LR(k) Parser

Consider a rightmost derivation in normal and reversed order.

$$S \Rightarrow^*_{rm} \delta A z \Rightarrow_{rm} \delta\alpha\beta z \Rightarrow^*_{rm} \delta\alpha w z \Rightarrow^*_{rm} \delta v w z \Rightarrow^*_{rm} uvwz\ .$$

$$uvyz \Leftarrow^*_{rm} \delta v y z \Leftarrow^*_{rm} \delta\alpha y z \Leftarrow^*_{rm} \delta\alpha\beta z \Leftarrow_{rm} \delta A z \Leftarrow^*_{rm} S\ .$$

$$\bullet uvyz \Leftarrow^*_{rm} \delta\bullet vyz \Leftarrow^*_{rm} \delta\alpha\bullet yz \Leftarrow^*_{rm} \delta\alpha\beta\bullet z \Leftarrow_{rm} \delta A\bullet z \Leftarrow^*_{rm} S\bullet.$$

$$\$ \mid uvyz\$ \Rightarrow^*_R \$\delta \mid vyz\$ \Rightarrow^*_R \$\delta\alpha \mid yz\$ \Rightarrow^*_R \$\delta\alpha\beta \mid z\$ \Rightarrow^{\alpha\beta\to A}_R \$\delta A \mid z \Rightarrow^*_R \$S \mid \$$$

We ***define*** *$[A \rightarrow \alpha\bullet\beta, k{:}z]$ to be a valid LR(k) item for a viable* ***prefix*** *$\delta\alpha$.*

Let R_k denote a set of valid LR(k) items.

We define $Valid_k$: $(N\cup\Sigma)^* \rightarrow 2^{R_k}$ or $\langle\cdot\rangle_k$ for short.

$$Valid_k(\gamma) = \langle\gamma\rangle_k$$

$$= \{[A \rightarrow \alpha\bullet\beta, x] \in R_k \mid S \Rightarrow^*_{rm} \delta A z \Rightarrow_{rm} \delta\alpha\beta z, \gamma = \delta\alpha, x = k{:}z\}$$

Lemma *If* $[A \rightarrow \alpha \bullet \beta\gamma, x] \in \langle\delta\alpha\rangle_k$, *then* $[A \rightarrow \alpha\beta \bullet \gamma, x] \in \langle\delta\alpha\beta\rangle_k$.

Proof

$\exists S \Rightarrow^*_{rm} \delta A z \Rightarrow_{rm} \delta\alpha\beta\gamma z$ *and* $x = k{:}z$.

$\therefore [A \rightarrow \alpha\beta \bullet \gamma, x] \in \langle\delta\alpha\beta\rangle_k$.

Cololary *If* $[A \rightarrow \alpha \bullet X\beta, x] \in \langle\delta\alpha\rangle_k$, $[A \rightarrow \alpha X \bullet \beta, x] \in \langle\delta\alpha X\rangle_k$.

Theorem *If* $[B \rightarrow \eta \bullet A\psi, x] \in \langle\delta\eta\rangle_k$ *and* $A \rightarrow \alpha \in P$, *then*

$[A \rightarrow \bullet\alpha, y] \in \langle\delta\eta\rangle_k$ *where* $y \in First_k(\psi x)$ *and vice versa.*

Proof $\exists S \Rightarrow^*_{rm} \delta B u \Rightarrow_{rm} \delta\eta A \underline{\psi} u \Rightarrow^*_{rm} \delta\eta A \underline{v} u \Rightarrow_{rm} \delta\eta\alpha v u = \gamma\alpha v u$,

$x = k{:}u$. $\psi \Rightarrow^* v$, $y = k{:}vu = k{:}vx \in First_k(\psi x)$.

Let $K \subseteq R_k$.

Definition $\partial_{LR(k)}: 2^{R_k} \rightarrow 2^{R_k}$, ∂_k *or* ∂ *for short.*

(***desc***$_{LR(k)}$, ***desc***$_k$ *in text*).

$$\partial_k K = \{[B \rightarrow \bullet\eta, y] \in R_k \mid [A \rightarrow \alpha\bullet B\beta, x] \in K, B \rightarrow \eta \in P, y \in First_k(\beta\cdot x)\}$$

Definition $\chi_{LR(k)}{}^X: 2^{R_k} \times (N \cup \Sigma) \rightarrow 2^{R_k}$, $\chi_k{}^X$ *or* χ^X *for short*

(***passes-X***$_{LR(k)}$, ***passes-X*** *in text*).

$$\chi_k{}^X K = \{[A \rightarrow \alpha X\bullet\beta, x] \in R_k \mid [A \rightarrow \alpha\bullet X\beta, x] \in K\}$$

If $K \in \langle\delta\rangle_k$, $\partial_k{}^* K \in \langle\delta\rangle_k$.

