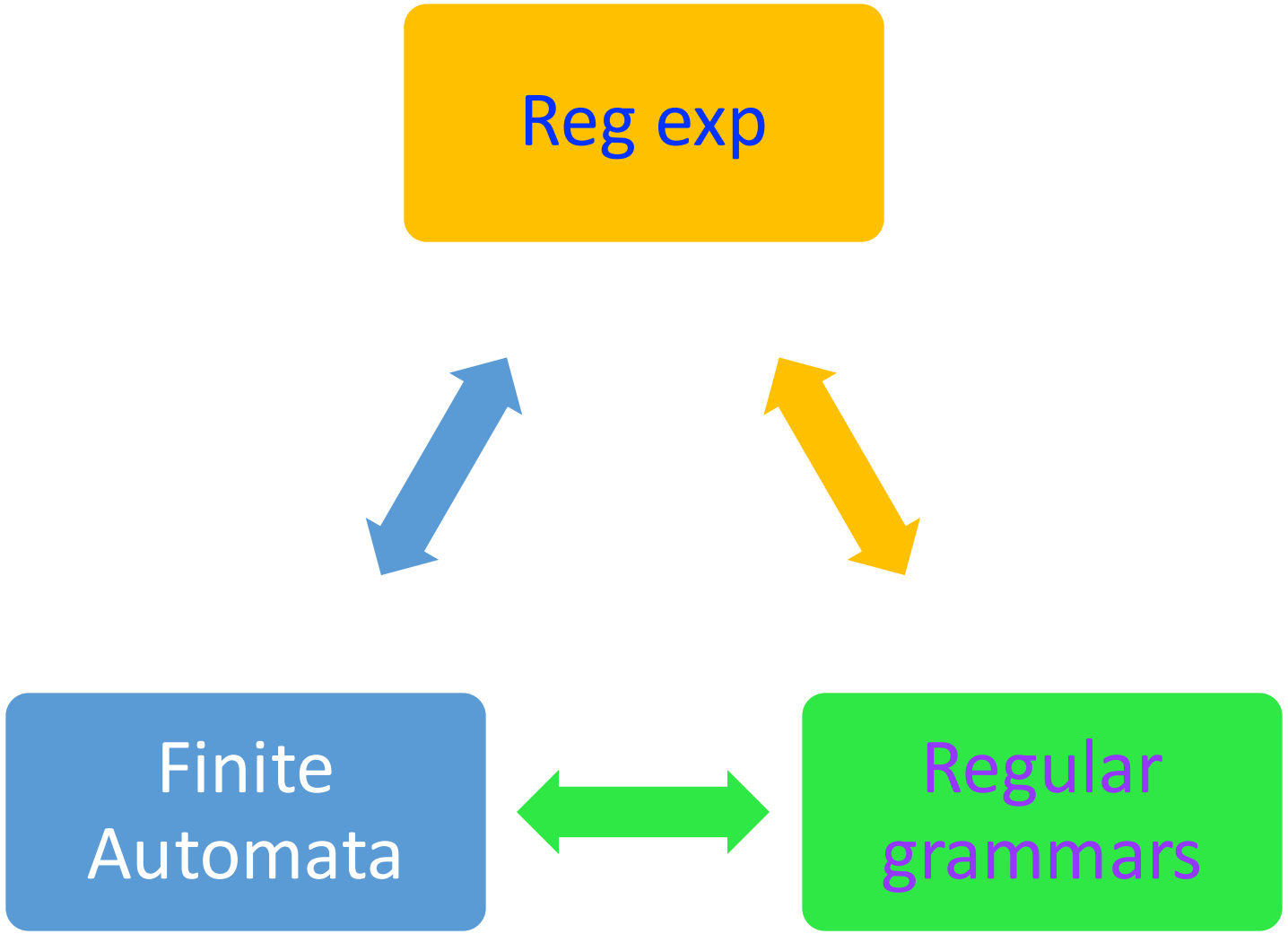
Course 4


Reg exp
Finite Automata
Regular grammars

Theorem 1: For any regular grammar G=(N, $\boldsymbol{\Sigma}$, P, S) there exists a FA M=(Q, $\boldsymbol{\Sigma}$, δ, q_0,F) such that L(G) = L(M)

