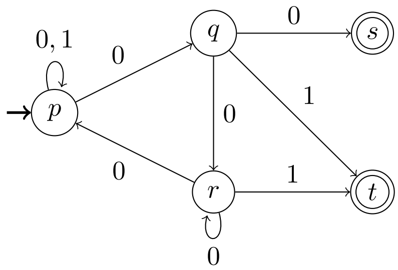
2.2.12 V minulé přednášce jsme si uvedli podmnožinovou konstrukci – konstrukci, která pro každý nedeterministický konečný automat zkonstruuje deterministický konečný automat, který přijímá stejný jazyk.

Podmnožinovou konstrukci si ukážeme na příkladě.

2.2.13 Příklad. Je dán NFA M tabulkou

		0	1
$\rightarrow$	p	$\{p,q\}$	$\{p\}$
	q	$\{r,s\}$	$\{t\}$
	r	$\{p,r\}$	$\{t\}$
$\leftarrow$	s	$\emptyset$	$\emptyset$
$\leftarrow$	t	$\emptyset$	$\emptyset$

se stavovým diagramem



Najděte DFA $\widehat{M}$, který přijímá stejný jazyk.

Řešení. DFA vytvoříme pomocí postupu 2.2.9. Na začátku práce algoritmu máme $I = \{p\}$, $Q = \{\{p\}\}$, $A = \{\{p\}\}$.

Množinu stavů Q i přechodovou funkci $\widehat{\delta}$ konstruujeme takto:

- Po prvním průchodu bodem 2 dostáváme:

$$\widehat{\delta}(\{p\},0) := \delta(p,0) = \{p,q\}, \quad \widehat{\delta}(\{p\},1) := \delta(p,1) = \{p\} \text{ a proto } B := \{\{p,q\}\}.$$

  V kroku 3

$$Q := \{\{p\},\{p,q\}\} \quad \text{a} \quad A := \{\{p,q\}\}.$$

- Po druhém průchodu bodem 2 dostáváme:

$$\widehat{\delta}(\{p,q\},0) := \delta(p,0) \cup \delta(q,0) = \{p,q,r,s\}$$

$$\widehat{\delta}(\{p,q\},1) := \delta(p,1) \cup \delta(q,1) = \{p,t\} \quad \text{a} \quad B := \{\{p,q,r,s\},\{p,t\}\}.$$

  V kroku 3

$$Q := \{\{p\},\{p,q\},\{p,q,r,s\},\{p,t\}\} \quad \text{a} \quad A := \{\{p,q,r,s\},\{p,t\}\}.$$

- Po třetím průchodu bodem 2 dostáváme:

$$\widehat{\delta}(\{p,q,r,s\},0) = \{p,q,r,s\} \quad \text{a} \quad \widehat{\delta}(\{p,q,r,s\},1) = \{p,t\},$$

$$\widehat{\delta}(\{p,t\},0) = \{p,q\}, \ \widehat{\delta}(\{p,t\},1) = \{p\} \text{ a } B := \emptyset.$$

  Protože $B = \emptyset$, přejdeme na krok 4 a definujeme:

$$\widehat{Q} := \{\{p\},\{p,q\},\{p,q,r,s\},\{p,t\}\}, \ \widehat{q_0} := \{p\}, \ \widehat{F} := \{\{p,q,r,s\},\{p,t\}\}.$$

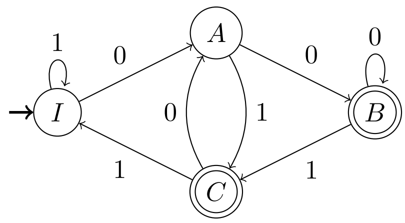
Předchozí postup je znázorněn v následující tabulce:

		0	1
$\rightarrow$	$\{p\}$	$\{p,q\}$	$\{p\}$
	$\{p,q\}$	$\{p,q,r,s\}$	$\{p,t\}$
$\leftarrow$	$\{p,q,r,s\}$	$\{p,q,r,s\}$	$\{p,t\}$
$\leftarrow$	$\{p,t\}$	$\{p,q\}$	$\{p\}$

Označíme-li $I = \{p\}$, $A = \{p,q\}$, $B = \{p,q,r,s\}$ a $C = \{p,t\}$, dostaneme následující tabulku.

