Automaty a gramatiky (BI-AAG)

9. Překladové automaty

Jan Holub

Katedra teoretické informatiky

Fakulta informačních technologií

ČVUT v Praze

© Jan Holub, 2014

Formální překlady

Definice

Formální překlad je binární relace $Z \subseteq L \times V$.
Definiční obor je množina L a obor hodnot je množina V.
Relace Z přiřazuje každému prvku w množiny L množinu jeho překladů $Z(w)$ z množiny V.
Pokud $Z(w)$ obsahuje pro každé $w \in L$ nejvýše jeden prvek, pak Z je funkce (případně parciální) a takový překlad se nazývá jednoznačný.

Formální překlady

Definice

T a D jsou abecedy. *Homomorfismem* nazýváme každé zobrazení h z T do D^*. Definiční obor homomorfismu h je možné rozšířit na T^* takto:

$h(\varepsilon) = \varepsilon$,

$h(xa) = h(x)h(a), \forall x, x \in T^*, a \in T$.

Formální překlady

Příklad

zobrazení řetězce dekadických číslic na řetězce binárních číslic

a	0	1	2	3	4	5	6	7	8	9
$h(a)$	0000	0001	0010	0011	0100	0101	0110	0111	1000	1001

$$h(1996) = 0001100110010110$$

Formální překlady

Definice

Prefixový a *postfixový zápis výrazu* E:

1. E je proměnná nebo konstanta, pak prefixový i postfixový zápis tohoto výrazu je E.
2. E je výraz tvaru $E_1 \; op \; E_2$, kde op je binární operátor, pak
   (a) $op \; E'_1 E'_2$ je prefixový zápis, kde E'_1 a E'_2 jsou prefixové zápisy.
   (b) $E''_1 E''_2 op$ je postfixový zápis, kde E''_1 a E''_2 jsou postfixové zápisy.
3. E je výraz tvaru (E_1), pak
   (a) prefixový zápis výrazu (E_1) je prefixový zápis výrazu E_1,
   (b) postfixový zápis výrazu (E_1) je postfixový zápis výrazu E_1.

Formální překlady

Příklad

Infixový zápis: $a * (b + c)$

Prefixový zápis: $*a + bc$

Postfixový zápis: $abc + *$

Příklad překladu:

$\{(x, y)$: x je infixový zápis výrazu, y je postfixový zápis téhož výrazu$\}$.

Překladové gramatiky

Definice

Překladová gramatika je $PG = (N, T, D, R, S)$, kde

- N je konečná množina neterminálních symbolů,
- T je konečná množina vstupních symbolů,
- D je konečná množina výstupních symbolů,
- R je konečná množina pravidel tvaru $A \to \alpha$, kde $A \in N$, $\alpha \in (N \cup T \cup D)^*$,
- S je počáteční symbol.

Přitom platí, že $T \cap D = \emptyset$ a $(T \cup D) \cap N = \emptyset$.

Překladové gramatiky

Definice

$PG = (N, T, D, R, S)$, $\alpha, \beta, \gamma \in (N \cup T \cup D)^*$, $A \in N$.

1. z α se *přímo derivuje* β, $\alpha \Rightarrow \beta$, právě když
   $\exists \tau, \omega, \gamma \in (N \cup T \cup D)^*$, $A \in N$, $\alpha = \tau A \, \omega$, $\beta = \tau \gamma \omega$,
   $A \ \rightarrow \gamma \in R$
2. z α se *derivuje* β, $\alpha \Rightarrow^* \beta$, právě když
   $\exists \alpha_1, \alpha_2, ..., \alpha_n \in (N \cup T \cup D)^*, (n \geq 1)$
   $\alpha = \alpha_1 \Rightarrow \alpha_2 \Rightarrow ... \Rightarrow \alpha_n = \beta$.

Posloupnost $\alpha_1, \alpha_2, ..., \alpha_n$ nazýváme překladovou derivací délky n řetězce β z řetězce α.

reflexivní a tranzitivní uzávěr: $\Rightarrow^*$

tranzitivní uzávěr: $\Rightarrow^+$

Překladové gramatiky

Příklad

$PG = (\{E\}, \{+, *, a\}, \{\oplus, \circledast, @\}, P, E)$, kde P:

$(1) E \to +EE\oplus \qquad (2) E \to *EE\circledast \qquad (3) E \to a@$

$$\begin{aligned} E &\Rightarrow +EE\oplus \\ &\Rightarrow +a@E\oplus \\ &\Rightarrow +a@ * EE\circledast\oplus \\ &\Rightarrow +a@ * a@E\circledast\oplus \\ &\Rightarrow +a@ * a@a@\circledast\oplus \end{aligned}$$

Překladové gramatiky

Definice

$PG = (N, T, D, R, S)$.

Vstupní homomorfismus h_i^{PG}: $h_i^{PG}(a) = \begin{cases} a & \text{pro } a \in T \cup N \\ \varepsilon & \text{pro } a \in D \end{cases}$

Výstupní homomorfismus h_o^{PG}: $h_o^{PG}(a) = \begin{cases} \varepsilon & \text{pro } a \in T \\ a & \text{pro } a \in D \cup N \end{cases}$

Definice

Překlad definovaný překladovou gramatikou $PG = (N, T, D, R, S)$:

$Z(PG) = \{(h_i^{PG}(w), h_o^{PG}(w)) : S \Rightarrow^* w, w \in (T \cup D)^*\}$.

Překladové gramatiky

Příklad

$PG = (\{E\}, \{+, *, a\}, \{\oplus, \circledast, ⓐ\}, P, E)$, kde P:

$(1) E \rightarrow +EE\oplus \qquad (2) E \rightarrow *EE\circledast \qquad (3) E \rightarrow a@$

PG generuje překlad výrazů z prefixového do postfixového zápisu.

$$
\begin{aligned}
E &\Rightarrow +EE\oplus \\
&\Rightarrow +a@E\oplus \\
&\Rightarrow +a@*EE\circledast\oplus \\
&\Rightarrow +a@*a@E\circledast\oplus \\
&\Rightarrow +a@*a@a@\circledast\oplus
\end{aligned}
$$

Derivace generuje dvojici $(+a*aa, @@@\circledast\oplus)$, která patří do překladu $Z(PG)$.

Překladové gramatiky

Definice

$PG = (N, T, D, R, S)$.

Vstupní gramatika překladové gramatiky PG je BG $G_i = (N, T, P_i, S)$, kde $P_i = \{A \to h_i(\alpha) : A \to \alpha \in R\}$.

Výstupní gramatika překladové gramatiky PG je BG $G_o = (N, D, P_o, S)$, kde $P_o = \{A \to h_o(\alpha) : A \to \alpha \in R\}$.

Definice

BG $G = (N, T \cup D, R, S)$ je *charakteristická gramatika překladové gramatiky* $PG = (N, T, D, R, S)$.

$L(G)$ je *charakteristický jazyk překladu Z(PG)*.

$w \in L(G)$ je *charakteristická věta dvojice* (x, y), kde $x = h_i(w)$, $y = h_o(w)$.

Překladové gramatiky

Definice
$PG = (N, T, D, R, S)$ je *regulární*, když všechna pravidla v R mají tvar:

- $A \rightarrow axB$ nebo $A \rightarrow ax$, kde $A, B \in N, a \in T, x \in D^*$
- $S \rightarrow \varepsilon$ v případě, že S není na pravé straně žádného pravidla.

