Automata and Grammars (BIE-AAG)

12. Pushdown automata

Jan Holub

Department of Theoretical Computer Science

Faculty of Information Technology

Czech Technical University in Prague

European social fund.
Prague & EU: Investing in your future

Pushdown automaton

Definition

Pushdown automaton is a 7-tuple $R = (Q, T, G, \delta, q_0, Z_0, F)$, where:

- Q is a finite set of states,
- T is a finite input alphabet,
- G is a finite pushdown store alphabet,
- δ is a mapping from $Q \times (T \cup \{\varepsilon\}) \times G^*$ into set of finite subsets $Q \times G^*$,
- q_0 is the initial state,
- Z_0 is the initial pushdown store symbol,
- $F \subseteq Q$ is a set of final states.

Pushdown automaton

Configuration of PDA R: $(q, w, \alpha) \in Q \times T^* \times G^*$, where
q is the current state,
w is the yet unprocessed part of input string,
α is the pushdown store content.

Initial configuration of PDA R: $(q_0, w, Z_0), w \in T^*$

$\delta(q, a, \alpha) = \{(p_1, \gamma_1), (p_2, \gamma_2), ..., (p_m, \gamma_m)\}$: PDA in state q reads symbol a, goes into state p_i $(i = 1, 2, ..., m)$ and string α on top of the pushdown store is replaced by string γ_i.

$\delta(q, \varepsilon, \alpha)$ = $\{(p_1, \gamma_1), (p_2, \gamma_2), ..., (p_m, \gamma_m)\}$: transition into a new state and change of pushdown store content without reading an input symbol.

Pushdown automaton

Definition

Transition of a pushdown automaton R is a relation on the set of configurations:
$(q, aw, \alpha\beta) \vdash (p, w, \gamma\beta)$ if $(p, \gamma) \in \delta(q, a, \alpha)$,
$a \in T \cup \{\varepsilon\}, \alpha, \beta, \gamma, \in G^*$.

$\vdash^k$: k-th power of relation $\vdash$,
$\vdash^+$: transitive closure of relation $\vdash$,
$\vdash^*$: transitive and reflexive closure of relation $\vdash$

Pushdown automaton

Definition

Language defined (accepted) by PDA $R = (Q, T, G, \delta, q_0, Z_0, F)$:

1. *by transition into a final state*
   $L(R) = \{w : (q_0, w, Z_0) \vdash^* (q, \varepsilon, \gamma),\ \gamma \in G^*, q \in F, w \in T^*\}$,
2. by *empty pushdown store*
   $L_\varepsilon(R) = \{w : (q_0, w, Z_0) \vdash^* (q, \varepsilon, \varepsilon),\ q \in Q, w \in T^*\}$.

Pushdown automaton

Example

PDA that accepts language $L(\text{PDA}) = \{ww^R : w \in \{a,b\}^*\}$:
$R = (\{q,p\}, \{a,b\}, \{a,b,S,Z\}, \delta, q, Z, \{p\})$, where
$\delta(q, a, \varepsilon) = \{(q, a)\}$,
$\delta(q, b, \varepsilon) = \{(q, b)\}$,
$\delta(q, \varepsilon, \varepsilon) = \{(q, S)\}$,
$\delta(q, \varepsilon, aSa) = \{(q, S)\}$,
$\delta(q, \varepsilon, bSb) = \{(q, S)\}$,
$\delta(q, \varepsilon, SZ) = \{(p, \varepsilon)\}$.
For input string $aabbaa$ the automaton R performs this sequence of transitions:

$$\begin{array}{lll} (q, aabbaa, Z) & \vdash (q, abbaa, aZ) & \vdash (q, bbaa, aaZ) \\ \vdash (q, baa, baaZ) & \vdash (q, baa, SbaaZ) & \vdash (q, aa, bSbaaZ) \\ \vdash (q, aa, SaaZ) & \vdash (q, a, aSaaZ) & \vdash (q, a, SaZ) \\ \vdash (q, \varepsilon, aSaZ) & \vdash (q, \varepsilon, SZ) & \vdash (p, \varepsilon, \varepsilon) \end{array}$$

Basic properties of PDA

Theorem

Let $P = (Q, T, G, \delta, q_0, Z_0, F)$ be a PDA. If $(q, w, A) \vdash_P^n (q', \varepsilon, \varepsilon)$, then $(q, w, A\,\alpha) \vdash_P^n (q', \varepsilon, \alpha)$ for all $A \in G$ a $\alpha \in G^*$.

