Automaty a gramatiky 2

Roman Barták, KTIML

bartak@ktiml.mff.cuni.cz
http://ktiml.mff.cuni.cz/~bartak

Co bylo minule

Formalizace a popis konečného automatu
- $A = (Q,X,\delta,q_0,F)$

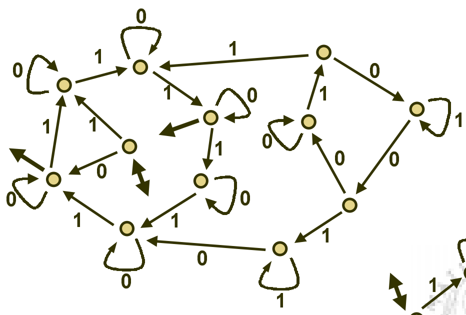
Jazyky a automaty
- jazyk přijímaný konečným automatem

Nerodova věta
- pravá kongruence
- charakteristika regulárních jazyků

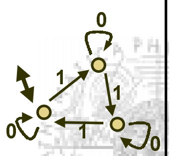
Ekvivalence automatů
- (automatový) homomorfismus

Trochu motivace

Co dělá tento automat?



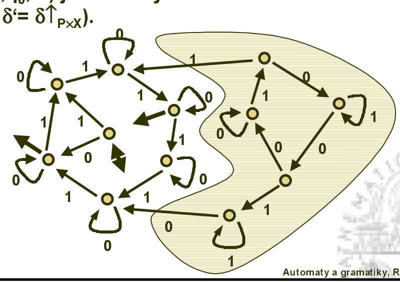
$L = \{w \mid w=\{0,1\}^* \ \&\ |w|_1=3k\}$



Dosažitelné stavy

$A = (Q,X,\delta,q_0,F)$ je konečný automat a $q \in Q$.
Řekneme, že stav q je *dosažitelný*, jestliže $\exists w \in X^*\ \delta^*(q_0,w)=q$

Věta: Nechť P jsou dosažitelné stavy automatu A. Potom $B=(P,X,\delta',q_0,F')$ je konečný automat ekvivalentní s A ($F'=F\cap P$, $\delta' = \delta\upharpoonright_{P\times X}$).



Hledání dosažitelných stavů

Iterační algoritmus (dosažitelnost po i krocích):

$M_0 = \{q_0\}$
repeat
$M_{i+1} = M_i \cup \{q \mid q\in Q, \exists p\in M_i\ \exists x\in X\ \delta(p,x)=q\}$
until $M_{i+1} = M_i$

Korektnost

$M_0 \subset M_1 \subset \ldots \subset Q$ (algoritmus je konečný)
každé M_i obsahuje pouze dosažitelné stavy

Úplnost

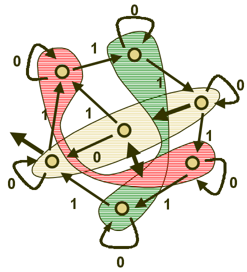
nechť q je libovolný dosažitelný stav, tj. $\exists w\in X^*\ \delta^*(q_0,w)=q$
vezměme nejkratší takové $w=x_1..x_n$ tž. $\delta^*(q_0,x_1..x_n)=q$
zřejmě $\delta^*(q_0,x_1..x_i) \in M_i$ (dokonce $M_i - M_{i-1}$)
tedy $\delta^*(q_0,x_1..x_n) \in M_n$
$q \in M_n$

Ekvivalentní stavy

Nechť $A = (Q,X,\delta,q_0,F)$ je konečný automat.

Stavy p,q jsou *ekvivalentní*, značíme $p\sim q$, jestliže:
$\forall w\in X^*\ \delta^*(p,w)\in F \Leftrightarrow \delta^*(q,w)\in F$

Výpočty startující z ekvivalentních stavů nelze rozlišit!



Ekvivalence po krocích

Ekvivalence po i krocích

$p \sim^i q \qquad \forall w \in X^*$ tž. $|w| \le i \;\; \delta^*(p,w) \in F \Leftrightarrow \delta^*(q,w) \in F$

$p \sim q \Leftrightarrow \forall i \;\; p \sim^i q$

Iterativní konstrukce $\sim^i$

$p \sim^0 q \quad \ldots \quad p \in F \Leftrightarrow q \in F$

$p \sim^{i+1} q \quad \ldots \quad p \sim^i q \;\;\&\;\; \forall x \in X \;\; \delta(p,x) \sim^i \delta(q,x)$

Je to v pořádku?

