Course 3

Formal Languages

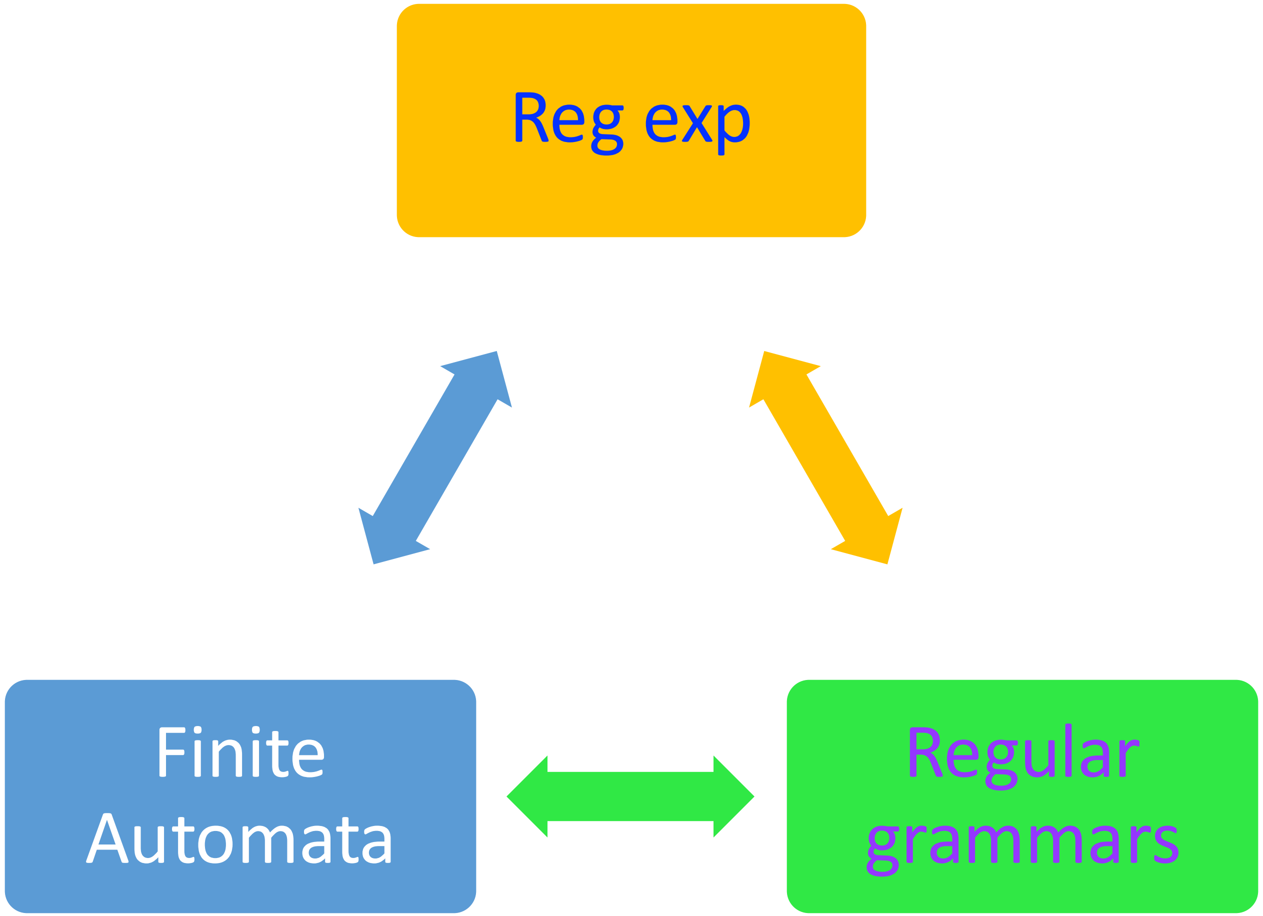
- Basic notions -

Changes in course schedule

- **27.10.2025 - Parallel and Distributed Programming**

- **31.10.2025 – Formal languages and Compiler Design**

Regular languages


Reg exp
Finite Automata
Regular grammars

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

Proof: construct M based on G

$Q = N \cup \{K\}, K \notin N$

$q_0 = S$

$F = \{K\} \cup \{S \mid \text{if } S \rightarrow \boldsymbol{\varepsilon} \in P\}$