Basic properties of PDA

Theorem

$L = L(P_1)$ is language accepted by PDA P_1 by empty pushdown store iff $L = L(P_2)$ is language accepted by PDA P_2 by transition into a final state.

Proof: First we show that from $L = L(P_2) \Rightarrow L = L(P_1)$, i.e. $L(P_2) \subseteq L(P_1)$.
Let $P_2 = (Q, T, G, \delta, q_0, Z_0, F)$ be a PDA such that $L = L(P_2)$.
Let $P_1 = (Q \cup \{q_\varepsilon, q_0'\}, T, G \cup \{X\}, \delta', q_0', X, \emptyset)$, where δ' is defined thusly:

1. $\delta'(q_0', \varepsilon, X) = \{(q_0, Z_0X)\}$,
2. $\delta'(q, a, Z) = \delta(q, a, Z), \forall q, a, Z, q \in Q, a \in T \cup \{\varepsilon\}, Z \in G$,
3. $\delta'(q, \varepsilon, Z) = \{(q_\varepsilon, \varepsilon)\}, \forall q, Z, q \in F, Z \in G \cup \{X\}$,
4. $\delta'(q_\varepsilon, \varepsilon, Z) = \{(q_\varepsilon, \varepsilon)\}, \forall Z, Z \in G \cup \{X\}$.

Basic properties of PDA

Theorem
$L = L(P_1)$ is language accepted by PDA P_1 by empty pushdown store iff $L = L(P_2)$ is language accepted by PDA P_2 by transition into a final state.

Proof (cont.): We now show that from $L = L(P_1) \Rightarrow L = L(P_2)$.
Let $P_1 = (Q, T, G, \delta, q_0, Z_0, \emptyset)$ be a PDA such that $L = (P_1)$.
Let $P_2 = (Q \cup \{q_0', q_f\}, T, G \cup \{X\}, \delta, q_0', X, \{q_f\})$, where δ' is defined thusly:

1. $\delta'(q_0', \varepsilon, X) = \{(q_0, Z_0 X)\}$,
2. $\delta'(q, a, Z) = \delta(q, a, Z), \forall q, a, Z, q \in Q, a \in T \cup \{\varepsilon\}, Z \in G$,
3. $\delta'(q, \varepsilon, X) = \{q_f, \varepsilon)\}, \forall q, q \in Q$.

Relationship betw. CFG and PDA

Theorem
If a CFG $G = (N, T, P, S)$ is given, we can create a PDA R such that $L(G) = L(R)$.

A. Construction of PDA (model of top-down syntactic analysis):
$R = (\{q\}, T, N \cup T, \delta, q, S, \emptyset)$, where δ:

1. $\delta(q, \varepsilon, A) = \{(q, \alpha) : A \rightarrow \alpha \in P\}, \forall A, A \in N$ (**expansion**),
2. $\delta(q, a, a) = \{(q, \varepsilon)\}, \forall a, a \in T$ (**comparison**).

Top of the pushdown store for this type of automaton is always on the left.

Relationshiop betw. CFG and PDA

Example
CFG $G = (\{E, T, F\}, \{+, *, (,), a\}, P, E)$, where P:

$$\begin{array}{lll} (1)\ E \rightarrow E + T & (2)\ E \rightarrow T & (3)\ T \rightarrow T * F \\ (4)\ T \rightarrow F & (5)\ F \rightarrow (E) & (6)\ F \rightarrow a. \end{array}$$

PDA $R = (\{q\}, \{+, *, (,), a\}, \{+, *, (,), a, E, T, F\}, \delta, q, E, \emptyset)$,
where δ:

$$\begin{aligned} \delta(q, \varepsilon, E) &= \{(q, E + T), (q, T)\} \\ \delta(q, \varepsilon, T) &= \{(q, T * F), (q, F)\} \\ \delta(q, \varepsilon, F) &= \{(q, (E)), (q, a)\} \\ \delta(q, b, b) &= \{(q, \varepsilon)\}, \forall b, b \in \{a, +, *, (,)\}. \end{aligned}$$

