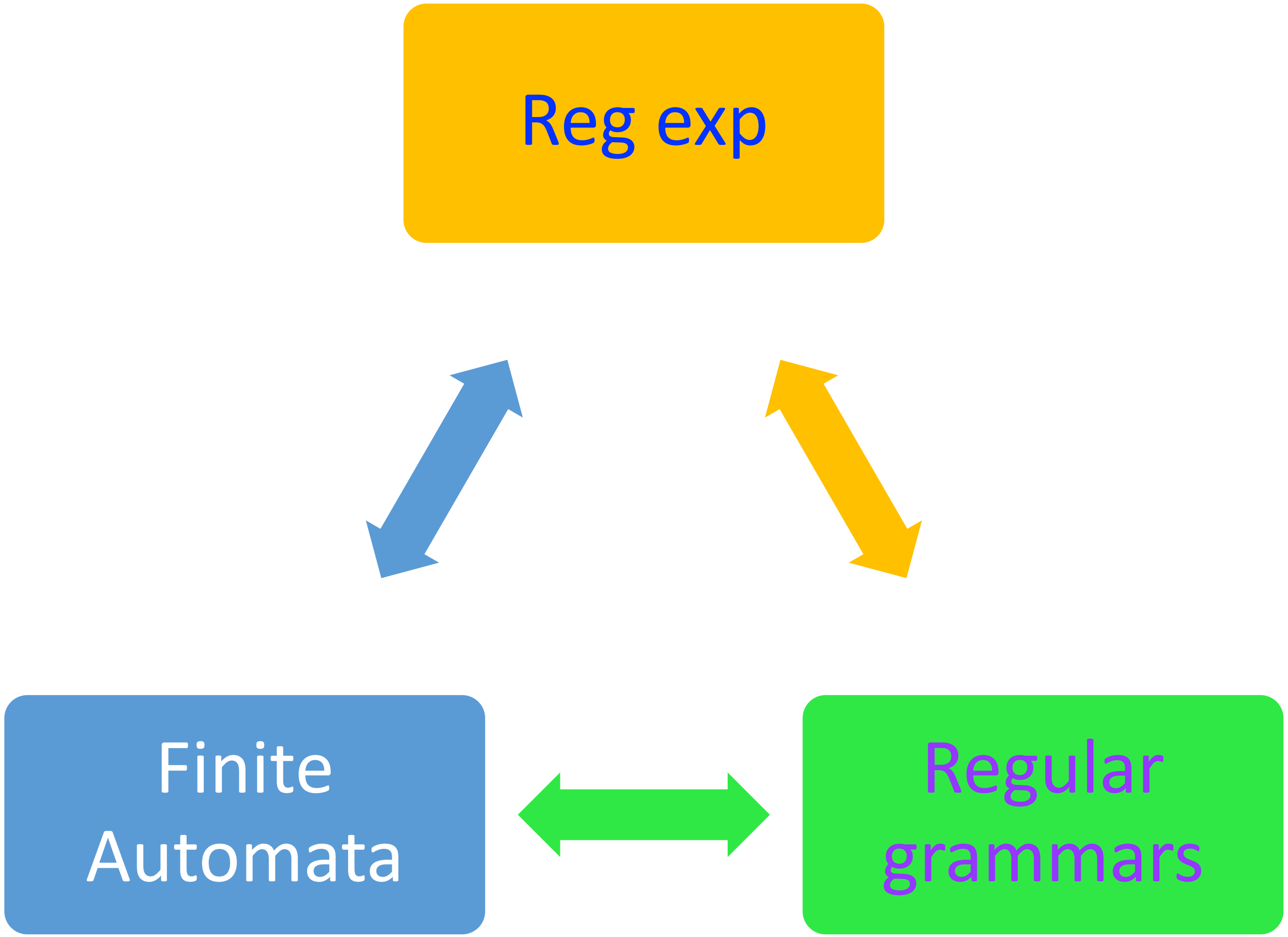
Course 4


Reg exp
Finite Automata
Regular grammars

Regular grammars

- $G = (N, \Sigma, P, S)$ right linear grammar if

$$\forall p \in P: A \to aB \text{ or } A \to b, \text{ where } A,B \in N \text{ and } a,b \in \Sigma$$