		0	1
$\rightarrow$	I	A	I
	A	B	C
$\leftarrow$	B	B	C
$\leftarrow$	C	A	I

Stavový diagram je



Ze stavového diagramu není těžké nahlédnout, že jazyk přijímaný automatem M (i $\widehat{M}$) se skládá ze všech binárních slov, které končí buď 01 nebo 00, tj. ze všech binárních slov, které mají jako předposlední symbol 0.

2.3 Nedeterministické automaty s ε přechody

2.3.1 Nedeterministický automat s ε přechody (ve zkratce ε-NFA) se liší od obyčejného NFA tím, že mezi některými stavy může automat přejít, aniž by četl nějaký vstupní symbol.

Definice. ε-NFA je pětice $M = (Q, \Sigma, \delta, I, F)$, kde Q, Σ, I a F mají stejný význam jako u nedeterministického automatu, viz 2.2.1, a přechodová funkce δ je zobrazení

$$\delta\colon Q \times (\Sigma \cup \{\varepsilon\}) \to \mathcal{P}(Q).$$

Stavový diagram nějakého ε-NFA definujeme takto: Vrcholy jsou stavy automatu, z vrcholu q do vrcholu p vede hrana označená ε právě tehdy, když $p \in \delta(q,\varepsilon)$, a hrana označená $a \in \Sigma$ právě tehdy, když $p \in \delta(q,a)$. □

2.3.2 Jazyk přijímaný ε-NFA (neformálně). Zde je opět výhodné dívat se na práci ε-NFA nad slovem w jako na množinu sledů, které začínají v některém počátečním stavu a jsou označené slovem w (kdekoli ve sledu může být hrana označená prázdným slovem ε). Slovo w je přijímáno ε-NFA, jestliže některý z těchto sledů končí v některém koncovém stavu.

K formální definici využijeme pojem rozšířené přechodové fuknce.

Rozšířená přechodová funkce (neformálně). Pro každé slovo u a stav q je $\delta^\star(q,u)$ definováno jako množina těch stavů p, do kterých ve stavovém diagramu vede z q orientovaný sled označený u. K formálně přesnější definici potřebujeme pojem ε-uzávěru.

2.3.3 ε-uzávěr. Neformálně, množina ε-UZ(X) je množina všech stavů, do kterých vede sled z některého stavu $q \in X$, jehož všechny hrany jsou označeny ε.

Definice. Je dán NFA s ε přechody $M = (Q, \Sigma, \delta, I, F)$. Pro množinu stavů X definujeme *ε-uzávěr* ε-UZ(X) indukcí takto:

- $X \subseteq \varepsilon\text{-UZ}(X)$;
- je-li $p \in \varepsilon\text{-UZ}(X)$, pak $\delta(p, \varepsilon) \subseteq \varepsilon\text{-UZ}(X)$.

□

Má-li množina X jediný stav q, píšeme ε-UZ(q) místo ε-UZ$(\{q\})$.

2.3.4 Rozšířená přechodová funkce. Definice. *Rozšířená přechodová funkce* $\delta^\star$ je definována indukcí takto:

- $\delta^\star(q, \varepsilon) = \varepsilon\text{-UZ}(q)$;
- $\delta^\star(q, ua) = \bigcup\{\varepsilon\text{-UZ}(\delta(p, a)) \mid p \in \delta^\star(q, u)\}$, pro $a \in \Sigma$, $u \in \Sigma^\star$.

