\} \cup \{0,1\}^2, \cup \{0,1\}^3 \cup \dots$

1. For each $a \in \Sigma$, a is a regular expression and denotes the set $\{a\}$

   $0 \quad \{0\}$ $\quad 1 \quad \{1\}$ $\quad \underline{(0+1)}^* \; \{0,1\} = \{\{0\} \cup \{1\}\}$

2. ϵ is a regular expression and denotes the set $\{\epsilon\}$
3. ϕ is a regular expression and denotes the empty set
4. If $\underline{r}$ and $\underline{s}$ are regular expressions denoting the set R and $\underline{S}$ respectively, then $\underline{(r+s)}$, $(\underline{rs})$, (r^*) are regular expressions denoting the languages $\underline{R \cup S}$, RS, R^* respectively.

$\underline{\text{Th}^m}$ A language is regular iff some regular expression specifies it

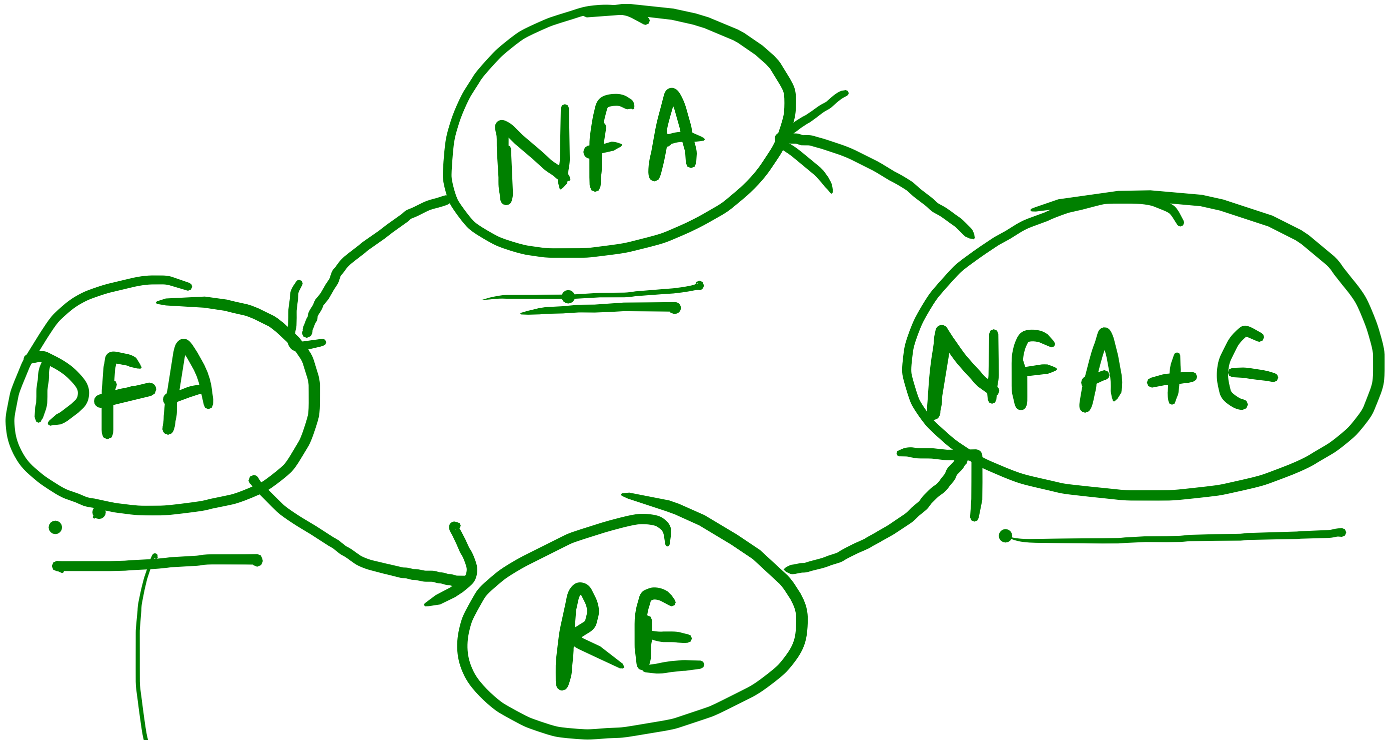
M_1 DFA $\qquad \underline{L(M_1)}$

$\|\!|$ $\qquad\qquad\qquad\qquad \|$

M_2 NFA $\qquad \underline{L(M_2)}$

$\qquad\qquad\qquad\qquad\quad \|$

$\underline{r}$ reg. $=$ $L(r)$



$$(0+1)^* = \{\epsilon, 0, 1, 00, \ldots$$

$$= \epsilon \cup \{0,1\} \cup \{0,1\}^2 \cup \{0,1\}^3 \cup \ldots$$

$$= \{\epsilon, 0, 1, 00, 01, 10, 11, 000, \ldots\}$$

€

$L = \{ 0^n 1^n \mid n \geq 0 \}$ is not regular

CFL

RL

$0^n 1^n$

CFL CFG Pushdown automata

DFA M regular regular expres

Context free Grammars

$L = \{0^n 1^n \mid n \geq 0\}$

CFG is a set of specification rules or production rules

$$\underbrace{lhs}_{\text{variables}} \rightarrow \underbrace{rhs}_{\substack{\text{string of}\\ \text{variables + constants}}}$$