Relationship betw. CFG and PDA

Example (cont.)
Sentence $a + a * a$ has the following left derivation in grammar G:

$$\begin{array}{lll} E & \Rightarrow E + T & (1) \\ & \Rightarrow T + T & (2) \\ & \Rightarrow F + T & (4) \\ & \Rightarrow a + T & (6) \\ & \Rightarrow a + T * F & (3) \\ & \Rightarrow a + F * F & (4) \\ & \Rightarrow a + a * F & (6) \\ & \Rightarrow a + a * a & (6) \end{array}$$

Relationship betw. CFG and PDA

Example (cont.)

$$
\begin{array}{lllrrl}
(q, a+a*a, E) & \vdash & (q, & a+a*a, & E+T) & (1) \\
 & \vdash & (q, & a+a*a, & T+T) & (2) \\
 & \vdash & (q, & a+a*a, & F+T) & (4) \\
 & \vdash & (q, & a+a*a, & a+T) & (6) \\
 & \vdash & (q, & +a*a, & +T) & \\
 & \vdash & (q, & a*a, & T) & \\
 & \vdash & (q, & a*a, & T*F) & (3) \\
 & \vdash & (q, & a*a, & F*F) & (4) \\
 & \vdash & (q, & a*a, & a*F) & (6) \\
 & \vdash & (q, & *a, & *F) & \\
 & \vdash & (q, & a, & F) & \\
 & \vdash & (q, & a, & a) & (6) \\
 & \vdash & (q, & \varepsilon, & \varepsilon) &
\end{array}
$$

Relationship betw. CFG and PDA

Example (cont.)

Rules applied in the left derivation of sentence $a + a * a$: 1, 2, 4, 6, 3, 4, 6, 6.

Relationship betw. CFG and PDA

B. Construction of PDA (bottom-up syntactic analysis):
$R = (\{q, r\}, T, N \cup T \cup \{\#\}, \delta, q, \#, \{r\})$, where δ:

1. $\delta(q, a, \varepsilon) = \{(q, a)\}, \forall a, a \in T$, **(shift)**,
2. $\delta(q, \varepsilon, \alpha) = \{(q, A) : A \to \alpha \in P\}$ **(reduce)**,
3. $\delta(q, \varepsilon, \#S) = \{(r, \varepsilon)\}$ **(přijetí)**.

Top of the pushdown store for this type of automaton is always on the right.

Relationship betw. CFG and PDA

Example

Let us have CFG $G = (\{E, T, F\}, \{+, *, (,), a\}, P, E)$, where P:

$(1)\ E \to E + T$	$(2)\ E \to T$	$(3)\ T \to T * F$
$(4)\ T \to F$	$(5)\ F \to (E)$	$(6)\ F \to a.$

$R = (\{q, r\}, \{+, *, (,), a\}, \{E, T, F, +, *, (,), a, \#\}, \delta, q, \#, \{r\})$,
where:

$\delta(q, b, \varepsilon) = \{(q, b)\}, \forall b, b \in \{a, +, *, (,)\}$
$\delta(q, \varepsilon, E + T) = \{(q, E)\}$
$\delta(q, \varepsilon, T) = \{(q, E)\}$
$\delta(q, \varepsilon, T * F) = \{(q, T)\}$
$\delta(q, \varepsilon, F) = \{(q, T)\}$
$\delta(q, \varepsilon, (E)) = \{(q, F)\}$
$\delta(q, \varepsilon, a) = \{(q, F)\}$
$\delta(q, \varepsilon, \#E) = \{(r, \varepsilon)\}$.

Relationship betw. CFG and PDA

Example (cont)

Sentence $a + a * a$ has the following right derivation in grammar G:

$$
\begin{array}{lll}
E & \Rightarrow E + T & (1) \\
 & \Rightarrow E + T * F & (3) \\
 & \Rightarrow E + T * a & (6) \\
 & \Rightarrow E + F * a & (4) \\
 & \Rightarrow E + a * a & (6) \\
 & \Rightarrow T + a * a & (2) \\
 & \Rightarrow F + a * a & (4) \\
 & \Rightarrow a + a * a & (6)
\end{array}
$$

Relationship betw. CFG and PDA

Example (cont.)