□

2.3.5 Jazyk přijímaný ε-NFA.

Definice. Slovo u je *přijímáno ε-NFA*, jestliže existuje počáteční stav q a koncový stav p takový, že $p \in \delta^\star(q, u)$. *Jazyk* $L(M)$ *přijímaný ε-NFA* M je množina právě všech slov přijímaných automatem M. □

2.3.6 Uvědomte si, že každý DFA je speciálním případem NFA a každý NFA je speciálním případem ε-NFA. Proto každý regulární jazyk je přijímán některým ε-NFA. Následující věta říká, že nedeterministické automaty s ε přechody nepřijímají „nic navíc.“

Věta. Jazyk přijímaný libovolným ε-NFA je regulární. □

Nástin důkazu:Větu dokážeme modifikací podmnožinové konstrukce z 2.2.8. Rozdíl je v tom, že množina stavů DFA se v případě ε-NFA neskládá ze všech podmnožin stavů automatu, ale jen z těch podmnožin, které jsou ε uzavřené. Přesněji:

Máme ε-NFA $M = (Q, \Sigma, \delta, I, F)$, který přijímá jazyk L. Označme Q_1 množinu všech podmnožin X množiny Q, které jsou ε uzavřené, tj. pro které platí $\varepsilon\text{-UZ}(X) = X$. Popíšeme variantu podmnožinové konstrukce, jejímž výsledkem bude DFA přijímající jazyk L.

Označme Q_1 množinu všech ε uzavřených podmnožin množiny Q.

Definujeme:

- $\delta_1(X, a) = \bigcup\{\varepsilon\text{-UZ}(\delta(q, a)) \mid q \in X\}$.
- $I_1 = \varepsilon\text{-UZ}(I)$.
- $F_1 = \{X \in Q_1 \mid X \cap F \neq \emptyset\}$.

Pak $M_1 = (Q_1, \Sigma, \delta_1, I_1, F_1)$ je deterministický automat, který přijímá jazyk L.

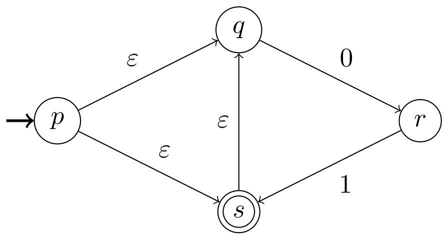
Poznámka. Opět se stačí zabývat jen dosažitelnými ε-uzavřenými množinami; tj. řídit se postupem analogickým 2.2.9. Uvědomte si, že sjednocením ε-uzavřených množin je opět ε-uzavřená množina.

Postup si ukážeme na následujícím příkladu.

2.3.7 Příklad. Je dán ε-NFA tabulkou

	ε	0	1
$\rightarrow$ p	$\{q,s\}$	$\emptyset$	$\emptyset$
q	$\emptyset$	$\{r\}$	$\emptyset$
r	$\emptyset$	$\emptyset$	$\{s\}$
$\leftarrow$ s	$\{q\}$	$\emptyset$	$\emptyset$

se stavovým diagramem



Najděte DFA $\widehat{M}$, který přijímá stejný jazyk a zredukujte ho.

Řešení. DFA vytvoříme pomocí postupu, který je analogický postupu 2.2.9. Máme:

$$\varepsilon\text{-UZ}(\{p\}) = \{p,q,s\},\ \varepsilon\text{-UZ}(\{q\}) = \{q\},\ \varepsilon\text{-UZ}(\{r\}) = \{r\},\ \varepsilon\text{-UZ}(\{s\}) = \{q,s\}.$$

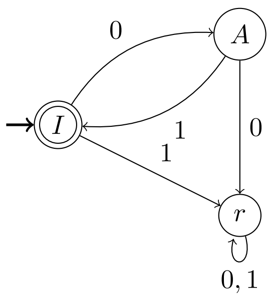
Proto počáteční stav $\widehat{M}$ je ε-UZ$(\{p\}) = \{p,q,s\}$. Další postup si znázorníme rovnou do tabulky.