If $K \in \langle\delta\rangle_k$, $\chi_k{}^X K \in \langle\delta\cdot X\rangle_k$.

We define $\rho_{LR(k)} \subseteq (N \cup \Sigma)^* \times (N \cup \Sigma)^*$ ρ_k *or* ρ *for short.*

$\gamma\ \rho_k\ \delta$, *if* $\langle\gamma\rangle_k = \langle\delta\rangle_k$.

ρ_k *is an **equivalent** binary relation on* $(N \cup \Sigma)^*$.

$[\gamma]_k = \{\delta \in (N \cup \Sigma)^* \mid \gamma\ \rho_k\ \delta\}$ *equivalent class on* $(N \cup \Sigma)^*$.

We extend the domain of $\langle\cdot\rangle_k$ *from* $(N \cup \Sigma)^*$ *to* $2^{(N \cup \Sigma)^*}$.

We may write $\langle\delta\rangle_k$ ***instead*** *of* $\langle[\delta]_k\rangle_k$. *since* $\langle\delta\rangle_k = \langle[\delta]_k\rangle_k$.

We may write $\langle\delta\rangle_k$ ***instead*** *of* $[\delta]_k$, *or vice versa, since* $\langle\delta\rangle_k \leftrightarrow^{1:1} [\delta]_k$.

The following statements are equivalent:

An LR(k) ***state***	$\langle\delta\rangle_k \leftrightarrow^{1:1} [\delta]_k$.
A set of valid ***LR(k) items***	$\langle\delta\rangle_k$.
A set of valid ***viable prefices***	$[\delta]_k$.
A set of valid ***stack strings***	$[\delta]_k$.

Canonical Collection of <u>*LR(k) states*</u>*,* C_k*, and* $Q: C_k \times (N \cup \Sigma) \rightarrow C_k$.

(<u>*sets of LR(k) items*</u>, <u>*equivalent classes of valid viable prefices*</u>).

$\langle\varepsilon\rangle_k := \partial_k^{\ *}([S' \rightarrow \bullet S, \varepsilon]);\ C_k := \{\langle\varepsilon\rangle_k\};\ Q := \emptyset;$

repaeat

  for $\langle\delta\rangle_k \in C_k$ ***do***

    for $X \in N \cup \Sigma$ ***where*** $K^X = \{[A \rightarrow \alpha \bullet X\beta, x] \in \langle\delta\rangle_k\}$ ***do***

      $\langle\delta \cdot X\rangle_k := \partial_k^{\ *}(\chi_k^{\ X}(K^X));$

      $C_k := C_k \cup \langle\delta \cdot X\rangle_k;$

      $Q := Q \cup \{\langle\delta\rangle_k \cdot X \rightarrow \langle\delta \cdot X\rangle_k\}$

    od ***od***

until *no more states are added to* C_k.

Fact & Construction of *shift actions of* Γ *in LR(k) parsing.*

If $[B \to \alpha \bullet \mathbf{X} \beta, x] \in \langle\delta\rangle_k \in C_k$ *and* $\mathbf{X} \in N \cup \Sigma$, *then*

$^{\exists}[B \to \alpha \mathbf{X} \bullet \beta, x] \in \langle\delta \cdot \mathbf{X}\rangle_k \in C_k$ *and*

$\textbf{Add } \langle\delta\rangle_k \mid \mathbf{X} \cdot z \to \langle\delta\rangle_k \langle\delta \cdot \mathbf{X}\rangle_k \mid z \in \Gamma,\ z \in First_{k-1}(\beta \cdot z).$

Fact & Construction of *reduce actions of* Γ *in LR(k) parsing.*

If $[B \to \alpha \bullet \mathbf{A} \beta, x] \in \langle\delta\rangle_k \in C_k$ *and* $\mathbf{A} \to X_1 \cdot X_2 \ldots \cdot X_n \in P$, *then*

$^{\exists}[\mathbf{A} \to \bullet X_1 \cdot X_2 \cdot \ldots \cdot X_n, y] \in \langle\delta\rangle_k$ *where* $y \in First_k(\beta \cdot x)$,

