Automaty a gramatiky

8

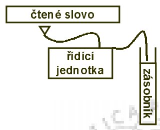
Roman Barták, KTIML

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

Co bylo minule

Bezkontextové gramatiky
- kanonické derivace (levá a pravá)
- derivační stromy
- jednoznačnost a nejednoznačnost gramatiky (jazyka)
- analýza shora

Zásobníkové automaty
- výpočet automatu
  - přijímání koncovým stavem
  - přijímání prázdným zásobníkem



Jak je to s determinismem/nedeterminismem ZA?

Jaký má zásobníkový automat vztah k bezkontextové gramatice?

Deterministické zásobníkové automaty

Jak je to s nedeterminismem zásobníkových automatů?
- Je nutný!
- Pro rozpoznání ww^R potřebujeme nedeterministicky uhádnout střed.

Kde je skryt nedeterminismus zásobníkového automatu?
- množina možných přechodů
- opracování zásobníku bez čtení vstupu (λ-přechod)

Definice: Říkáme, že *zásobníkový automat* $M=(Q,X,Y,\delta,q_0,Z_0,F)$, je *deterministický*, jestliže platí:
- $\forall p\in Q,\ \forall a\in X\cup\{\lambda\},\ \forall Z\in Y \qquad |\delta(p,a,Z)|\leq 1$
- $\forall p\in Q,\ \forall Z\in Y\ (\ \delta(p,\lambda,Z)\neq\emptyset \Rightarrow \forall a\in X\ \delta(p,a,Z)=\emptyset\ )$

Každý krok výpočtu je přesně určen.

Příklady zásobníkových automatů

Deterministický zásobníkový automat (prázdný zásobník)

$L = \{0^n1^n \mid n>0\}$

$\delta(p,0,Z) = \{(p,A)\}$ … čte první symbol 0
$\delta(p,0,A) = \{(p,AA)\}$ … čte další symboly 0
$\delta(p,1,A) = \{(q,\lambda)\}$ … čte první symbol 1
$\delta(q,1,A) = \{(q,\lambda)\}$ … čte další symboly 1

(Klasický) zásobníkový automat (prázdný zásobník)

$L = \{ww^R \mid w\in\{0,1\}^*\}$

$\delta(p,0,X) = \{(p,AX)\}$ … čte 0 v první půlce (uschovává na zásobníku)
$\delta(p,1,X) = \{(p,BX)\}$ … čte 1 v první půlce (uschovává na zásobníku)
$\delta(p,\lambda,X) = \{(q,X)\}$ … překlopení do druhé půlky (nedeterminismus)
$X\in\{Z,A,B\}$
$\delta(q,0,A) = \{(q,\lambda)\}$ … čte 0 symetricky v druhé půlce
$\delta(q,1,B) = \{(q,\lambda)\}$ … čte 1 symetricky v druhé půlce
$\delta(q,\lambda,Z) = \{(q,\lambda)\}$ … ukončuje výpočet (není nedeterminismus)

Deterministické a bezprefixové jazyky

Již víme, že determinismus je u ZA slabší než nedeterminismus (ww^R).

Jak je to s přijímáním slov?

Deterministické bezkontextové jazyky jsou jazyky rozpoznávané DZA koncovým stavem.

Tvrzení: Regulární jazyk je deterministický.
Zásobník nemusíme využívat.

Bezprefixové bezkontextové jazyky jsou jazyky rozpoznávané DZA prázdným zásobníkem.

Tvrzení: Bezprefixový jazyk je deterministický.
Převod nepřidává nedeterminismus.

Pozorování: M je DZA, $u\in N(M) \Rightarrow \forall w\in X^+\ uw\notin N(M)$
Jakmile jednou vyprázdníme zásobník (po přečtení *u*, které je přijato), nemůže výpočet pokračovat (čtením *w*).

Koncové stavy vs. prázdný zásobník

Je také každý deterministický jazyk bezprefixový?

NE!

Z vlastnosti M je DZA, $u\in N(M) \Rightarrow \forall w\in X^+\ uw\notin N(M)$ snadno sestrojíme příslušný jazyk.

Potřebujeme:
základní deterministický jazyk, který
obsahuje slovo, jež je prefixem jiného přijímaného slova

Například $\{0^n1^m \mid 0<n\leq m\}$ (uděláme DZA, q_F je koncový stav).

$\delta(p,0,Z) = \{(p,AZ)\}$ … čte první symbol 0
$\delta(p,0,A) = \{(p,AA)\}$ … čte další symboly 0
$\delta(p,1,A) = \{(q,\lambda)\}$ … čte první symbol 1
$\delta(q,1,A) = \{(q,\lambda)\}$ … čte další symboly 1
$\delta(q,\lambda,Z) = \{(q_F,Z)\}$ … nyní se počet 0 a 1 vyrovnal
$\delta(q_F,1,Z) = \{(q_F,Z)\}$ … dále čteme jen 1
to už „prázdný zásobník“ nemůže

Jak na bezprefixové jazyky?

Jak se od deterministického jazyka dostaneme k bezprefixovému?

Co nám vadí?

po přečtení prefixu, který patří do jazyka, máme prázdný zásobník.

Řešení:

přidáme speciální znak na konec slova (až po přečtení tohoto znaku vyprázdníme zásobník)

Nechť $L \subseteq X^*$ je deterministický jazyk a $\# \notin X$, potom $L\#$ je bezprefixový jazyk.