Překladové gramatiky

Příklad

$RPG = (\{S, A, P, K\}, \{a, +, *\}, \{@, \oplus, \circledast\}, R, S)$, kde R:

$$
\begin{array}{ll}
S \to a@A & A \to *K \\
S \to a@ & A \to +P \\
K \to a@\circledast A & P \to a@\oplus A \\
K \to a@\circledast & P \to a@\oplus
\end{array}
$$

$$
\begin{aligned}
S &\Rightarrow a@A \\
&\Rightarrow a@ + P \\
&\Rightarrow a@ + a@\oplus A \\
&\Rightarrow a@ + a@\oplus * K \\
&\Rightarrow a@ + a@\oplus * a@\circledast.
\end{aligned}
$$

Překlad: $(a + a * a, @@\oplus@\circledast)$.

Konečné překladové automaty

Definice

Konečný překladový automat $KPA = (Q, T, D, \delta, q_0, F)$, kde

- Q je konečná množina vnitřních stavů,
- T je konečná množina vstupních symbolů,
- D je konečná množina výstupních symbolů,
- δ je zobrazení z $Q \times (T \cup \{\varepsilon\})$ do množiny $2^{Q \times D^*}$,
- $q_0 \in Q$ je počáteční stav,
- $F \subseteq Q$ je množina koncových stavů.

Konečné překladové automaty

Definice

Konfigurace konečného překladového automatu $KPA = (Q, T, D, \delta, q_0, F)$ je trojice $(q, x, y) \in Q \times T^* \times D^*$.

(q_0, x, ε) je *počáteční konfigurace*.

(q, ε, y), $q \in F$ je *koncová konfigurace*.

Relace přechodu $\vdash\ a$ definovaná na množině konfigurací:

Jestliže $\delta(q, a)$ obsahuje (r, z), pak $(q, ax, y) \vdash (r, x, yz)$, $x \in T^*, y \in D^*$.

Tranzitivní uzávěr: $\vdash^+$

Reflexivní a tranzitivní uzávěr: $\vdash^*$

k-tá mocnina: $\vdash^k$

Konečné překladové automaty

Definice

Překlad definovaný konečným překladovým automatem $KPA = (Q, T, D, \delta, q_0, F)$:
$Z(KPA) = \{(u, v) : (q_0, u, \varepsilon) \vdash^* (q, \varepsilon, v), q \in F\}$

Definice

KPA je *deterministický*, když pro všechny jeho stavy platí:

1. $|\delta(q, a)| \leq 1, \forall a \in T$ a $\delta(q, \varepsilon) = \emptyset$ nebo
2. $|\delta(q, \varepsilon)| \leq 1$ a $\delta(q, a) = \emptyset, \forall a \in T$.

Konečné překladové automaty

Příklad

KPA, který binární čísla dělitelná třemi dělí třemi.

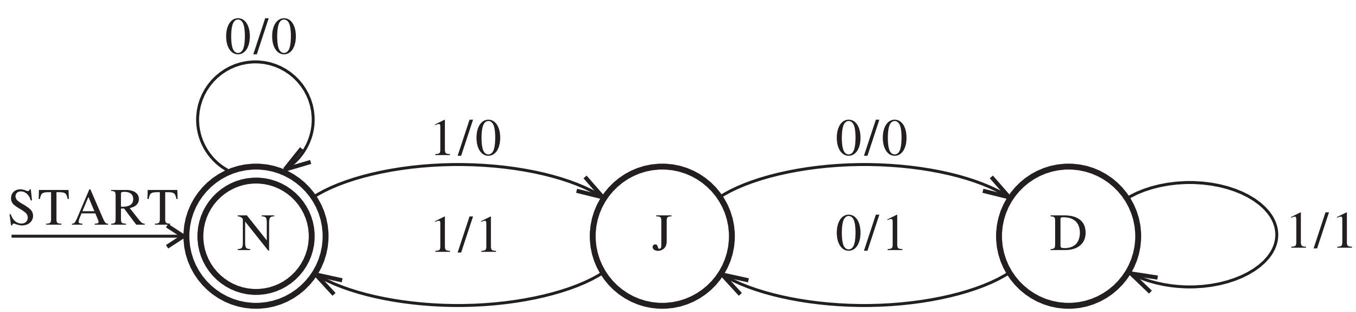
$KPA = (\{N, J, D\}, \{0, 1\}, \{0, 1\}, \delta, N, \{N\})$, kde δ:

$\delta(N, 0) = \{(N, 0)\}$ $\quad$ $\delta(J, 1) = \{(N, 1)\}$

$\delta(N, 1) = \{(J, 0)\}$ $\quad$ $\delta(D, 0) = \{(J, 1)\}$

$\delta(J, 0) = \{(D, 0)\}$ $\quad$ $\delta(D, 1) = \{(D, 1)\}$

δ	0	1
N	$(N, 0)$	$(J, 0)$
J	$(D, 0)$	$(N, 1)$
D	$(J, 1)$	$(D, 1)$



Převod RPG na KPA

Věta

Je-li dána $RPG = (N, T, D, R, S)$, pak existuje $KPA = (Q, T, D, \delta, q_0, F)$ takový, že $Z(RPG) = Z(KPA)$.

Důkaz: Pro zadanou $RPG = (N, T, D, R, S)$ sestrojíme $KPA = (Q, T, D, \delta, q_0, F)$, kde $Q = N \cup \{X\}$, $X \notin N$.
Zobrazení δ je definováno takto ($y \in D^*$, $B, C \in N$):

$$(C, y) \in \delta(B, a), \text{ jestliže } B \rightarrow ayC \in R, \forall a \in T,$$
$$(X, y) \in \delta(B, a), \text{ jestliže } B \rightarrow ay \in R, \forall a \in T,$$

$q_0 = S$
$F = \{S, X\}$, jestliže $S \rightarrow \varepsilon \in R$
$F = \{X\}$, jestliže $S \rightarrow \varepsilon \notin R$.

Důkaz, že $Z(RPG) = Z(KPA)$: indukcí podle délky derivace v RPG a podle délky posloupnosti přechodů KPA. $\square$