Proof: construct M based on G

Q = N U {K}, K ∉ N

q_0 = S

F = {K} U {S| if S→$\boldsymbol{\varepsilon}$ ∈ P}

δ: if A →aB ∈ P then δ(A,a) = B

if A →a ∈ P then δ(A,a) = K

Theorem 2: For any FA M=(Q, $\boldsymbol{\Sigma}$, δ, q_0,F) there exists a right linear grammar G=(N, $\boldsymbol{\Sigma}$, P, S) such that L(G) = L(M)

Proof: construct G based on M

N = Q

S = q_0

P: if δ(q,a) = p then q →ap ∈ P

if p ∈ F then q →a ∈ P

if q_0 ∈ F then S → $\boldsymbol{\varepsilon}$

Regular sets

Definition: Let $\boldsymbol{\Sigma}$ be a finite alphabet. We define regular sets over $\boldsymbol{\Sigma}$ recursively in the following way:

1. $\boldsymbol{\phi}$ is a regular set over $\boldsymbol{\Sigma}$ (empty set)
2. $\{\boldsymbol{\varepsilon}\}$ is a regular set over $\boldsymbol{\Sigma}$
3. {a} is a regular set over $\boldsymbol{\Sigma}$, $\forall$ a$\in\boldsymbol{\Sigma}$
4. If P, Q are regular sets over $\boldsymbol{\Sigma}$, then P$\cup$Q, PQ, P* are regular sets over $\boldsymbol{\Sigma}$
5. Nothing else is a regular set over $\boldsymbol{\Sigma}$

Regular expressions

Definition: Let $\boldsymbol{\Sigma}$ be a finite alphabet. We define regular expressions over $\boldsymbol{\Sigma}$ recursively in the following way:

1. $\boldsymbol{\phi}$ is a regular expression denoting the regular set $\boldsymbol{\phi}$ (empty set)
2. $\boldsymbol{\varepsilon}$ is a regular expression denoting the regular set $\{\boldsymbol{\varepsilon}\}$
3. **a** is a regular expression denoting the regular set {a}, $\forall$ a$\in\boldsymbol{\Sigma}$
4. If **p,q** are regular expression denoting the regular sets P, Q then:
   - **p+q** is a regular expression denoting the regular set P$\cup$Q,
   - **pq** is a regular expression denoting the regular set PQ,
   - **p*** is a regular expression denoting the regular set P*
5. Nothing else is a regular expression

Remarks:

Examples

O1* denotes {0,01,011, 0111,...}
(01)* denotes {Ɛ, 01, 0101,...}
0*1* denotes {Ɛ,0,1,01,00,11,...}
01+10* denotes {01, 1,10,100,...}

1. $p^+ = pp^*$
2. Use paranthesis to avoid ambiguity
3. Priority of operations: *, concat, + (from high to low)
4. For each regular set we can find at least one regular exp to denote it (there is an infinity of reg exp denoting them)
5. For each regular exp, we can construct the corresponding regular set
6. 2 regular expressions are **equivalent** **iff** they denote the same regular set

Algebraic properties of regular exp

Let α, β, γ be regular expressions.

1. $\alpha + \beta = \beta + \alpha$
2. $\boldsymbol{\Phi}^* = \boldsymbol{\varepsilon}$
3. $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$
4. $\alpha(\beta\gamma) = (\alpha\beta)\gamma$
5. $\alpha\,(\beta + \gamma) = \alpha\beta + \alpha\gamma$
6. $(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma$
7. $\alpha\,\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}\,\alpha = \alpha$
8. $\boldsymbol{\Phi}\alpha = \alpha\boldsymbol{\Phi} = \boldsymbol{\Phi}$
9. $\alpha^* = \alpha + \alpha^*$
10. $(\alpha^*)^* = \alpha^*$
11. $\alpha + \alpha = \alpha$
12. $\alpha + \boldsymbol{\Phi} = \alpha$

