Теория автоматов и формальных языков

Атрибутные грамматики и магазинные преобразователи

Автор: Екатерина Вербицкая

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»

29 ноября 2016г.

В предыдущей серии

- Полезно не только распознавать предложения или строить деревья их разбора, но и осуществлять трансляцию произвольного вида
- Трансляция — перевод предложения на одном языке в предложение на другом языке
- Для этого существует несколько механизмов
  - S-атрибутные грамматики
    - Все атрибуты синтезируемые (атрибуты узла и его детей)
  - L-атрибутные грамматики
    - Все атрибуты наследуемые (атрибуты узлов предков или братьев слева)
  - Схема синтаксически управляемой трансляции
- Есть ли общий механизм работы с трансляциями?

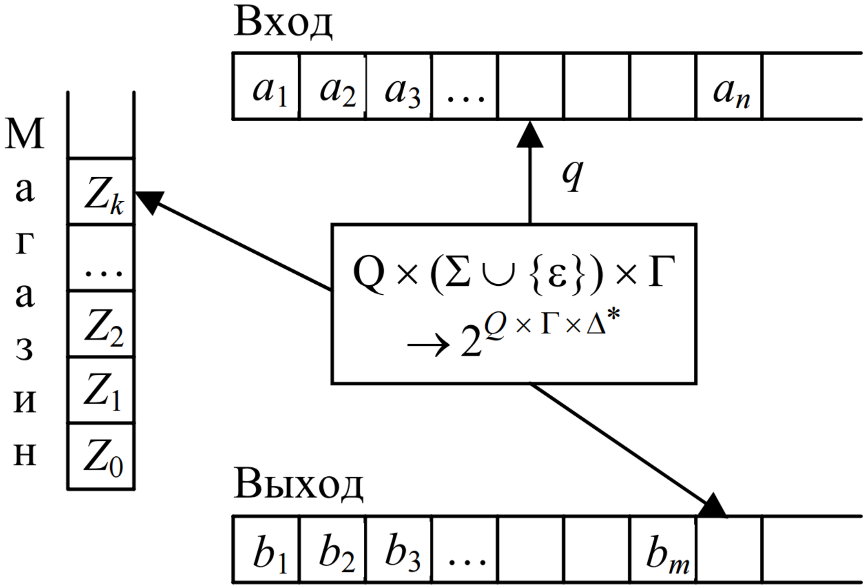
В предыдущей серии: простые СУ-схемы

Простая схема синтаксически управляемой трансляции — пятерка (N, Σ, Π, P, S)

- N — конечное множество нетерминальных символов
- Σ — конечный входной алфавит
- Π — конечный выходной алфавит
- $S \in N$ — стартовый нетерминал
- P — конечное множество правил трансляции вида $A \to \alpha, \beta$, где $\alpha \in (N \cup \Sigma)^*, \beta \in (N \cup \Pi)^*$
  - Нетерминалы входят в цепочку β в том же порядке, в каком они входят в α
  - Если нетерминалы повторяются больше одного раза, то их различают по индексам: $E \to E^l + E^r, +E^l\,E^r$

Такие схемы можно моделировать **магазинным преобразователем**

Что такое магазинный преобразователь



Что такое магазинный преобразователь: неформально

- Магазинный автомат, который при каждом переходе пишет что-то в выходную строку

Формальное определение

Магазинный преобразователь это набор $(Q, \Sigma, \Gamma, \Delta, \delta, q_0, Z_0, F)$

- Q — конечное множество состояний
- Σ — конечное множество символов, входной алфавит
- Γ — конечное множество символов, стековый алфавит
- Δ — конечное множество символов, выходной алфавит
- $\delta \subseteq Q \times (Z \cup \varepsilon) \times \Gamma \to 2^{Q \times \Gamma^* \times \Delta^*}$ — отношение переходов
- $q_0 \in Q$ — стартовое состояние
- $Z_0 \in \Gamma$ — начальный элемент стека
- $F \subseteq Q$ — множество принимающих (конечных) состояний

Отношение переходов

$\delta(p, a, Z) = \{(q_i, \gamma_i, \alpha_i) \,|\, 1 \leq i \leq n\}$ означает

- Если магазинный преобразователь находится в состоянии $p \in Q$, на вершине стека находится $Z \in \Gamma$, а со входа читается символ $a \in \Sigma \cup \varepsilon$, то для некоторого i:
  - Изменяем состояние на $q_i \in Q$
  - Снимаем со стека символ Z, записываем на стек строку $\gamma_i \in \Gamma^*$
  - В выходную строку дописываем $\alpha_i \in \Delta^*$
- $\Sigma \cup \varepsilon$ сигнализирует о том, что вход можно и не читать
- Если $\gamma_i = \varepsilon$, символ со стека стирается
- Если $\alpha_i = \varepsilon$, в выходную строку ничего не пишем