$p \sim^0 q \quad \ldots \quad \delta^*(p,\lambda) = p \in F \;\Leftrightarrow\; \delta^*(q,\lambda) = q \in F$

$p \sim^{i+1} q \quad \ldots \quad p \sim^i q$ tj. platí pro slova w tž. $|w| \le i$

slova délky i+1, $w = xu$, $|u| = i$

$\delta(p,x) \sim^i \delta(q,x)$ tj. $\delta^*(\delta(p,x),u) \in F \Leftrightarrow \delta^*(\delta(q,x),u) \in F$

dohromady $\delta^*(p,xu) \in F \Leftrightarrow \delta^*(q,xu) \in F$

Vlastnosti ekvivalence po krocích

1) $\forall i \ge 0$ je $\sim^i$ ekvivalence na Q, označme $\mathcal{R}_i = Q/\sim^i$
2) $\mathcal{R}_{i+1}$ zjemňuje $\mathcal{R}_i$
3) $\mathcal{R}_{i+1} = \mathcal{R}_i \Rightarrow \forall t > 0 \;\; \mathcal{R}_{i+t} = \mathcal{R}_i$
4) nechť $|Q| = n$, potom $\exists k \le n-1 \;\; \mathcal{R}_{k+1} = \mathcal{R}_k$
5) $\mathcal{R}_{k+1} = \mathcal{R}_k \Rightarrow (p \sim q \Leftrightarrow p \sim^k q)$

Důkaz:

1) a 2) přímo z definice

3) $p \sim^{i+1} q \Leftrightarrow p \sim^i q \;\;\&\;\; \forall x \in X \;\; \delta(p,x) \sim^i \delta(q,x)$

$p \sim^{i+2} q \Leftrightarrow p \sim^{i+1} q \;\;\&\;\; \forall x \in X \;\; \delta(p,x) \sim^{i+1} \delta(q,x)$

4) přímo z max. počtu tříd rozkladu (n)

5) $p \sim q \;\; \Leftrightarrow \forall i \ge 0 \;\; p \sim^i q$

$\Leftrightarrow \forall 0 \le i \le k \;\; p \sim^i q \;\&\; \forall i > k \;\; p \sim^i q$

$\Leftrightarrow p \sim^k q$

Hledání ekvivalence stavů

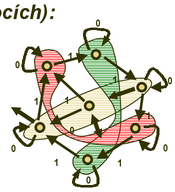
Iterační algoritmus (ekvivalence po i krocích):

sestroj $\mathcal{R}_0$

repeat

sestroj $\mathcal{R}_{i+1}$ z $\mathcal{R}_i$

until $\mathcal{R}_{i+1} = \mathcal{R}_i$



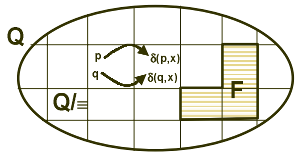
Příklad:

	0	1	$\mathcal{R}_0$	0	1	$\mathcal{R}_1$	0	1	$\mathcal{R}_2$
a	b	c	A	A	C	A	A	C	A
b	b	c	A	A	C	A	A	C	A
c	c	d	C	C	C	C	C	D	C
d	d	e	C	C	A	D	D	A	D
e	e	f	A	A	C	A	A	C	A
f	f	g	C	C	C	C	C	D	C
g	g	b	C	C	A	D	D	A	D

Automatová kongruence

Nechť $\equiv$ je relace ekvivalence na Q. Říkáme, že $\equiv$ je *automatovou kongruencí*, jestliže:

$\forall p,q \in Q \;\; p \equiv q \Rightarrow (p \in F \Leftrightarrow q \in F) \;\&\; \forall x \in X \;\; (\delta(p,x) \equiv \delta(q,x))$



Tvrzení: Ekvivalence stavů ~ je automatovou kongruencí.