δ: if $A \rightarrow aB \in P$ then $\delta(A,a) = B$

if $A \rightarrow a \in P$ then $\delta(A,a) = K$

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

Proof: construct M based on G

$Q = N \cup \{K\}, K \notin N$

$q_0 = S$

$F = \{K\} \cup \{S \mid \text{if } S \to \varepsilon \in P\}$

δ: if $A \to aB \in P$ then $\delta(A,a) = B$

if $A \to a \in P$ then $\delta(A,a) = K$

Prove that L(G) = L(M) ($w \in L(G) \Leftrightarrow w \in L(M)$):

$S \overset{*}{\Rightarrow} w \Leftrightarrow (S,w) \overset{*}{\vdash} (qf, \varepsilon)$

$w = \varepsilon$: $S \overset{*}{\Rightarrow} \varepsilon \Leftrightarrow (S, \varepsilon) \overset{*}{\vdash} (S, \varepsilon)$ – true

$w=a_1a_2...a_n$: $S \overset{*}{\Rightarrow} w \Leftrightarrow (S,w) \overset{*}{\vdash} (K, \varepsilon)$

$S \Rightarrow a_1A_1 \Rightarrow a_1a_2A_2 \Rightarrow ... \Rightarrow a_1a_2...a_{n-1}A_{n-1} \Rightarrow a_1a_2...a_{n-1}a_n$

$S \Rightarrow a_1A_1$ exists if $S \to a_1A_1$ and then $\delta(S,a_1)=A_1$

$A_1 \to a_2A_2$: $\delta(A_1,a_2)=A_2$ ...

$A_{n-1} \to a_n$: $\delta(A_{n-1},a_n)=K$

$(S,a_1a_2...a_n) \vdash (A_1,a_2...a_n) \vdash (A_2,a_3...a_n) \vdash ... \vdash (A_{n-1},a_n) \vdash (K, \varepsilon)$, $K \in F$

Theorem 2: For any FA $M=(Q, \boldsymbol{\Sigma}, \delta, q_0, F)$ 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 $\delta(q,a) = p$ then $q \rightarrow ap \in P$

if $p \in F$ then $q \rightarrow a \in P$

if $q_0 \in F$ then $S \rightarrow \boldsymbol{\varepsilon}$

Theorem 2: For any FA $M=(Q, \boldsymbol{\Sigma}, \delta, q_0, F)$ there exists a <u>right linear</u> 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 $\delta(q,a) = p$ then $q \rightarrow ap \in P$

if $p \in F$ then $q \rightarrow a \in P$

if $q_0 \in F$ then $S \rightarrow \boldsymbol{\varepsilon}$

Prove that L(M) = L(G) *($w \in L(M) \Leftrightarrow w \in L(G)$):*

$P(i)$: $q \overset{i+1}{\Rightarrow} x \Leftrightarrow (q,x) \overset{i}{\vdash} (q_f, \boldsymbol{\varepsilon})$, $q_f \in F$ -prove by induction

Apply P : $q_0 \overset{i+1}{\Rightarrow} w \Leftrightarrow (q_0,w) \overset{i}{\vdash} (q_f, \boldsymbol{\varepsilon})$, $q_f \in F$

If i=0: $q \Rightarrow x \Leftrightarrow (q,x) \overset{0}{\vdash} (q_f, \boldsymbol{\varepsilon})$ $(x= \boldsymbol{\varepsilon}, q=q_f)$ $q \Rightarrow \boldsymbol{\varepsilon} \Leftrightarrow q_0 \rightarrow \boldsymbol{\varepsilon}$, $q_0 \in F$

Assume $\forall$ $k \leq i$ P is true

$q \overset{i+1}{\Rightarrow} x \Leftrightarrow (q,x) \overset{i}{\vdash} (q_f, \boldsymbol{\varepsilon})$

For $q \in N$ apply "$\Rightarrow$" : $q \Rightarrow ap \overset{i}{\Rightarrow} ax$