Převod RPG na KPA

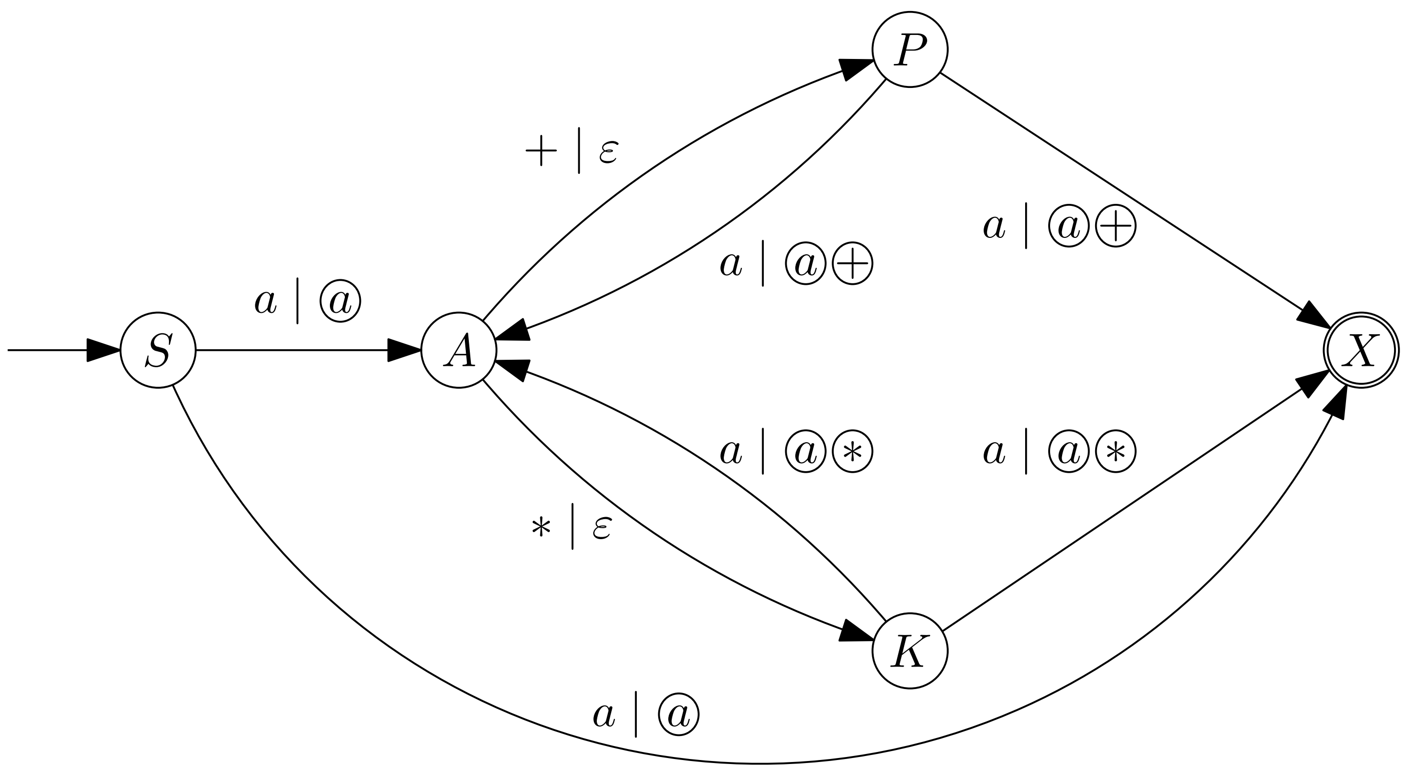
Příklad

$RPG = (\{S, A, P, K\}, \{a, +, *\}, \{@, \oplus, \circledast\}, R, S)$, kde R:

$S \rightarrow a@A \qquad A \rightarrow *K \qquad S \rightarrow a@ \qquad A \rightarrow +P$

$K \rightarrow a@\circledast A \qquad P \rightarrow a@\oplus A \qquad K \rightarrow a@\circledast \qquad P \rightarrow a@\oplus$

$KPA = (\{S, X, A, P, K\}, \{a, +, *\}, \{@\oplus\circledast\}, \delta, S, \{X\})$



Převod KPA na RPG

Věta
Je-li dán konečný překladový automat KPA, pak existuje regulární překladová gramatika RPG taková, že $Z(KPA) = Z(RPG)$.

Důkaz: Pro zadaný $KPA = (Q, T, D, \delta, q_0, F)$ sestrojíme $RPG = (N \cup \{S\}, T, D, R, S)$, kde $S \notin N$, takto:

1. $N = Q$.
2. Vytvoříme množinu pravidel R', pro všechna $\forall a \in T$ a $y \in D^*$:
   $B \to ayC$, když $(C, y) \in \delta(B, a)$,
   $B \to ay$, když $(C, y) \in \delta(B, a)$ a $C \in F$,
   $S \to \varepsilon$, když $q_0 \in F$.
3. $R = R' \cup \{S \to x : q_0 \to x \in R'\}$.

Důkaz rovnosti $Z(RPG) = Z(KPA)$: indukcí podle délky derivace v RPG a podle délky posloupnosti přechodů KPA. $\square$

Převod KPA na RPG

Příklad

$KPA = (\{N, J, D\}, \{0, 1\}, \{⓪, ①\}, \delta, N, \{N\})$, kde δ:

$\delta(N, 0) = \{(N, ⓪)\}$ $\quad$ $\delta(J, 1) = \{(N, ①)\}$

$\delta(N, 1) = \{(J, ⓪)\}$ $\quad$ $\delta(D, 0) = \{(J, ①)\}$

$\delta(J, 0) = \{(D, ⓪)\}$ $\quad$ $\delta(D, 1) = \{(D, ①)\}$

$RPG = (\{S, N, J, D\}, \{0, 1\}, \{⓪, ①\}, R, S)$, kde R:

$S \to \varepsilon$	$N \to 0⓪N$	$J \to 0⓪D$	$D \to 0①J$
$S \to 0⓪N$	$N \to 1⓪J$	$J \to 1①N$	$D \to 1①D$
$S \to 1⓪J$	$N \to 0⓪$	$J \to 1①$	$S \to 0⓪$

Sekvenční zobrazení

Definice

Sekvenční zobrazení S:

1. Zachovává délku řetězce, tj. je-li $y = S(x)$, pak $|x|$ = $|y|$.
2. Mají-li dva vstupní řetězce stejnou předponu délky $k > 0$, pak mají i jim odpovídající výstupní řetězce stejné předpony alespoň délky k. To znamená, když $S(xx_1) = yy_1$ a $S(xx_2) = yy_2$ a $|x| = |y|$.

Sekvenční zobrazení

Definice

Mealyho automat $M = (Q, T, D, \delta, \lambda, q_o, F)$, kde

- Q je konečná množina vnitřních stavů,
- T je konečná množina vstupních symbolů,
- D je konečná množina výstupních symbolů,
- δ je zobrazení z $Q \times T$ do Q nazývané *přechodová funkce*,
- λ je zobrazení z $Q \times T$ do D nazývané *výstupní funkce*,
- $q_0 \in Q$ je počáteční stav,
- $F \subseteq Q$ je množina koncových stavů.

Sekvenční zobrazení

Definice

Mooreův automat $M = (Q, T, D, \delta, \lambda, q_o, F)$, kde

- Q je konečná množina vnitřních stavů,
- T je konečná množina vstupních symbolů,
- D je konečná množina výstupních symbolů,
- δ je zobrazení z $Q \times T$ do Q nazývané *přechodová funkce*,
- λ je zobrazení z Q do D nazývané *výstupní funkce*,
- $q_0 \in Q$ je počáteční stav,
- $F \subseteq Q$ je množina koncových stavů.