	0	1
$\leftrightarrow$ $\{p,q,s\}$	$\{r\}$	$\emptyset$
$\{r\}$	$\emptyset$	$\{q,s\}$
$\leftarrow$ $\{q,s\}$	$\{r\}$	$\emptyset$
$\emptyset$	$\emptyset$	$\emptyset$

Po redukci a při označení $I = \{p,q,s\}$, $A = \{r\}$ a $O = \emptyset$, dostaneme výsledný DFA:

	0	1
$\leftrightarrow$ I	A	O
A	O	I
O	O	O

se stavovým diagramem



Není těžké nahlédnout, že $L(\widehat{M}) = \{(01)^n \mid n \geq 0\}$.

2.4 Uzávěrové vlastnosti regulárních jazyků

2.4.1 Operace s jazyky. V dalším textu si nejprve připomeneme množinové operace s jazyky a pak zavedeme ještě další operace specifické pro jazyky.

Jsou-li L_1 a L_2 dva jazyky nad stejnou abecedou, pak můžeme vytvořit jejich sjednocení $L_1 \cup L_2$ a průnik $L_1 \cap L_2$. Dále pro jazyk L je jeho doplněk jazyk $\overline{L} = \Sigma^\star \setminus L$ (tj. $w \in \overline{L}$ právě tehdy, když $w \notin L$). □

2.4.2 Zřetězení jazyků. Jsou dány jazyky L_1 a L_2 nad abecedou Σ. *Zřetězení* jazyků L_1 a L_2 je jazyk L_1L_2 definovaný

$$L_1L_2 = \{uv \mid u \in L_1, v \in L_2\}.$$

□

2.4.3 Kleeneho operace $\star$. Je dán jazyk L nad abecedou Σ. Definujeme $L^0 = \{\varepsilon\}$, $L^{i+1} = L^iL$ pro $i \geq 0$. Jazyk $L^\star$ je definován

$$L^\star = \{\varepsilon\} \cup L \cup L^2 \cup L^3 \cup \ldots = \bigcup_{i=0}^{\infty} L^i.$$

□

Poznamenejme, že operaci $\star$ se též říká *Kleeneho operátor*.

2.4.4 Reverze, též obrácení. Je dán jazyk L nad abecedou Σ. Definujeme reverzi (obrácení) jazyka L

$$L^R = \{w^R \mid w \in L\}.$$

□

Připomeňme, že např. pro slovo $u = a_1a_2 \ldots a_{n-1}a_n$ je $u^R = a_na_{n-1} \ldots a_2a_1$.

2.4.5 Poznámka. Uvědomte si, že pro jazyky L_1 a L_2 obecně **neplatí**:

1. $(L_1 \cup L_2)^\star = L_1^\star \cup L_2^\star$.
2. $(L_1 \cap L_2)^\star = L_1^\star \cap L_2^\star$.
3. $L_1^\star L_2^\star = (L_1L_2)^\star$.

2.4.6 Věta. Třída regulárních jazyků je uzavřena na následující operace: 1) sjednocení, 2) doplněk, 3) průnik, 4) zřetězení, 5) Kleeneho operaci $\star$ a 6) reverzi.

Přesněji, jestliže L, L_1 a L_2 jsou regulární jazyky, pak také $L_1 \cup L_2$, $\overline{L}$, $L_1 \cap L_2$, L_1L_2, $L^\star$ a L^R jsou regulární jazyky. □

Nástin důkazu: Pro L, L_1 a L_2 uvažujme DFA, které je přijímají; označme je po řadě $M = (Q, \Sigma, \delta, q_0, F)$, $M_1 = (Q_1, \Sigma, \delta_1, q_1, F_1)$ a $M_2 = (Q_2, \Sigma, \delta_2, q_2, F_2)$. Z jejich stavových diagramů vyrobíme stavové diagramy ε-NFA pro požadované jazyky. Předpokládáme, že množiny stavů Q_1 a Q_2 jsou disjunktní (uvědomte si, že bez tohoto předpokladu by naše konstrukce nebyly správné).