$$
\begin{array}{llrll}
(q, a + a * a, \#) & \vdash (q, & +a * a, & \#a &) \\
 & \vdash (q, & +a * a, & \#F &) \quad (6) \\
 & \vdash (q, & +a * a, & \#T &) \quad (4) \\
 & \vdash (q, & +a * a, & \#E &) \quad (2) \\
 & \vdash (q, & a * a, & \#E + &) \\
 & \vdash (q, & *a, & \#E + a &) \\
 & \vdash (q, & *a, & \#E + F &) \quad (6) \\
 & \vdash (q, & *a, & \#E + T &) \quad (4) \\
 & \vdash (q, & a, & \#E + T* &) \\
 & \vdash (q, & \varepsilon, & \#E + T * a &) \\
 & \vdash (q, & \varepsilon, & \#E + T * F &) \quad (6) \\
 & \vdash (q, & \varepsilon, & \#E + T &) \quad (3) \\
 & \vdash (q, & \varepsilon, & \#E &) \quad (1) \\
 & \vdash (r, & \varepsilon, & \varepsilon &)
\end{array}
$$

Relationship betw. CFG and PDA

Example (cont.)

Rules applied in the left derivation of sentence $a + a * a$: 6, 4, 2, 6, 4, 6, 3, 1.

Deterministic PDA

Definition

Pushdown automaton $R = (Q, T, G, \delta, q_0, Z_0, F)$ is *deterministic*, if the following holds:

1. $|\delta(q, a, \gamma)| \leq 1, \forall q, a, \gamma, q \in Q, a \in (T \cup \{\varepsilon\}), \gamma \in G^*$.
2. If $\delta(q, a, \alpha) \neq \emptyset$, $\delta(q, a, \beta) \neq \emptyset$ and $\alpha \neq \beta$, then α is not a suffix of β and β is not a suffix of α (i.e. $\gamma\alpha \neq \beta, \alpha \neq \gamma\beta$).
3. If $\delta(q, a, \alpha) \neq \emptyset$, $\delta(q, \varepsilon, \beta) \neq \emptyset$, then α is not a suffix of β and β is not a suffix of α (i.e. $\gamma\alpha \neq \beta, \alpha \neq \gamma\beta$).

Deterministic PDA

Construction of det. PDA by the top-down method (**A**) for CFG in Greibach normal form:

(Greibach normal form: all rules are in the form $A \to a\alpha$, where $a \in T, \alpha \in N^*$)

$R = (\{q\}, T, N, \delta, q, S, \emptyset)$, where
$\delta(q, a, A) = \{(q, \alpha) : A \to a\alpha \in P\}, \forall A, A \in N$.

Deterministic PDA

Construction of determinitic PDA by a bottom-up method (**B**):

Example

CFG $G = (\{S, A, B\}, \{a, b, c, d\}, P, S)$, where P:

$S \to Aa \qquad A \to Bb \mid c \qquad B \to d$

$R = (\{q, r\}, \{a, b, c, d\}, \{S, A, B, a, b, c, d, \#\}, \delta, q, \#, \{r\})$, where
δ:

$$
\begin{array}{ll}
(1) & \delta(q, a, \varepsilon) = (q, a) \\
 & \delta(q, b, \varepsilon) = (q, b) \\
 & \delta(q, c, \varepsilon) = (q, c) \\
 & \delta(q, d, \varepsilon) = (q, d) \\
(2) & \delta(q, \varepsilon, Aa) = (q, S) \\
 & \delta(q, \varepsilon, Bb) = (q, A) \\
 & \delta(q, \varepsilon, c) = (q, A) \\
 & \delta(q, \varepsilon, d) = (q, B) \\
(3) & \delta(q, \varepsilon, \#S) = (r, \varepsilon)
\end{array}
$$

Relationship betw. CFG and PDA

Example (cont.)

PDA is nondeterministic due to shifts by (1). These shifts can be made depending on the contents of the pushdown store:

$(1)'$ $\delta(q, a, A) = \{(q, Aa)\}$ – symbol a is present in the sentential form only after symbol A,

$\delta(q, b, B) = \{(q, Bb)\}$ – symbol b is present in the sentential form only after symbol B,

$\delta(q, c, \#) = \{(q, \#c)\}$,

$\delta(q, d, \#) = \{(q, \#d)\}$ – symbols c, d can be present only at the beginning of the sentential form.