Zásobníkové překladové automaty

Definice

Zásobníkový překladový automat je osmice
$ZPA = (Q, T, G, D, \delta, q_0, Z_0, F)$, kde

- Q je konečná množina vnitřních stavů,
- T je konečná množina symbolů (*vstupní symboly*),
- D je konečná množina symbolů (*výstupní symboly*),
- G je konečná množina symbolů (*zásobníkové symboly*),
- δ je konečné zobrazení z $Q \times (T \cup \{\varepsilon\}) \times G^*$ do množiny konečných podmnožin $Q \times G^* \times D^*$,
- $q_0 \in Q$ je počáteční stav,
- $Z_0 \in G$ je počáteční symbol zásobníku,
- $F \subseteq Q$ je množina koncových stavů. $\square$

Zásobníkové překladové automaty

Definice

Konfiguraci zásobníkového překladového automatu $ZPA = (Q, T, G, D, \delta, q_0, Z_0, F)$ definujme jako čtveřici $(q, x, \alpha, y) \in Q \times T^* \times G^* \times D^*$.
$(q_0, x, Z_0, \varepsilon)$ – *počáteční konfigurace* (q_0 je počáteční stav a Z_0 je počáteční symbol zásobníku)
relace přechodu $\vdash$: Jestliže $\delta(q, a, u)$ obsahuje (r, α, v), pak budeme psát $(q, ax, u\gamma, y) \vdash (r, x, \alpha\gamma, yv)$ pro libovolné $x \in T^*, \gamma \in G^*, y \in D^*$. $\square$

Zásobníkové překladové automaty

Definice

Překlad definovaný zásobníkovým překladovým automatem ZPA přechodem do koncového stavu je množina dvojic
$Z(ZPA) = \{(x, y) : (q_0, x, Z_0, \varepsilon) \vdash^* (q, \varepsilon, \alpha, y), q \in F, \alpha \in G^*\}$.
Překlad definovaný zásobníkovým překladovým automatem ZPA přechodem do konfigurace s prázdným zásobníkem je množina dvojic
$Z_\varepsilon(ZPA) = \{(x, y) : (q_0, x, Z_0, \varepsilon) \vdash^* (q, \varepsilon, \varepsilon, y), q \in Q\}$. □

Zásobníkové překladové automaty

Příklad

$ZPA = (\{q\}, \{a, +, *\}, \{+, *, E\}, \{a, +, *\}, \delta, q, E, \{q\})$, kde δ:

$\delta(q, a, E) = \{(q, \varepsilon, a)\}$

$\delta(q, +, E) = \{(q, EE+, \varepsilon)\}$

$\delta(q, *, E) = \{(q, EE*, \varepsilon)\}$

$\delta(q, \varepsilon, +) = \{(q, \varepsilon, +)\}$

$\delta(q, \varepsilon, *) = \{(q, \varepsilon, *)\}$.

Pro vstup $+ * aaa$ provede tento automat posloupnost přechodů

$$
\begin{array}{lrrrr}
(q, + * aaa, E, \varepsilon) & \vdash (q, & *aaa, & EE+, & \varepsilon) \\
& \vdash (q, & aaa, & EE * E+, & \varepsilon) \\
& \vdash (q, & aa, & E * E+, & a) \\
& \vdash (q, & a, & *E+, & aa) \\
& \vdash (q, & a, & E+, & aa*) \\
& \vdash (q, & \varepsilon, & +, & aa * a) \\
& \vdash (q, & \varepsilon, & \varepsilon, & aa * a+).
\end{array}
$$

Zásobníkové překladové automaty

Definice
Zásobníkový překladový automat $ZPA = (Q, T, G, D, \delta, q_0, Z_0, F)$ se nazývá *deterministický*, jestliže platí následující podmínky:

(a) $\delta(q, a, Z)$ obsahuje nejvýše jeden prvek pro všechna $q \in Q, a \in T \cup \{\varepsilon\}$ a $Z \in G^*$,
(b) je-li $\delta(q, a, \alpha) \neq \emptyset$ a $\delta(q, a, \beta) \neq \emptyset$ a $\alpha \neq \beta$, pak ani jeden z řetězců α, β není prefixem druhého.
(c) je-li $\delta(q, a, \alpha) \neq \emptyset$ a $\delta(q, \varepsilon, \beta) \neq \emptyset$, pak ani jeden z řetězců α, β není prefixem druhého. $\square$

Zásobníkové překladové automaty

Věta
Nechť $PG = (N, T, D, R, S)$ je překladová gramatika. Pak existuje zásobníkový překladový automat ZPA takový, že $Z_\varepsilon(ZPA) = Z(PG)$.

Důkaz: Pro zadanou $PG = (N, T, D, R, S)$ sestrojíme $ZPA = (\{q\}, T, N \cup T \cup D, D, \delta, q, S, \emptyset)$, kde δ:

1. Expanze: $\delta(q, \varepsilon, A) = \{(q, \alpha, \varepsilon) : A \rightarrow \alpha \in R\}$,
2. Srovnání: $\delta(q, a, a) = \{(q, \varepsilon, \varepsilon)\}$ pro všechna $a \in T$.
3. Výstup: $\delta(q, \varepsilon, b) = \{(q, \varepsilon, b)\}$ pro všechna $b \in D$.

Důkaz, že $Z(PG) = Z_\varepsilon(ZPA)$, se provede indukcí podle délky derivace v gramatice PG a podle délky posloupnosti přechodů automatu ZPA. $\square$

Zásobníkové překladové automaty

Příklad

$PG = (\{E, T, F\}, \{+, *, (,), a\}, \{\oplus, \circledast, @\}, R, E)$, kde R:

$E \rightarrow E + T\oplus$ $\qquad$ $E \rightarrow T$

$T \rightarrow T * F\circledast$ $\qquad$ $T \rightarrow F$

$F \rightarrow (E)$ $\qquad$ $F \rightarrow a@$.

$ZPA = (\{q\}, \{+, *, (,), a\}, \{+, *, (,), a, E, T, F, \oplus, \circledast, @\},$
$\{\oplus, \circledast, @\}, \delta, q, E, \emptyset)$, kde δ:

$\delta(q, \varepsilon, E) = \{(q, E + T\oplus, \varepsilon), (q, T, \varepsilon)\},$

$\delta(q, \varepsilon, T) = \{(q, T * F\circledast, \varepsilon), (q, F, \varepsilon)\},$

$\delta(q, \varepsilon, F) = \{(q, (E), \varepsilon), (q, a@, \varepsilon)\},$

$\delta(q, c, c) = \{(q, \varepsilon, \varepsilon)\}$ pro všechna $c \in \{+, *, (,), a\},$

$\delta(q, \varepsilon, b) = \{(q, \varepsilon, b)\}$ pro všechna $b \in \{\oplus, \circledast, @\}.$

Zásobníkové překladové automaty

$$
\begin{array}{rlrrr}
(q, a+a*a, E, \varepsilon) & \vdash(q, & a+a*a, & E+T\oplus, & \varepsilon) \\
& \vdash(q, & a+a*a, & T+T\oplus, & \varepsilon) \\
& \vdash(q, & a+a*a, & F+T\oplus, & \varepsilon) \\
& \vdash(q, & a+a*a, & a@+T\oplus, & \varepsilon) \\
& \vdash(q, & +a*a, & @+T\oplus, & \varepsilon) \\
& \vdash(q, & +a*a, & +T\oplus, & @) \\
& \vdash(q, & a*a, & T\oplus, & @) \\
& \vdash(q, & a*a, & T*F\circledast\oplus, & @) \\
& \vdash(q, & a*a, & F*F\circledast\oplus, & @) \\
& \vdash(q, & a*a, & a@*F\circledast\oplus, & @) \\
& \vdash(q, & *a, & @*F\circledast\oplus, & @) \\
& \vdash(q, & *a, & *F\circledast\oplus, & @@) \\
& \vdash(q, & a, & F\circledast\oplus, & @@) \\
& \vdash(q, & a, & a@\circledast\oplus, & @@) \\
& \vdash(q, & \varepsilon, & @\circledast\oplus, & @@) \\
& \vdash(q, & \varepsilon, & \circledast\oplus, & @@@) \\
& \vdash(q, & \varepsilon, & \oplus, & @@@\circledast) \\
& \vdash(q, & \varepsilon, & \varepsilon, & @@@\circledast\oplus).
\end{array}
$$

Atributované překlady

- Atribut – veličina, která nabývá hodnot z nějaké množiny (obor hodnot atributu).
  Např. proměnná v programu, která má definován typ.
- Atributovaný symbol – symbol abecedy, ke kterému je přiřazena konečná množina atributů (i prázdná).
- Atributovaný řetězec – řetězec atributovaných symbolů.