- $G = (N, \Sigma, P, S)$ regular grammar if
  - G is right linear grammar

  and
  - $A \to \varepsilon \notin P$, with the exception that $S \to \varepsilon \in P$, in which case S does not appear in the rhs (right hand side) of any other production

A->aA|a ok ✓
S->aA| ε and A->b ok ✓
S->aA| ε and A-> ε NOT ok ✗
S->aA| ε and A->bS|a NOT ok ✗

- $L(G) = \{w \in \Sigma^* \mid S \overset{*}{\Rightarrow} w\}$ - right linear language

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

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

$q_0 = S$

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

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

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

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

Proof: construct M based on G

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

$q_0 = S$

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

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∈L(G) ⇔ w∈L(M)):

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

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

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

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

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

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

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

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

EX 1

$G = (\{S,A\}, \{0,1\}, P, S)$

P: $S \rightarrow 0S \mid 0A$

$A \rightarrow 1A \mid 1$

M:

$Q = \{S,A,K\}$

$q_0 = S$

$F = \{K\}$

δ

	0	1
S	S, A	
A		A,K
K		

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 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 $S \rightarrow \varepsilon$

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

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

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

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

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

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

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 $(p,x) \overset{i-1}{\vdash} (q_f, \varepsilon)$, $qf \in F$

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

M:

$Q = \{S,A,K\}$

$q_0 = S$

$F = \{K\}$

δ

	0	1
S	S, A	
A		A,K
K		

N = {S,A,K}
S = S
P:
S-> 0S | 0A
A -> 1A| 1K | 1

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 $\mathbf{\Sigma}$ be a finite alphabet. We define regular expressions over $\mathbf{\Sigma}$ 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 $\{a\}$, $\forall\ a \in \mathbf{\Sigma}$
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 $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ be regular expressions.