Reg exp equations

- Normal form: **X = aX + b**

  where a,b – reg exp

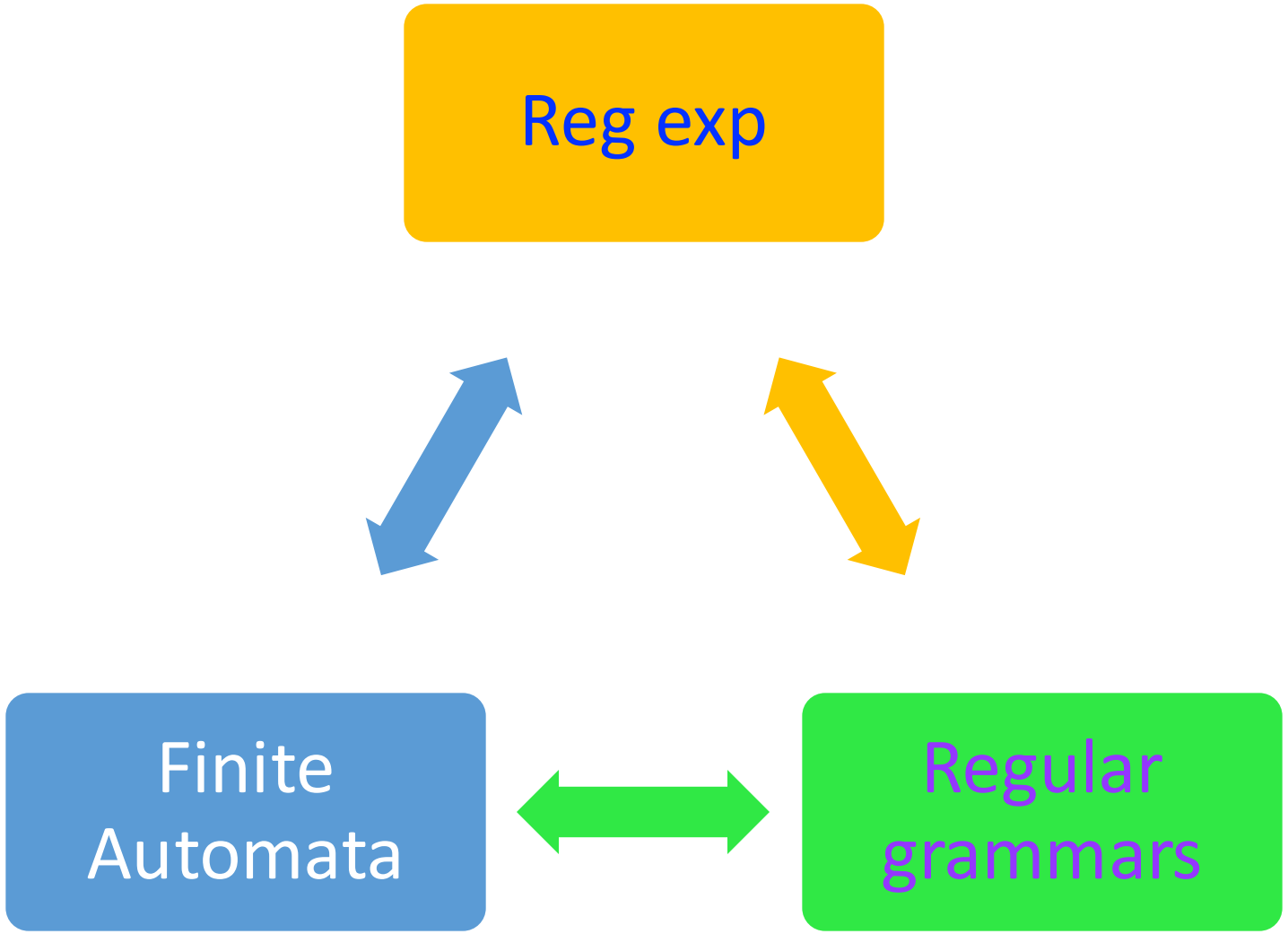
- Solution: **X = a*b**

a a*b + b = (aa* +**ε**)b = a*b

- System of reg exp equations:

$$\begin{cases} X = a_1 X + a_2 Y + a_3 \\ Y = b_1 X + b_2 Y + b_3 \end{cases}$$

- Solution: Gauss method (replace X_i and solve X_n)


Reg exp
Finite Automata
Regular grammars

Lemma 1:$\boldsymbol{\Phi}$,{$\boldsymbol{\varepsilon}$}, {a},$\forall a \in \boldsymbol{\Sigma}$ are accepted by FA