Např.

$x.a$

$x[x.a_1, x.a_2, ..., x.a_n]$

Atributované překlady

Definice

Atributovaný překlad je relace $Z_A \subseteq T^* \times D^*$, kde T^* je množina vstupních atributovaných řetězců, D^* je množina výstupních atributovaných řetězců.

Příklad

Vstup: atributované řetězce nad abecedou $\{a, +, *, (,)\}$. Symbolu a je přiřazen atribut x (oborem hodnot je množina celých čísel).
Výstup: v, který má atribut y (oborem hodnot je množina celých čísel).

$(a[10] \qquad\qquad\qquad\quad , v[10]),$

$(a[5] + a[6] \qquad\quad , v[11]),$

$(a[3] * a[4] + a[2], v[14]).$

Atributované překlady

Příklad

$a.x$ bude relativní adresa, na které je uložena hodnota operandu a.

Adresa p, ke které jsou vztaženy relativní adresy operandů.

Funkce $\texttt{vyber}(x)$, která z adresy x vybere hodnotu.

$p = 100$, $\texttt{vyber}(102) = 3$, $\texttt{vyber}(103) = 5$, $\texttt{vyber}(104) = 6$

$(a[3] * a[4] + a[2], v[33])$

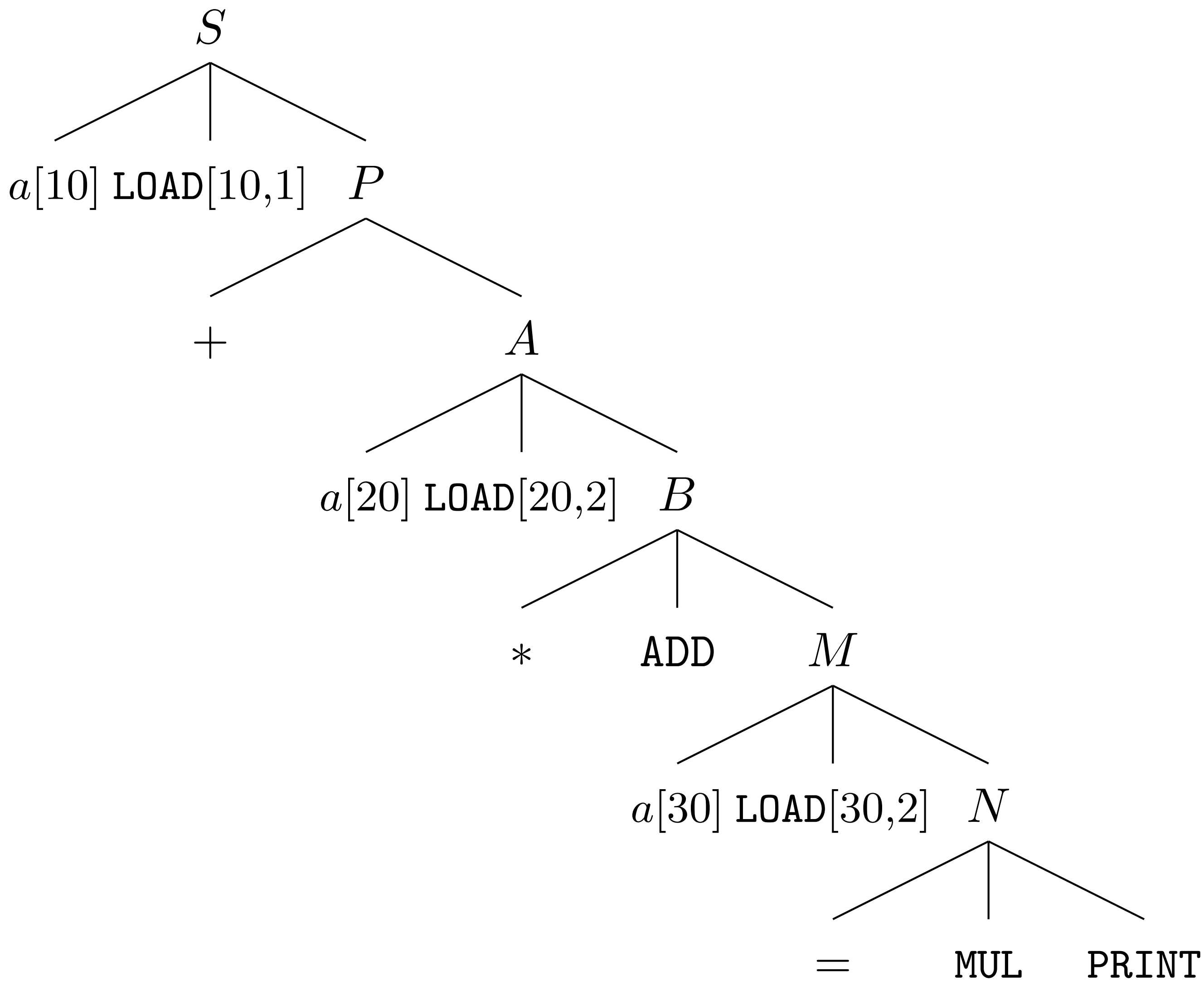
Atributované překlady

Hodnoty atributů výstupních symbolů mohou záviset:

- na hodnotách atributů vstupních symbolů,
- na struktuře vstupního řetězce,
- na zadaných parametrech.

Atributovaný překladový strom

Atributovaný překladový strom



Atributované překlady

Atribut:

- syntetizovaný — hodnota závisí na informacích obsažených uvnitř podstromu,
- dědičný — hodnota závisí na kontextu, ve kterém se příslušný podstrom nachází.

Atributovaný překladový strom

atributovaná výstupní věta:

- atributovaný řetězec výstupních symbolů, kterými jsou ohodnoceny listy atributovaného překladového stromu AT

sémantické vyhodnocení překladového stromu:

- výpočet hodnot atributů v překladovém stromu
  - stanovíme hodnoty dědičných atributů kořene AT
  - stanovíme hodnoty syntetizovaných atributů vstupních uzlů
  - v nějakém pořadí probíráme uzly stromu a vyhodnocujeme jejich atributy

Atributovaný překladový strom

Příklad

Realizujte překlad výrazu s operátory $+$, $*$, $=$, a operandy. Zadány adresy, na kterých jsou uloženy operandy. Výstupním jazykem tohoto překladu je strojový jazyk počítače, který má tyto instrukce:

`LOAD` adr, r	...	provede přesun obsahu adresy adr do registru r,
`ADD`	...	provede součet obsahu registrů 1 a 2 a výsledek uloží do registru 1,
`MUL`	...	provede součin obsahu registrů 1 a 2 a výsledek uloží do registru 1,
`PRINT`	...	provede výpis obsahu registru 1.