- Pro sjednocení jazyků stačí příslušné stavové diagramy položit vedle sebe. Přesněji NFA $\overline{M} = (Q_1 \cup Q_2, \Sigma, \overline{\delta}, \{q_1, q_2\}, F_1 \cup F_2)$, kde pro $q \in Q_1$ je $\overline{\delta}(q, a) = \delta_1(q, a)$ a pro $q \in Q_2$ je $\overline{\delta}(q, a) = \delta_2(q, a)$. Takto definovaný NFA přijímá jazyk $L_1 \cup L_2$.

- Pro doplněk $\overline{L}$ stačí v M přehodit koncové a nekoncové stavy, tj. koncové stavy DFA přijímajícího $\overline{L}$ je množina $Q \setminus F$. (Uvědomte si, že k tomu, aby stačilo přejmenovat koncové a nekoncové stavy, je nutné, aby automat byl deterministický; pro nedeterministický automat by automat, kde jsem pouze přehodili koncové a nekoncové stavy, ne vždy příjímal doplněk.)

- Protože průnik $L_1 \cap L_2$ je roven doplňku sjednocení doplňků, tj. $L_1 \cap L_2 = \overline{\overline{L_1} \cup \overline{L_2}}$, tvrzení proto vyplývá z předchozích dvou bodů.

- Pro zřetězení jazyků L_1 a L_2 položíme automaty M_1 a M_2 za sebe a z každého koncového stavu automatu M_1 vedeme ε přechod do počátečního stavu automatu M_2 (tj. pro $q \in F_1$ přidáme přechod $\overline{\delta}(q, \varepsilon) = \{q_2\}$.) Výsledný automat je pak $(Q_1 \cup Q_2, \Sigma, \overline{\delta}, \{q_1\}, F_2)$.

- Pro Kleeneho operaci $\star$, tj. pro jazyk $L^\star$, vezmeme automat M, k množině Q přidáme nový počáteční stav p_0, který bude současně novým koncovým stavem, a k přechodům automatu M přidáme tyto ε přechody: z nového počátečního stavu p_0 a z každého koncového stavu $q \in F$ přidáme ε přechod do původního počátečního stavu q_0 (všechny koncové stavy automatu M zůstanou koncové). Tím dostaneme ε-NFA, který přijímá $L^\star$.

- Pro reverzi jazyka L stačí v automatu M obrátit šipky a zaměnit počáteční a koncové stavy. (Uvědomte si, že z DFA se většinou stane NFA.)

2.4.7 Poznámka – součinová konstrukce. Fakt, že průnik a sjednocení dvou regulárních jazyků je opět regulární jazyk, se dá dokázat také pomocí součinové konstrukce. Důvody, proč zde součinovou konstrukci uvádíme, jsou dva. První je to, že pomocí ní se dá přímo z DFA pro jazyky L_1 a L_2 zkonstruovat DFA přijímající jazyk $L_1 \cap L_2$. Druhý je fakt, že se jedná o základní konstrukci, kterou využijeme i později při studiu zásobníkových automatů.

Součinová konstrukce Jsou dány DFA $M_1 = (Q_1, \Sigma, \delta_1, q_1, F_1)$ a $M_2 = (Q_2, \Sigma, \delta_2, q_2, F_2)$ takové, že $L_1 = L(M_1)$ a $L_2 = L(M_2)$. Definujme DFA $M = (Q, \Sigma, \delta, q_0, F)$ takto:

- $Q = Q_1 \times Q_2$,
- $q_0 = (q_1, q_2)$,
- $\delta((p, q), a) = (\delta_1(p, a), \delta_2(q, a))$,
- $F = F_1 \times F_2$.

Není těžké ukázat, že $L(M) = L_1 \cap L_2$. □

Výše uvedenou součinovou konstrukci jsme mohli využít i pro konstrukci DFA M', který přijímá jazyk $L(M') = L_1 \cup L_2$. Jediný rozdíl mezi M a M' je v množině koncových stavů – M' má množinu koncových stavů rovnu $F' = (F_1 \times Q_2) \cup (Q_1 \times F_2)$.