Důkaz: nechť $p \sim q$, potom: $p \in F \Leftrightarrow q \in F$ (položme $w = \lambda$, $\delta^*(p,\lambda) = p$)

nechť $p \sim q$, potom: $\forall x \in X \;\forall u \in X^* \;(\delta^*(p,xu) \in F \Leftrightarrow \delta^*(q,xu) \in F)$

$\Leftrightarrow \forall x \in X \;\forall u \in X^* \;(\delta^*(\delta(p,x),u) \in F \Leftrightarrow \delta^*(\delta(q,x),u) \in F)$

$\Leftrightarrow \forall x \in X \;\; \delta(p,x) \sim \delta(q,x)$

Podílový automat

$A = (Q,X,\delta,q_0,F)$ je konečný automat a $\equiv$ je automatová kongruence.

Potom $A/\equiv \; = \; (Q/\equiv, X, \delta_\equiv, [q_0]_\equiv, \{[q]_\equiv \mid q \in F\})$, kde $\delta_\equiv([q]_\equiv,x) = [\delta(q,x)]_\equiv$ je konečný automat (nazvěme ho *podílovým automatem*) ekvivalentní s A.

1) Je $A/\equiv$ skutečně konečný automat?

Množiny $Q/\equiv$ a X jsou neprázdné a konečné.

$\delta_\equiv$ je definována korektně (vlastnosti automatové kongruence)

2) Jsou oba automaty ekvivalentní?

Stačí najít homomorfismus A do $A/\equiv$ (věta o ekvivalenci automatů)!

Definujme $h: Q \to Q/\equiv$ takto $h(q) = [q]_\equiv$

$h(q_0) = [q_0]_\equiv$

$h(\delta(q,x)) = [\delta(q,x)]_\equiv = \delta_\equiv([q]_\equiv,x) = \delta_\equiv(h(q),x)$

$q \in F \Leftrightarrow h(q) = [q]_\equiv \in F_\equiv$ ($F_\equiv$ jsou koncové stavy automatu $A/\equiv$)

Podílový automat a ekvivalence stavů

A je konečný automat a ~ je ekvivalence stavů.

Potom A/~ je konečný automat ekvivalentní s A a žádné stavy A/~ nejsou ekvivalentní.

1) Ekvivalence stavů ~ je automatová kongruence, a tedy víme, že A/~ je konečný automat ekvivalentní s A.

2) V A/~ nejsou ekvivalentní stavy.

Sporem: nechť $[p]_\sim$ a $[q]_\sim$ jsou různé ekvivalentní stavy (tj. $\neg\, p \sim q$)

vezměme libovolné $w \in X^*$:

$\delta_\sim([p]_\sim,w) \in F_\sim \Leftrightarrow \delta_\sim([q]_\sim,w) \in F_\sim$ ($[p]_\sim$ a $[q]_\sim$ jsou ekvivalentní)

$\delta_\sim(h(p),w) \in F_\sim \Leftrightarrow \delta_\sim(h(q),w) \in F_\sim$ ($h(p) = [p]_\sim$)

$h(\delta(p,w)) \in F_\sim \Leftrightarrow h(\delta(q,w)) \in F_\sim$ (h je homomorfismus)

$\delta(p,w) \in F \Leftrightarrow \delta(q,w) \in F$

$p \sim q$ spor

Redukce automatu

Konečný automat je *redukovaný*, jestliže:

- nemá nedosažitelné stavy,
- žádné dva stavy nejsou ekvivalentní.

Konečný automat B je *reduktem* automatu A, jestliže:

- B je redukovaný,
- A a B jsou ekvivalentní.

Věta: Ke každému konečnému automatu existuje nějaký jeho redukt.