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

Reg exp equations

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

  where a,b – reg exp

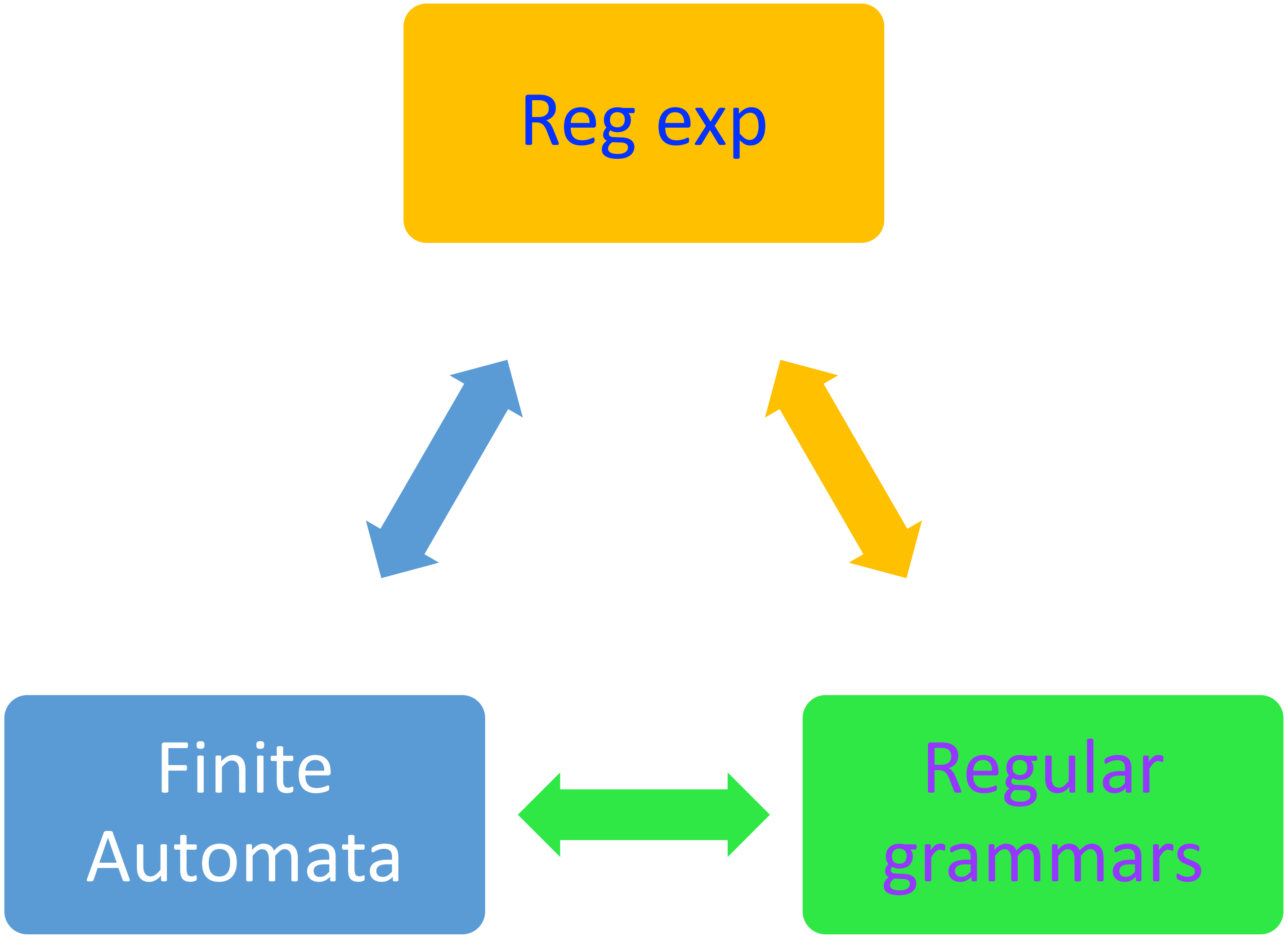
- Solution: **X = a*b**

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

- 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

Prop: *Regular sets are right linear languages*

Lemma 1: $\boldsymbol{\Phi}$,{$\boldsymbol{\varepsilon}$}, {a},∀a∈$\boldsymbol{\Sigma}$ 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

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

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

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

$N_3 = N_1 \cup N_2 \cup \{S_3\}$; $\Sigma_3 = \Sigma_1 \cup \Sigma_2$

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

$\{S_3 \to \boldsymbol{\alpha}_1 \mid S_1 \to \boldsymbol{\alpha}_1 \in P_1\} \cup \{S_3 \to \boldsymbol{\alpha}_2 \mid S_2 \to \boldsymbol{\alpha}_2 \in P_2\}$

G_3 – right linear language

and

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

PROOF!!! Homework

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

$N_4 = N_1 \cup N_2$; $S_4 = S_1$; $\Sigma_4 = \Sigma_1 \cup \Sigma_2$

$P_4 = \{A \to aB \mid \text{if } A \to aB \in P_1\} \cup$
$\{A \to aS_2 \mid \text{if } A \to a \in P_1\} \cup P_2 \cup$
$\{S_1 \to bA_2 \mid \text{if } S_2 \to bA_2 \in P_2 \text{ and } S_1 \to \boldsymbol{\varepsilon}\}$

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

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

$P_5 = \quad P_1 \cup \{S_5 \to \boldsymbol{\varepsilon}\} \cup$

$\{S_5 \to \boldsymbol{\alpha}_1 \mid S_1 \to \boldsymbol{\alpha}_1 \in P_1\} \cup$

$\{A \to aS_1 \mid \text{if } A \to a \in P_1\}$

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**

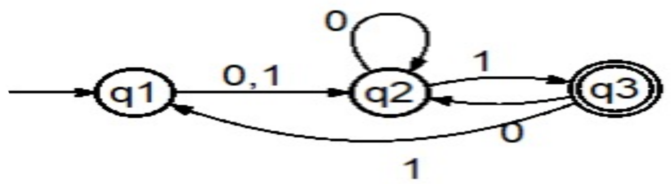
A → 0B | 1A

B → 0S | 1

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

Regular exp = solution corresponding to S

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



Proof:

=> Apply lemma 1’ and lemma 2’ (to follow, similar to RG)

<= 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*

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

Regular exp = union of solutions corresponding to final states