$(a[20] =$, `LOAD`[20,1] `PRINT`[] $)$

$(a[10] + a[30] =$, `LOAD`[10,1] `LOAD`[30,2] `ADD`[] `PRINT`[] $)$

$(a[10] + a[20] * a[30] =$, `LOAD`[10,1] `LOAD`[20,2] `ADD`[] `LOAD`[30,2] `MUL`[] `PRINT`[]$)$

Atributovaný překladový strom

Příklad (pokračování)

Základní gramatika atributové překladové gramatiky $PG = (\{S, P, A, B, M, N\}, \{a, +, *, =\}, \{\mathtt{LOAD}, \mathtt{ADD}, \mathtt{MUL}, \mathtt{PRINT}\}, R, S)$, kde R:

1. $S \to a\ \mathtt{LOAD}\ P$,
2. $P \to +\ A$,
3. $P \to *\ M$,
4. $P \to =\ \mathtt{PRINT}$,
5. $A \to a\ \mathtt{LOAD}\ B$,
6. $M \to a\ \mathtt{LOAD}\ N$,
7. $B \to +\ \mathtt{ADD}\ A$,
8. $B \to *\ \mathtt{ADD}\ M$,
9. $B \to =\ \mathtt{ADD}\ \mathtt{PRINT}$,
10. $N \to +\ \mathtt{MUL}\ A$,
11. $N \to *\ \mathtt{MUL}\ M$,
12. $N \to =\ \mathtt{MUL}\ \mathtt{PRINT}$.

Symboly	Dědičné atributy	Syntetizované atributy
a		adr
LOAD	adr, r	
ADD		
MUL		
PRINT		

Atributovaný překladový strom

Příklad (pokračování)

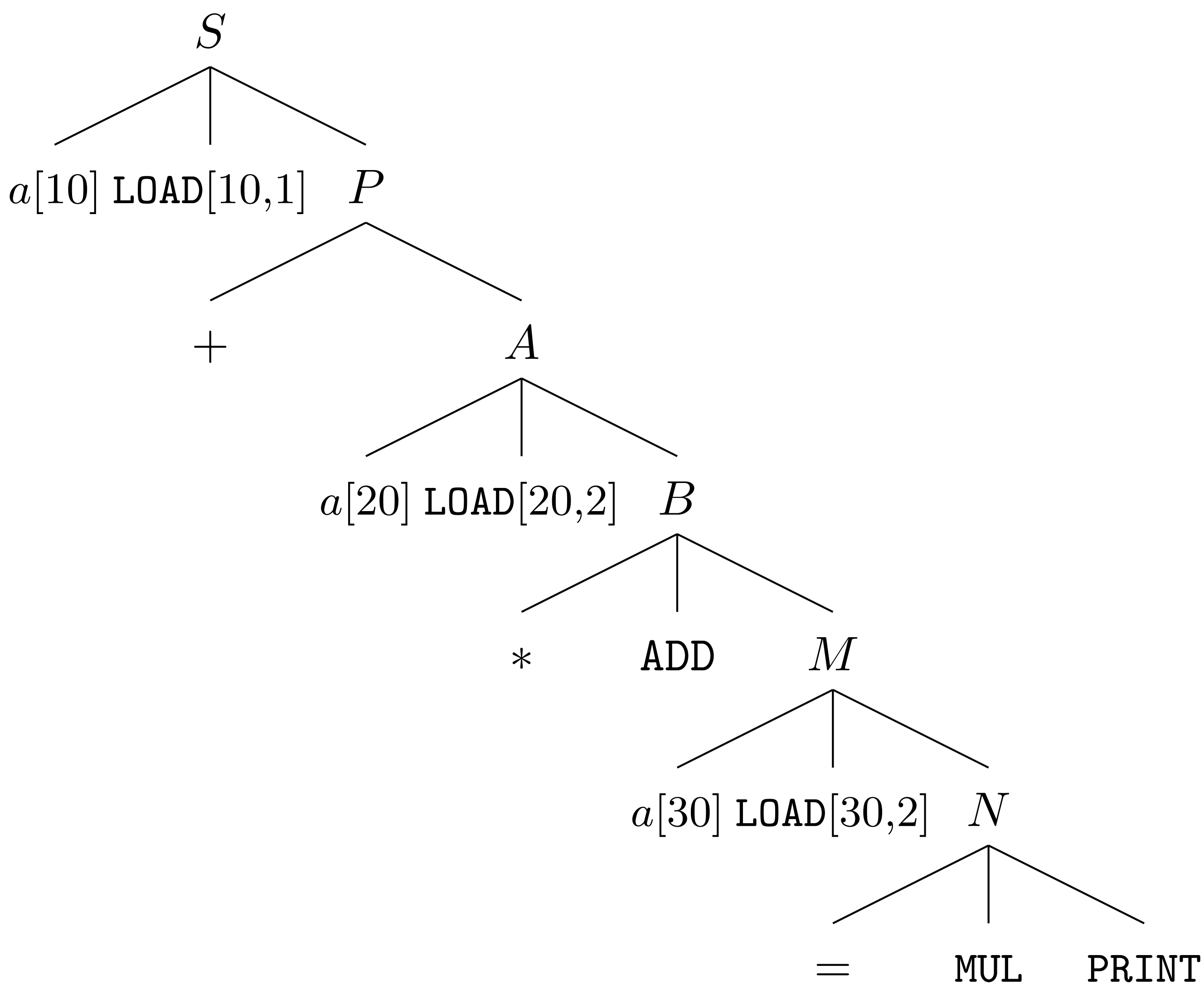
Sémantická pravidla a jejich přidělení pravidlům:

Syntax	Sémantika
1. $S \rightarrow a\ \mathtt{LOAD}\ P,$	$\mathtt{LOAD}.adr \leftarrow a.adr \quad \mathtt{LOAD}.r \leftarrow 1$
2. $P \rightarrow +\ A,$	
3. $P \rightarrow *\ M,$	
4. $P \rightarrow =\ \mathtt{PRINT},$	
5. $A \rightarrow a\ \mathtt{LOAD}\ B,$	$\mathtt{LOAD}.adr \leftarrow a.adr \quad \mathtt{LOAD}.r \leftarrow 2$
6. $M \rightarrow a\ \mathtt{LOAD}\ N,$	$\mathtt{LOAD}.adr \leftarrow a.adr \quad \mathtt{LOAD}.r \leftarrow 2$
7. $B \rightarrow +\ \mathtt{ADD}\ A,$	
8. $B \rightarrow *\ \mathtt{ADD}\ M,$	
9. $B \rightarrow =\ \mathtt{ADD}\ \mathtt{PRINT},$	
10. $N \rightarrow +\ \mathtt{MUL}\ A,$	
11. $N \rightarrow *\ \mathtt{MUL}\ M,$	
12. $N \rightarrow =\ \mathtt{MUL}\ \mathtt{PRINT}.$	

Atributovaný překladový strom

Příklad (pokračování)

Atributovaný překladový strom pro vstupní řetězec $a[10] + a[20] * a[30] =$.