Reg exp	FA
$\boldsymbol{\Phi}$	$M = (Q, \boldsymbol{\Sigma}, \delta, q_0, \boldsymbol{\Phi})$
$\boldsymbol{\varepsilon}$	$M = (Q, \boldsymbol{\Sigma}, \boldsymbol{\Phi}, q_0, \{q_0\})$
$a, \forall a \in \boldsymbol{\Sigma}$	$M = (\{q_0,q_1\}, \boldsymbol{\Sigma}, \{\delta(q_0,a) = q_1\}, q_0, \{q_1\})$

Lemma 2: If L_1 and L_2 are accepted by a FA then: $L_1 \cup L_2$, L_1L_2 and L_1^* are accepted by FA

Proof:

$M_1 = (Q_1, \Sigma_1, \delta_1, q_{01}, F_1)$ such that $L_1 = L(M_1)$

$M_2 = (Q_2, \Sigma_2, \delta_2, q_{02}, F_2)$ such that $L_2 = L(M_2)$

$M_3 = (Q_3, \Sigma_{1U}, \delta_3, q_{03}, F_3)$

$Q_3 = Q_1 \cup Q_2 \cup \{q_{03}\}$; $\sum_3 = \sum_1 \cup \sum_2$

$F_3 = F_1 \cup F_2 \cup \{q_{03} \mid \text{if } q_{01} \in F_1 \text{ or } q_{02} \in F_2\}$

$\delta_3 = \delta_1 \cup \delta_2 \cup \{\delta_3(q_{03},a) = p \mid \exists \delta_1(q_{01},a) = p\} \cup$

$\{\delta_3(q_{03},a) = p \mid \exists \delta_2(q_{02},a) = p\}$

$L(M_3) = L(M_1) \cup L(M_2)$

PROOF!!! Homework

$M_4 = (Q_4, \Sigma_4, \delta_4, q_{04}, F_4)$

$Q_4 = Q_1 \cup Q_2$; $q_{04} = q_{01}$;

$F_4 = F_2 \cup \{q \in F_1 \mid \text{if } q_{02} \in F_2\}$

$$\delta_4(q,a) = \begin{cases} \delta_1(q,a), \text{ if } q \in Q_1 - F_1 \\ \delta_1(q,a) \cup \delta_2(q_{02},a) \text{ if } q \in F_1 \\ \delta_2(q,a), \text{ if } q \in Q_2 \end{cases}$$

$L(M_3) = L(M_1)L(M_2)$

PROOF!!! Homework

$M_5 = (Q_5, \Sigma_1, \delta_5, q_{05}, F_5)$

$Q_5 = Q_1$; $q_{05} = q_{01}$

$F_5 = F_1 \cup \{q_{01}\}$

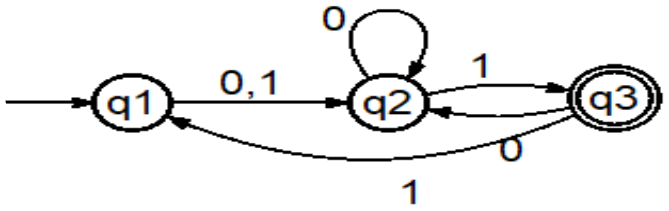
$\delta_5(q,a) = \delta_1(q,a)$, if $q \in Q_1 - F_1$

$\delta_1(q,a) \cup \delta_1(q_{01},a)$ if $q \in F_1$

$L(M_3) = L(M_1)^*$

PROOF!!! Homework

Theorem: *A language is a regular set if and only if is accepted by a FA*



Proof:

⇒Apply lemma 1 and lemma 2

<= construct a system of regular exp equations where:

- Indeterminants – states
- Coefficients – terminals
- Equation for A: all the possibilities that put the FA in state A
- Equation of the form: X=Xa+b => solution **X=ba***

$$\begin{cases} q_1 = q_3 0 + \boldsymbol{\varepsilon} \\ q_2 = q_1 0 + q_1 1 + q_2 0 + q_3 0 \\ q_3 = q_2 1 \end{cases}$$

Regular exp = union of solutions corresponding to final states

Prop:*Regular sets are right linear languages*

Lemma 1': $\boldsymbol{\Phi}$,{$\boldsymbol{\varepsilon}$}, {a},$\forall a \in \boldsymbol{\Sigma}$ are right linear languages