$[1 \le {}^{\forall} i \le n:\ {}^{\exists}[A \to X_1 X_2 \ldots X_i \bullet \ldots X_n, y] \in \langle\delta X_1 X_2 \ldots X_i\rangle_k$ *and*

$^{\exists}\langle\delta \cdot X_1 \cdot \ldots \cdot X_{i-1}\rangle_k \mid X_i \cdot z \to \langle\delta \cdot X_1 \cdot \ldots \cdot X_{i-1}\rangle_k \langle\delta \cdot X_1 \cdot \ldots \cdot X_{i-1} \cdot X_i\rangle_k \mid z \in \Gamma,$

where $z \in First_k(X_{i+1} \ldots X_n \cdot y)]$

$^{\exists}[B \to \alpha \mathbf{A} \bullet \beta, x] \in \langle\delta \cdot \mathbf{A}\rangle_k \in C_k$ *and* $^{\exists}\langle\delta\rangle_k \mid y \to \langle\delta\rangle_k \langle\delta \cdot \mathbf{A}\rangle_k \mid y \in \Gamma.$

$\textbf{Add } \langle\delta\rangle_k \langle\delta \cdot X_1\rangle_k \langle\delta \cdot X_1 \cdot X_2\rangle_k \ldots \langle\delta \cdot X_1 \cdot X_2 \ldots \cdot X_n\rangle_k \mid \mathbf{y} \to \langle\delta\rangle_k \langle\delta \cdot \mathbf{A}\rangle_k \mid \mathbf{y} \in \Gamma.$

LR(k) parser is a pdt $M_k = \{C_k, \Sigma, \Gamma, P, \tau, [\varepsilon]_k, \{[\varepsilon]_k[S]_k\}, \$, \mid)$ *where*

$\Gamma = \{[\delta]_k[\delta\cdot X_1]_k[\delta\cdot X_1\cdot X_2]\ldots[\delta\cdot X_1\cdot X_2\cdot\ldots\cdot X_n]_k \mid \mathbf{x} \to [\delta]_k[\delta\cdot \mathbf{A}]_k \mid \mathbf{x} \mid$

$[\mathbf{A} \to X_1\cdot X_2\cdot\ldots\cdot X_n\bullet, \mathbf{x}] \in \langle\delta\cdot X_1\cdot X_2\cdot\ldots\cdot X_n\rangle_k\}$

$\cup \{[\delta]_k \mid \mathbf{a}x \to [\delta]_k[\delta\cdot\mathbf{a}]_k \mid x \mid [A \to \alpha\bullet\mathbf{a}\beta, y] \in \langle\delta\rangle_k, x \in First_{k-1}(\beta y)\}$

$\tau([\delta]_k[\delta\cdot X_1]_k\ldots[\delta\cdot X_1\cdot X_2\cdot\ldots\cdot X_n]_k \mid \mathbf{x} \to [\delta]_k[\delta\cdot\mathbf{A}]_k \mid \mathbf{x}) = A \to X_1\cdot\ldots\cdot X_n \in P,$

$\tau([\delta]_k \mid \mathbf{a}x \to [\delta]_k[\delta\cdot\mathbf{a}]_k \mid x) = \varepsilon.$

$[B \to \alpha\bullet\mathbf{A}\beta, y] \in \langle\delta\rangle_k$ $\qquad$ $[B \to \alpha\mathbf{A}\bullet\beta, y] \in \langle\delta\cdot\mathbf{A}\rangle_k$

$[A \to \bullet X_1\cdot X_2\cdot\ldots\cdot X_n, x] \in \langle\delta\rangle_k$

$[A \to X_1\bullet\cdot X_2\cdot\ldots\cdot X_n, x] \in \langle\delta\cdot X_1\rangle_k$ $\qquad$ $[A \to \alpha\bullet\mathbf{a}\beta, y] \in \langle\delta\rangle_k$

$[A \to X_1\cdot X_2\bullet\cdot\ldots\cdot X_n, x] \in \langle\delta\cdot X_1\cdot X_2\rangle_k$ $\qquad$ $[A \to \alpha\mathbf{a}\bullet\beta, y] \in \langle\delta\cdot\mathbf{a}\rangle_k$

…

$[A \to X_1\cdot X_2\cdot\ldots\cdot X_n\bullet, x] \in \langle\delta\cdot X_1\cdot\ldots\cdot X_n\rangle_k$