Atributovaný překladový strom

Definice

Atributovaný překladový strom AT atributované vstupní věty w v $APG = (PG, A, V, F)$ je překladový strom této věty w sestrojený v PG (bez uvažování atributů) a rozšířený následujícím způsobem:

1. Ke každému uzlu, který je ohodnocen symbolem $X \in N \cup T \cup D$, jsou přidruženy atributy dané množinou $A(X)$.
2. Jsou stanoveny hodnoty dědičných atributů kořene stromu AT.
3. Hodnoty syntetizovaných atributů vstupních uzlů AT jsou dány vstupní větou w.
4. Nechť u_0 je libovolný vnitřní uzel X a $u_1, u_2, \ldots, u_n$ $(n \geq 0)$ jsou přímí následníci uzlu u_0 ohodnocení $X_1, X_2, \ldots, X_n$ a nechť $X_0 \to X_1 X_2 \ldots X_n$ je syntaktické pravidlo (r). Pak pro hodnoty atributů přidružených k uzlům u_k, $0 \leq k \leq n$, platí: je-li $t := f_{rtk}(a_1, a_2, \ldots, a_m)$ sémantické pravidlo pro výpočet hodnoty atributu t přidruženého k uzlu u_k, pak hodnota atributu t je určena tímto sémantickým pravidlem.

Atributová překladová gramatika

Definice

Atributová překladová gramatika $APG = (PG, A, V, F)$, kde:

- $PG = (N, T, D, R, S_o)$ – *základní překladová gramatika*, kde R:

  $$(r)\ X_0 \rightarrow X_1 X_2 \dots X_{n_r},$$

  kde $n_r \geq 0$, $X_0 \in N$, $X_k \in (N \cup T \cup D)$ pro $1 \leq k \leq n_r$.
- A – je konečná množina atributů ($A = S \cup I, S \cap I = \emptyset$, S je množina syntetizovaných atributů, I je množina dědičných atributů). Pro každý atribut a je zadán *obor hodnot* $H(a)$.
- V – zobrazení: každému $X \in N$ přiřazuje množinu $A(X) \in A$. Vstupní symboly mají syntetizované atributy, výstupní mají dědičné.
- F – konečná množina sémantických pravidel. $\forall X_k\ (1 \leq k \leq n_r)$ na pravé straně pravidla $r \in R$ a jeho dědičný atribut d:

  $$d := f_{rdk}(a_1, a_2, ..., a_m),$$ kde $a_1, a_2, \dots, a_m$ jsou atributy symbolů v pravidle r.

  $\forall$ syntetizovaný atribut s symbolu X_o na levé straně pravidla $r \in R$:

  $$s := f_{rso}(a_1, a_2, \dots, a_m),$$ kde $a_1, a_2, \dots, a_m$ jsou atributy symbolů v pravidle r.

Atributová překladová gramatika

Aby bylo možno určit hodnoty všech atributů, musíme předpokládat, že jsou zadány:

- hodnoty dědičných atributů počátečního symbolu,
- hodnoty syntetizovaných atributů vstupních symbolů.

Regulární atributované překlady

Definice
APG je *regulární* APG, jestliže platí:

1. Základní překladová gramatika je regulární.
2. Neterminální symboly mají jen dědičné atributy.

Regulární atributované překlady

Příklad

Překlad desítkových celých čísel: $RPG = (\{C\}, \{d\}, \{ⓥ\}, P, C)$, kde
P: $C \rightarrow dC | dⓥ$.

Symboly	Dědičné atributy	Syntetizované atributy
d		*kód*
C	$hodnota$	
ⓥ	$hodnota$	

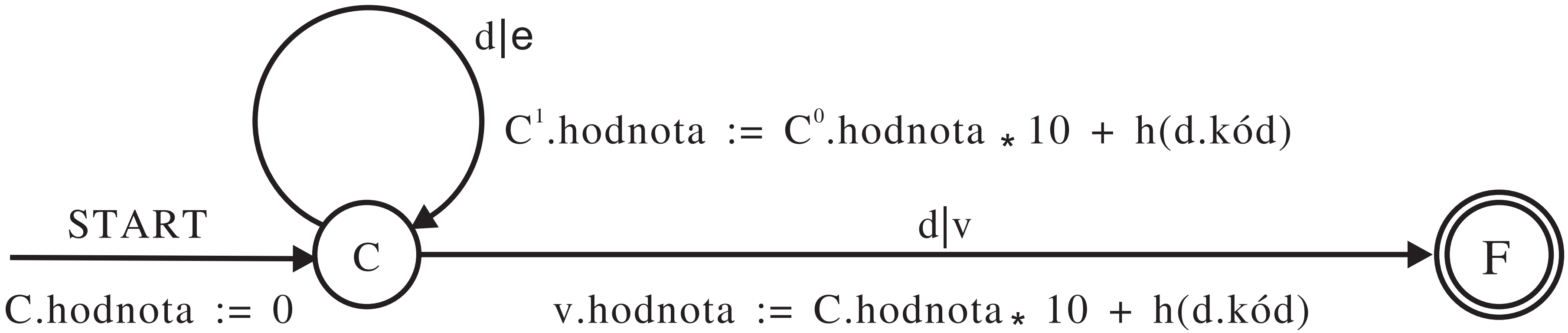
Syntax	Sémantika
$C^0 \rightarrow dC^1$	$C^1.hodnota := C^0.hodnota * 10 + h(d.$*kód*$)$
$C \rightarrow dⓥ$	$ⓥ.hodnota := C.hodnota * 10 + h(d.$*kód*$)$

Funkce $h(x)$ má jako argument kód číslice a jako funkční hodnotu číselnou hodnotu číslice. Počáteční hodnota atributu počátečního symbolu $C.hodnota$ se rovná nule.

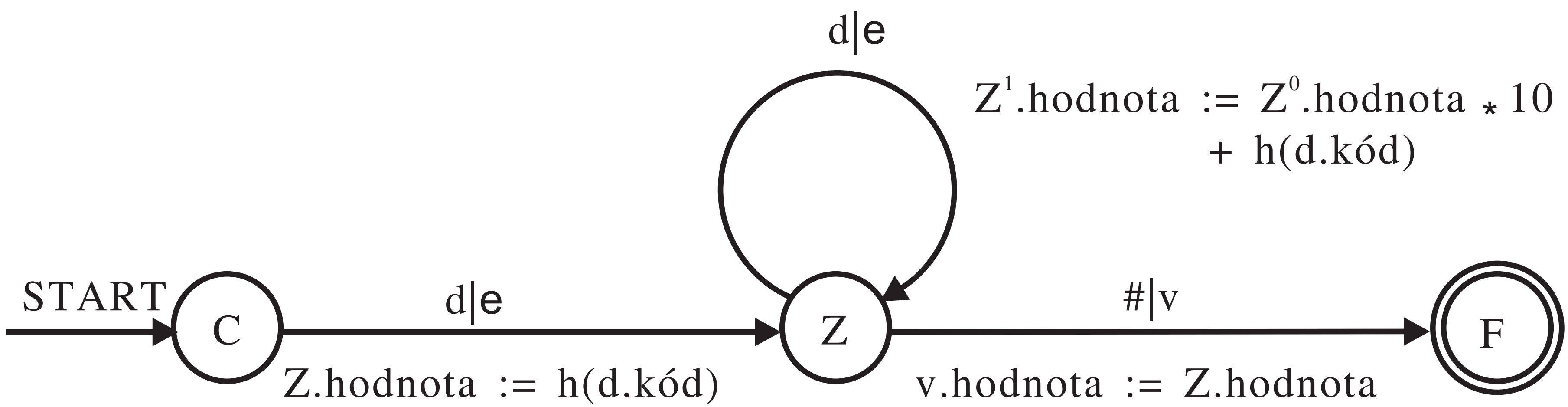
Regulární atributované překlady

Příklad (pokračování)

NKPA



DKPA



Regulární atributované překlady

Příklad (pokračování)

Upravená překladová gramatika

Syntax	Sémantika
$C \rightarrow dZ$	$Z.hodnota := h(d.\textit{kód})$
$Z^0 \rightarrow dZ^1$	$Z^1.hodnota := Z^0.hodnota * 10 + h(d.\textit{kód})$
$Z \rightarrow \#ⓥ$	$ⓥ.hodnota := Z.hodnota$

Regulární atributované překlady

Příklad

RAPG pro model kalkulačky s klávesami 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, $+$, $*$, $=$. Vstupem budou výrazy s operátory $+$ a $*$. Symbol $=$ bude konec výrazu. Operátory $+$ a $*$ mají stejnou prioritu.

dva registry = dva atributy x a y

$$RPG = (\{S, P, A, M, N\}, \{d, +, *, =\}, \{ⓥ\}, P, S).$$

Zápis $d.h$ představuje hodnotu vstupní číslice.