\# opraví nedeterministické λ-kroky při převodu automatů

Příklad: $\{0^n1^m\# \mid 0<n\leq m\}$

$\delta(p,0,Z) = \{(p,AZ)\}$ … čte první symbol 0

$\delta(p,0,A) = \{(p,AA)\}$ … čte další symboly 0

$\delta(p,1,A) = \{(q,\lambda)\}$ … čte první symbol 1

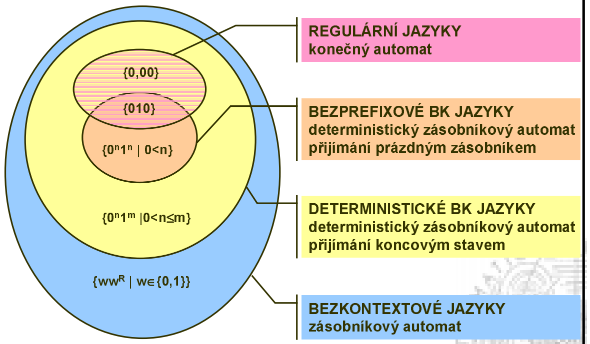
$\delta(q,1,A) = \{(q,\lambda)\}$ … čte další symboly 1

$\delta(q,\lambda,Z) = \{(q_{OK},Z)\}$ … nyní se počet 0 a 1 vyrovnal

$\delta(q_{OK},1,Z) = \{(q_{OK},Z)\}$ … dále čteme jen 1

$\delta(q_{OK},\#,Z) = \{(q_{OK},\lambda)\}$ … na konci vyprázdníme zásobník

Jazyky a automaty - přehled



Od gramatiky k automatu

Každý bezkontextový jazyk je rozpoznáván zásobníkovým automatem prázdným zásobníkem.

Důkaz (konstruktivní):

idea:

- ze vstupu čteme terminály, na zásobníku terminály i neterminály
- pravidla gramatiky → pravidla automatu (opracování zásobníku)
- přidáme pravidla pro krácení terminálu na vstupu a na zásobníku
- stačí jediný stav

formálně:

$G=(V_N,V_T,S,P) \rightarrow M=(\{p\},V_T,V_N\cup V_T,\delta,p,S,\{\})$

$\delta(p,\lambda,X)=\{(p,w) \mid (X\rightarrow w)\in P\}$ $\forall X\in V_N$ … opracování zásobníku

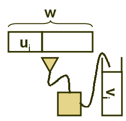
$\delta(p,a,a)=\{(p,\lambda)\}$ $\forall a\in V_T$ … krácení terminálů

$A \rightarrow bAc \mid \lambda$ … $\delta(p,\lambda,A) = \{ (p,bAc), (p,\lambda) \}$

Od gramatiky k automatu - pokračování

$L(G)\supseteq N(M)$?

Situaci v kroku *i* popíšeme pomocí slova $w_i=u_iv_i$



u_i - dosud přečtená část vstupního slova

v_i - ještě nezpracovaná část (zásobník)

provedení kroku

a) $v_i = Zv_i'$ kde $Z\in V_N$, potom

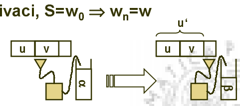
$u_{i+1} = u_i$ (nic nečteme), $v_{i+1} = vv_i'$ (dle pravidla gramatiky $Z\rightarrow v$)

zřejmě tedy $w_i \Rightarrow w_{i+1}$ (přímý přepis v gramatice)

b) $v_i = av_i'$ kde $a\in V_T$, potom $u_{i+1} = u_ia$, $v_{i+1} = v_i'$, tedy $w_{i+1} = w_i$

výpočet automatu tedy definuje derivaci, $S=w_0 \Rightarrow w_n=w$

$L(G)\subseteq N(M)$?



vezmeme levé derivace

$u\alpha \Rightarrow^* u'\beta$, kde $u'=uv$, $u,u'\in V_T^*$ potom $(p,v,\alpha) \vdash^* (p,\lambda,\beta)$

dohromady $S \Rightarrow^* w$ a tedy $(p,w,S) \vdash^* (p, \lambda, \lambda)$

Převod BKG na ZA - příklad

$L= \{a^ib^jc^k \mid i+j=k\}$

Gramatika G = ({S,A},{a,b,c},S,P)

$S \rightarrow aSc \mid A$

$A \rightarrow bAc \mid \lambda$

$S \Rightarrow aSc \Rightarrow aAc \Rightarrow abAcc \Rightarrow abbAccc \Rightarrow abbccc$

Zásobníkový automat M=({p},{a,b,c},{S,A,a,b,c},δ,p,S)

$\delta(p,\lambda,S) = \{ (p,aSc), (p,A) \}$

$\delta(p,\lambda,A) = \{ (p,bAc), (p,\lambda) \}$

$\delta(p,x,x) = \{ (p,\lambda) \}$ $\forall x\in\{a,b,c\}$

$(p,abbccc,S) \vdash (p,abbccc,aSc) \vdash (p,bbccc,Sc) \vdash (p,bbccc,Ac)$
$\vdash (p,bbccc,bAcc) \vdash (p,bccc,Acc) \vdash (p,bccc,bAccc) \vdash (p,ccc,Accc)$
$\vdash (p,ccc,ccc) \vdash (p,cc,cc) \vdash (p,c,c) \vdash (p, \lambda, \lambda)$