If $q \Rightarrow ap$ then $\delta(q,a)= p$; if $p \overset{i}{\Rightarrow} ax$ then $(p,x) \overset{i-1}{\vdash} (q_f, \boldsymbol{\varepsilon})$, $qf \in F$

THEN $(q,ax) \overset{i}{\vdash} (q_f, \boldsymbol{\varepsilon})$, $qf \in F$

Regular sets

Definition: Let Σ be a finite alphabet. We define regular sets over Σ recursively in the following way:

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

Regular expressions

Definition: Let Σ be a finite alphabet. We define <u>regular expressions</u> over Σ recursively in the following way:

1. Φ is a regular expression denoting the regular set Φ (empty set)
2. ε is a regular expression denoting the regular set $\{\varepsilon\}$
3. **a** is a regular expression denoting the regular set $\{a\}$, $\forall\ a \in \Sigma$
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

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 $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ be regular expressions.

1. $\alpha + \beta = \beta + \alpha$
2. $\Phi^* = \varepsilon$
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. $\alpha\varepsilon = \varepsilon\alpha = \alpha$
8. $\Phi\alpha = \alpha\Phi = \Phi$
9. $\alpha^* = \alpha + \alpha^*$
10. $(\alpha^*)^* = \alpha^*$
11. $\alpha + \alpha = \alpha$
12. $\alpha + \Phi = \alpha$