Regulární atributované překlady

Příklad (pokračování)

Syntax	Sémantika	
$S \rightarrow dP$	$P.x := d.h$	$P.y := 0$
$P^0 \rightarrow dP^1$	$P^1.x := P^0.x * 10 + d.h$	$P^1.y := P_0.y$
$P \rightarrow + A$	$A.x := P.x + P.y$	
$P \rightarrow * M$	$M.x := P.x + P.y$	
$P \rightarrow = ⓥ$	$ⓥ.x := P.x + P.y$	
$A \rightarrow dP$	$P.x := d.h$	$P.y := A.x$
$M \rightarrow dN$	$N.x := d.h$	$N.y := M.x$
$N^0 \rightarrow dN^1$	$N^1.x := N^0.x * 10 + d.h$	$N^1.y := N_0.y$
$N \rightarrow + A$	$A.x := N.x * N.y$	
$N \rightarrow * M$	$M.x := N.x * N.y$	
$N \rightarrow = ⓥ$	$ⓥ.x := N.x * N.y$	

Bezkontextové atributované překlady

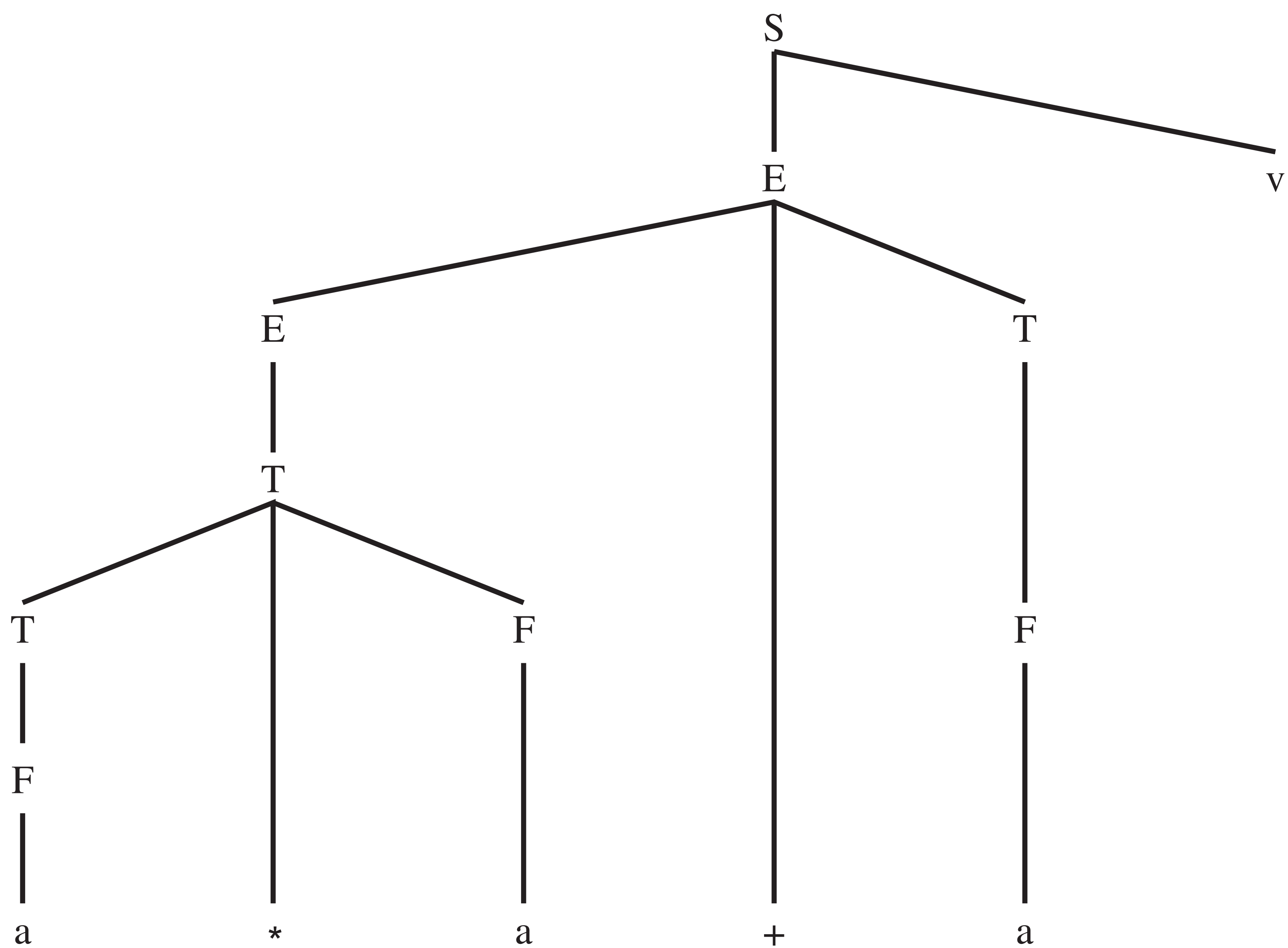
Gramatika respektující priority operací:

$PG = (\{S, E, T, F\}$, $\{a, +, *, (,)\}$, $\{v\}, R, S)$, kde R:

$$S \to Ev$$
$$E \to E + T \mid T$$
$$T \to T * F \mid F$$
$$F \to a \mid (E)$$

Bezkontextové atributované překlady

Překladový strom výrazu $(a * a + a)$



Bezkontextové atributované překlady

Atributy přidělené jednotlivým symbolům

Symboly	Atributy
S	p
E, T, F	h, p
a	x
v	y

Bezkontextové atributované překlady

Pravidla pro výpočet atributů

Pravidla překladové gramatiky	Pravidla pro výpočet atributů
$S \to Ev$	$E.p := S.p \quad v.y := E.h$
$E^0 \to E^1 + T$	$E^1.p := E^0.p \quad E^0.h := E^1.h + T.h$ $T.p := E^0.p$
$E \to T$	$T.p := E.p \quad E.h := T.h$
$T^0 \to T^1 * F$	$T^1.p := T^0.p \quad T^0.h := T^1.h * F.h$ $F.p := T^0.p$
$T \to F$	$F.p := T.p \quad T.h := F.h$
$F \to a$	$F.h := \texttt{vyber}(a.x + F.p)$
$F \to (E)$	$E.p := F.p \quad F.h := E.h$

Bezkontextové atributované překlady

Atributovaný překladový strom věty $a[3] * a[4] + a[2]$