Семантика магазинного преобразователя

- Мгновенное описание МП: $(p, \omega, \beta, \alpha) \in Q \times \Sigma^* \times \Gamma^* \times \Delta^*$
  - p — текущее состояние автомата
  - ω — непрочитанный фрагмент входного потока
  - β — содержимое стека (верхушка записана первой)
  - α — содержимое выходной ленты
- Отношение $\vdash$ на мгновенных описаниях (шаг)
  - Для каждого $(q, \gamma, \alpha) \in \delta(p, a, Z)$, верно $(p, ax, Z\eta, \zeta) \vdash (q, x, \gamma\eta, \alpha\zeta)$ для произвольных $x \in \Sigma^*, \eta \in \Gamma^*, \zeta \in \Delta^*$
- Шаг не определен, если стек пуст

Семантика магазинного преобразователя: вычисление

- Вычисление — последовательность шагов
  - $\vdash^*$ — транизитивно рефлексивное замыкание отношения $\vdash$
- Начальное мгновенное описание $(q_0, \omega, Z_0, \varepsilon)$
- Два варианта окончания работы
  - По достижении конечного состояния
    - $\tau(M) = \{(\omega, \alpha) \,|\, \omega \in \Sigma^*, \alpha \in \Delta^*, (q_0, \omega, Z_0, \varepsilon) \vdash^* (f, \varepsilon, \gamma, \alpha), f \in F, \gamma \in \Gamma^*\}$
  - По опустошении стека
    - $\tau_\varepsilon(M) = \{(\omega, \alpha) \,|\, \omega \in \Sigma^*, \alpha \in \Delta^*, (q_0, \omega, Z_0, \varepsilon) \vdash^* (q, \varepsilon, \varepsilon, \alpha), q \in Q\}$
  - Эти варианты эквивалентны: по преобразователю, завершающемуся по первой схеме, можно посмотроить преобразователь, завершающийся по второй схеме, и наоборот

Детерминированные магазинные преобразователи

Магазинный преобразователь является **детерминированным**, если

- $\forall q \in Q, a \in \Sigma \cup \{\varepsilon\}, Z \in \Gamma . |\delta(q, a, Z)| \leq 1$
- Если $\delta(q, \varepsilon, Z) \neq \varnothing$, то $\forall a \in \Sigma . \delta(q, a, Z) = \varnothing$
- Детерминированный магазинный преобразователь является частным случаем недетерминированного

Пример: преобразование префиксных арифметических выражений в постфиксные

$M = \{\{q\}, \{a, +, *\}, \{E, +, *\}, \{a, +, *\}, \delta, q, E, \{q\}\}$

$$\begin{array}{rcl}
\delta(q, a, E) & = & \{(q, \varepsilon, a)\} \\
\delta(q, +, E) & = & \{(q, EE+, \varepsilon)\} \\
\delta(q, *, E) & = & \{(q, EE*, \varepsilon)\} \\
\delta(q, \varepsilon, +) & = & \{(q, \varepsilon, +)\} \\
\delta(q, \varepsilon, *) & = & \{(q, \varepsilon, *)\}
\end{array}$$

$(q, + * aaa, E, \varepsilon) \vdash (q, *aaa, EE+, \varepsilon) \vdash (q, aaa, EE * E+, \varepsilon) \vdash$
$(q, aa, E * E+, a) \vdash (q, a, *E+, aa) \vdash (q, a, E+, aa*) \vdash (q, \varepsilon, +, aa * a) \vdash$
$(q, \varepsilon, \varepsilon, aa * a+)$

Взаимоотношение между простыми СУ-схемами и магазинными преобразователями

Теорема

По простой СУ-схеме $(N, \Sigma, \Delta, R, S)$ *можно построить магазинный преобразователь, задающий эквивалентную трансляцию*

Теорема

По магазинному преобразователю $P = (Q, \Sigma, \Gamma, \Delta, \delta, q_0, Z_0, \varnothing)$ *можно построить простую СУ-схему, задающую эквивалентную трансляцию*

Теорема

Класс трансляций, задаваемых простыми СУ-трансляциями совпадает с классом трансляций, задаваемых магазинными автоматами

Однозначные СУ-схемы и левосторонний вывод

Однозначная СУ-схема — СУ-схема, в которой не существует двух правил $A \to \alpha, \beta, A \to \alpha, \gamma : \beta \neq \gamma$

Теорема

Выходная цепочка однозначной СУ-схемы может быть сгенерирована при левостороннем выводе входной цепочки