Reg exp equations

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

  where a,b – reg exp

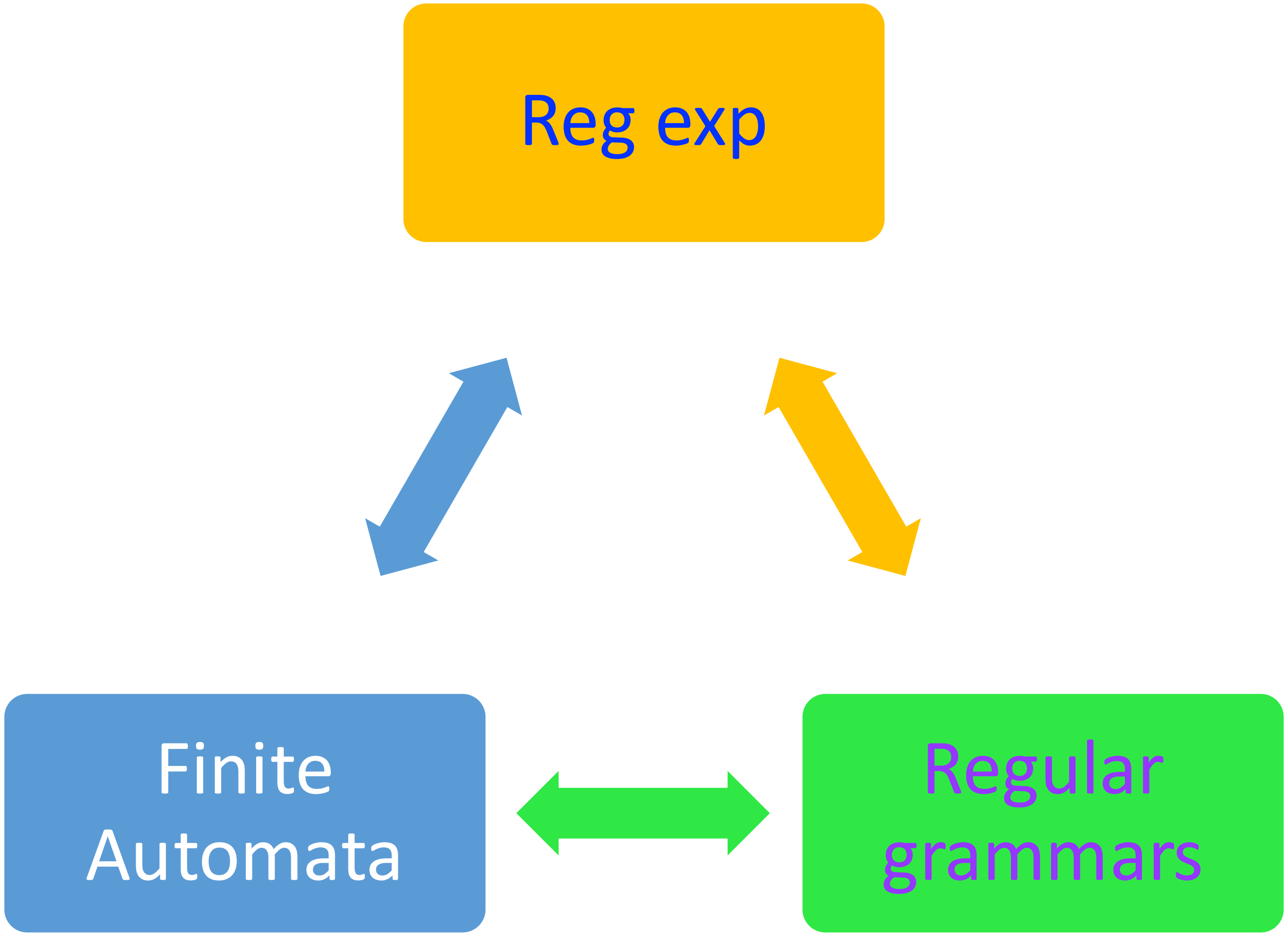
- Solution: **X = a*b**

$\boxed{a\,a^*b + b = (aa^* + \boldsymbol{\varepsilon})b = a^*b}$

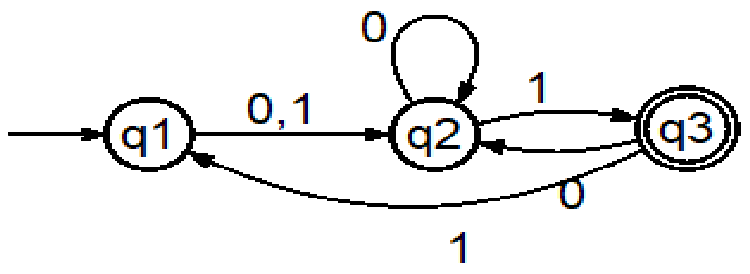
- System of reg exp equations:

$$\begin{cases} X = a_1X + a_2Y + a_3 \\ Y = b_1X + b_2Y + b_3 \end{cases}$$

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


Reg exp
Finite Automata
Regular grammars

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



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

Proof:

=> Apply lemma 1 and lemma 2 (to follow)

<= 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*

Regular exp = union of solutions corresponding to final states

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 $L_2 = L(M_2)$

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

$Q_3 = Q_1 \cup Q_2 \cup \{q_{03}\}$; $\Sigma_3 = \Sigma_1 \cup \Sigma_2$

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

$\delta_3 = \delta_1 \cup \delta_2 \cup \{\delta_3(q_{03},a) = p \mid \exists \delta_1(q_{01},a) = p\} \cup$
$\{\delta_3(q_{03},a) = p \mid \exists \delta_2(q_{02},a) = p\}$

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

PROOF!!! Homework

$M_4 = (Q_4, \boldsymbol{\Sigma}_4, \delta_4, q_{04}, F_4)$

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

$F_3 = F_2 \cup \{q \in F_1 \mid \text{if } q_{02} \in F_2\}$

$$\delta_3(q,a) = \begin{cases} \delta_1(q,a), \text{ if } q \in Q_1\text{-}F_1 \\ \delta_1(q,a) \cup \delta_2(q_{02},a) \text{ if } q \in F_1 \\ \delta_2(q,a), \text{ if } q \in Q_2 \end{cases}$$

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

PROOF!!! Homework

$M_5 = (Q_5, \boldsymbol{\Sigma}_1, \delta_5, q_{05}, F_5)$ *//IDEA: concatenate with itself*

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

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

$$\delta_5(q,a) = \begin{cases} \delta_1(q,a), \text{ if } q \in Q_1 - F_1 \\ \delta_1(q,a) \cup \delta_1(q_{01},a) \text{ if } q \in F_1 \end{cases}$$

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

PROOF!!! Homework