Proof: constructive

i. G = ({S}, $\boldsymbol{\Sigma}$, $\boldsymbol{\Phi}$, S) – regular grammar such that L(G) = $\boldsymbol{\Phi}$

ii. G = ({S}, $\boldsymbol{\Sigma}$,{S→$\boldsymbol{\varepsilon}$}, S) – regular grammar such that L(G) ={$\boldsymbol{\varepsilon}$}

iii. G = ({S}, $\boldsymbol{\Sigma}$,{S→a}, S) – regular grammar such that L(G) ={a}

Lemma 2': If L_1 and L_2 are right linear languages then:

$L_1 \cup L_2$, L_1L_2 and L_1^* are right linear languages.

Proof: constructive

L_1, L_2 right linear languages => $\exists G_1, G_2$ such that

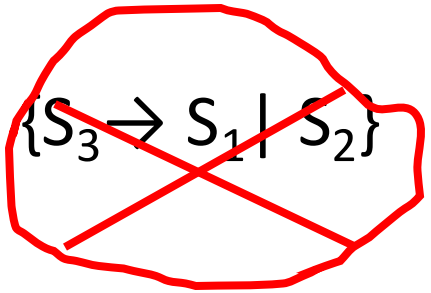
$G_1 = (N_1, \mathbf{\Sigma}_1, P_1, S_1)$ and $L_1 = L(G_1)$

$G_2 = (N_2, \mathbf{\Sigma}_2, P_2, S_2)$ and $L_2 = L(G_2)$ assume $N_1 \cap N_2 = \emptyset$

i. $G_3 = (N_3, \Sigma, P_3, S_3)$

$N_3 = N_1 \cup N_2 \cup \{S_3\}$; $\sum_3 = \sum_1 \cup \sum_2$

$P_3 = P_1 \cup P_2 \cup \{S_3 \rightarrow S_1 \mid S_2\}$



$$\{S_3 \rightarrow \alpha_1 \mid S_1 \rightarrow \alpha_1 \in P_1\} \cup \{S_3 \rightarrow \alpha_2 \mid S_2 \rightarrow \alpha_2 \in P_2\}$$

G_3 – right linear language

and

$L(G_3) = L(G_1) \cup L(G_2)$

PROOF!!! Homework

ii. $G_4 = (N_4, \mathbf{\Sigma}, P_4, S_4)$

$N_4 = N_1 U\ N_2$; $S_4 = S_1$; $\sum_4 = \sum_1 U \sum_2$

$P_4 =$ $\{A \rightarrow aB \mid \text{if } A \rightarrow aB \in P_1\}\ U$

$\{A \rightarrow aS_2 \mid \text{if } A \rightarrow a \in P_1\}\ U\ P_2\ U$

$\{ S_1 \rightarrow bA_2 \mid \text{if } S_2 \rightarrow bA_2 \in P_2 \text{ and } S_1 \rightarrow \boldsymbol{\varepsilon}\}$

G_4 – right linear language

and

$L(G_4) = L(G_1)\ L(G_2)$

PROOF!!! Homework

iii. $G_5 = (N_5, \Sigma_1, P_5, S_5)$

//IDEA: concatenate L_1 with itself

$N_5 = N_1 \cup \{S_5\}$;

$P_5 = P_1 \cup \{S_5 \rightarrow \varepsilon\} \cup$
$\{S_5 \rightarrow \alpha_1 \mid S_1 \rightarrow \alpha_1 \in P_1\} \cup$
$\{A \rightarrow aS_1 \mid \text{if } A \rightarrow a \in P_1\}$

G_5 – right linear language

and

$L(G_5) = L(G_1)^*$

PROOF!!! Homework

Theorem: *A language is a regular set if and only if is a right linear language*

Proof:

=> Apply lemma 1' and lemma 2'

<= construct a system of regular exp equations where:

- Indeterminants – nonterminals
- Coefficients – terminals
- Equation for A: all the possible rewritings of A

Example: G=({S,A,B},{0,1}, P, S)

P: S → 0A | 1B | **ε**

A → 0B | 1A

B → 0S | 1

$$\begin{cases} S = 0A + 1B + \varepsilon \\ A = 0B + 1A \\ B = 0S + 1 \end{cases}$$

Regular exp = solution corresponding to S