Konstruktivní důkaz:

1) vyřazení nedosažitelných stavů
   faktorizace podle ekvivalence stavů (nezpůsobí nedosažitelnost)
2) faktorizace podle ekvivalence stavů
   vyřazení nedosažitelných stavů (nezpůsobí změnu ekvivalence)

Příklad redukce automatů

Co dělá tento automat (jaký jazyk přijímá)?

		a	b	$\mathcal{R}_0$	a	b	$\mathcal{R}_1$	a	b	$\mathcal{R}_2$
→	1	2	3	I	I	III	I	II	III	I
	2	2	4	I	I	I	II	II	II	II
←	3	3	5	III	III	III	III	III	III	III
	4	2	7	I	I	I	II	II	II	II
←	5	6	3	III	III	III	III	III	III	III
←	6	6	6	III	III	III	III	III	III	III
	7	7	4	I	I	I	II	II	II	II
	~~8~~	~~2~~	~~3~~	~~I~~	~~I~~	~~III~~	~~I~~	~~II~~	~~III~~	~~I~~
	~~9~~	~~9~~	~~4~~	~~I~~	~~I~~	~~I~~	~~II~~	~~II~~	~~II~~	~~II~~

Nedosažitelné stavy

Redukovaný automat

		a	b
→	I	II	III
	II	II	II
←	III	III	III

L = {w | w=bu, u∈{a,b}*}

Věta o isomorfismu reduktů

Pro libovolná dva redukované konečné automaty jsou následující dvě tvrzení ekvivalentní:

a) automaty jsou ekvivalentní,
b) automaty jsou isomorfní.

Důsledky:

Dva redukty libovolných dvou ekvivalentních konečných automatů se shodují až na isomorfismus.

Pro každý konečný automat je jeho redukt určen až na isomorfismus jednoznačně.

Ve třídě navzájem ekvivalentních konečných automatů existuje „minimální“ automat.

Důkaz věty o isomorfismu reduktů

a) isomorfismus ⇒ ekvivalence (víme)

b) ekvivalence reduktů ⇒ isomorfismus

hledáme homomorfismus $h:Q_1 \rightarrow Q_2$, který je „prostý a na“ tj. pro každé $q \in Q_1$ hledáme právě jedno $p \in Q_2$

q je dosažitelný stav, tudíž $\exists u \in X^*\ \delta_1(q_1,u)=q$

položme $h(q)= \delta_2(q_2,u)$

je to skutečně funkce?

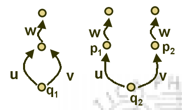
$\delta_1(q_1,u)=\delta_1(q_1,v) \Leftrightarrow \delta_2(q_2,u)=\delta_2(q_2,v)$ (*)

sporem, nechť $\delta_1(q_1,u)=\delta_1(q_1,v)\ \&\ \delta_2(q_2,u)\neq\delta_2(q_2,v)$

z A_1 víme $\forall w \in X^*\ uw \in L \Leftrightarrow vw \in L$

automaty jsou ekvivalentní, tedy $p_1 \sim p_2$

spor - automat A_2 je redukovaný



funkce h je „prostá a na“ (vlastnost (*))

$h(q_1)=q_2$ (pro $u=\lambda$)

$h(\delta_1(q,x)) = \delta_2(h(q),x)$ ($\delta_1(q_1,v)=q$, $u=vx$)

$q \in F_1 \Leftrightarrow h(q) \in F_2$ (pro $u \in L$ + ekvivalentní automaty)

Normalizace automatu

Jak najít isomorfismus automatů?

Normovaný tvar automatu

1) fixujme pořadí písmen v abecedě
2) počáteční stav označme 1
3) tabulku (automatu) vyplňujme po řádcích zleva doprava a pokud narazíme na nový stav, přiřadíme mu první volné číslo

Příklad:

		a	b
	A	B	A
↔	B	D	C
	C	A	D
←	D	A	B

		a	b
↔	1(B)	2	3
←	2(D)	4	1
	3(C)	4	2
	4(A)	1	4

Poznámky k redukci a ekvivalenci

Algoritmicky umíme řešit:

- zjištění ekvivalence automatů
  zredukujeme, znormalizujeme a porovnáme
- zjištění zda L(A)=∅
  žádný koncový stav není dosažitelný
- zjištění zda L(A)=X*
  po redukci dostaneme jednostavový automat (s koncovým stavem)

Umíme najít nejkratší slovo rozlišující dva stavy

$p \sim^i q\ \&\ \neg\ p \sim^{i+1} q$

$\exists a_1 \in X\quad \delta(p,a_1) \sim^{i-1} \delta(q,a_1)\ \ \&\ \ \neg\ \delta(p,a_1) \sim^i \delta(q,a_1)$

…

iterací najdeme slovo $a_1 \ldots a_{i+1}$