where one variable is designated as start symbol

Language generated by the grammar G

CFG $\quad G_1$

Context free Language

CFL

$L(G_1) = \{ 0\#1,\ 00\#11,\ 000\#111, \ldots \}$

$= \{ 0^n \# 1^n \mid n \geq 1 \}$

G_1

CFG

$S \to 0A1$

$A \to 0A1$

$A \to B$

$B \to \#$

Variables

S, A, B

Constants 0, 1, #

S start symbol

Variables = nonterminals

Constants = terminals

$0^n \# 1^n$

$S \Rightarrow 0A1 \Rightarrow 0B1 \Rightarrow 0\#1$

$S \Rightarrow 0A1 \Rightarrow 00A11 \Rightarrow 000A111 \Rightarrow 000B111 \Rightarrow 000\#111$

$0\,0A11 \underset{G}{\Rightarrow} 000A111$

$S \overset{*}{\underset{G}{\Rightarrow}} 000\#111$

Formal Defn

A context free grammar is a 4-tuple

$G = \{V, T, P, S\}$

1. V is a finite set of variables

2.

$V = \{A, B\}$

$T = \{0, 1\}$

$(V \cup T)^* = \epsilon \cup \{A, B, 0, 1\} \cup$

Formal Def^n

A context free grammar is a 4-tuple

$G = \{V, T, P, S\}$

1. V is a finite set of variables $\{A, B, 0, 1\}^2, \ldots$
2. T is a finite set of terminals
3. $V \cap T = \phi$
4. P is a finite set of production rules

each production rule is of the form $A \to \alpha$

A variable, $\alpha \in \underline{(V \cup T)^*}$

$A \to \alpha_1$, $A \to \alpha_2$, . . . $A \to \alpha_n$

$$\underbrace{A \to \alpha_1 \mid \alpha_2 \mid \cdots \mid \alpha_n}$$

Notation

1. Capital letters A, B, C, D, E, S
denote variables
S start symbol

2. ~~a~~, lower case letters a, b, c, d, e, digits
are terminals

3. The capital letters X, Y, Z denote terminals or variables

4. lower case letters u, v, w, x, y z, denote string of terminals i.e. $\in \underline{T^*}$

5. Greek letters α, β, γ strings of variables and terminals $\in \underline{(V \cup T)^*}$

Direct Derivation $\underline{G}$

Df $\alpha, \beta, \gamma \in (V \cup T)^*$

$$A \to \beta \in P$$

then

$$\underbrace{\alpha A \gamma}_{\alpha_1} \underset{G}{\Rightarrow} \underbrace{\alpha \beta \gamma}_{\alpha_2}$$

$\alpha A \gamma$ directly derives $\alpha \beta \gamma$

Derivation

suppose $\alpha_1, \alpha_2, \ldots \alpha_n \in (V \cup T)^*$

$$\underbrace{\alpha_1 \underset{G}{\Rightarrow} \alpha_2}, \quad \underbrace{\alpha_2 \underset{G}{\Rightarrow} \alpha_3}, \ldots \quad \underbrace{\alpha_{m-1} \underset{G}{\Rightarrow} \alpha_m}$$

we say $\underbrace{\alpha_1 \overset{*}{\underset{G}{\Rightarrow}} \alpha_m}$

α_1 derives α_m

$\underset{G}{\Rightarrow}$ directly derive

$\overset{*}{\underset{G}{\Rightarrow}}$

If $G = (V, T, P, S)$ is a CFG.

$$\underbrace{L(G)} = \{ \underline{w} \mid w \in T^{*} \text{ and } S \underset{G}{\overset{*}{\Rightarrow}} w \}$$

$L(G)$ a context free Language

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

We call L a ~~cal~~ CFL if L is $L(G)$ for some grammar G.

$T = \Sigma$

$x = ab \qquad abba$

$\underbrace{u}_{\in T^{*}} = \underbrace{b}ba$

terminal

T set of terminal

Exp Construct a grammar for language

$\{0^n 1^n \mid n \geq 0\} \cup \{1^n 0^n \mid n \geq 0\}$

$G_1 \begin{cases} S_1 \to 0S_1 1 \\ S_1 \to \epsilon \end{cases}$

$S_1 \to 0S_1 1 \mid \epsilon$

$G_2 \ [S_2 \to 1S_2 0 \mid \epsilon$

$S_1 \Rightarrow 0S_1 1 \Rightarrow 01$

$S_1 \Rightarrow 0S_1 1 \Rightarrow 00S_1 11 \Rightarrow 0011$

$= 0^2 1^2$

$S_1 \Rightarrow 0^3 1^3$

$\underline{\underline{S}} \to S_1 \mid S_2$

$[S_1 \to 0S_1 1 \mid \epsilon$

$[S_2 \to 1S_2 0 \mid \epsilon$