17. Активные префиксы. LR(0)-пункты. Теорема об LR(0)-автомате. Следствия из теоремы об LR(0)-автомате.

Определение

LR(k)-пункт $[A \to \beta_1 \cdot \beta_2, v]$ *допустим для активного префикса* γ, если $\gamma = \gamma_1\beta_1$ и существует правый вывод $S' \Rightarrow^* \gamma_1 A w \Rightarrow \gamma_1\beta_1\beta_2 w \Rightarrow^* uw$, v — префикс $w\dashv$.

Допустимость пункта определяет действия распознавателя в случае, когда содержимое стека равно $\gamma = \gamma_1\beta_1$. Если $\beta_2 = \lambda$, то производится свёртка, иначе — перенос. Для данного активного префикса могут быть допустимы несколько пунктов. Далее будем рассматривать LR(0)- и LR(1)-пункты.

опр. Активный префикс — это содержимое стека в любой момент времени работы распознавателя, "перенос-свертка" (префикс r-формы не выходящий за правый конец основы)

Определение

Автомат LR(0)-пунктов расширенной грамматики — это λ-НКА $\mathcal{I}_G = (\Sigma \cup \Gamma, I, \delta, i_0, I)$, где:
I — множество LR(0)-пунктов грамматики G;
$i_0 = [S' \to \cdot S]$;
множество переходов δ состоит из переходов двух типов:
$[A \to \beta_1 \cdot X\beta_2] \xrightarrow{X} [A \to \beta_1 X \cdot \beta_2]$ (базисные) и
$[A \to \beta_1 \cdot B\beta_2] \xrightarrow{\lambda} [B \to \cdot\beta]$.

Базисными пунктами называются пункты i_0 и те пункты, где $\beta_1 \neq \lambda$.

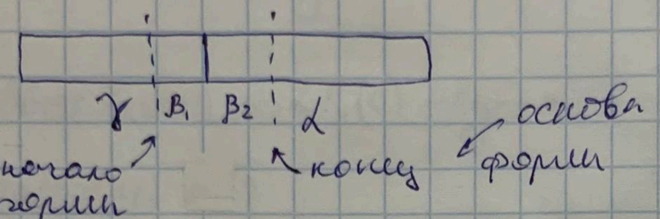
Основная теорема LR-анализа

Пункт $[A \to \beta_1 \cdot \beta_2]$ в грамматике G допустим для активного префикса γ тогда и только тогда, когда в автомате $\mathcal{I}_G$ существует путь из i_0 в $[A \to \beta_1 \cdot \beta_2]$, помеченный γ.

Лемма 1 $\forall$ АП γ $\exists$ допустимый для него базовый пункт.

Док-во (леммы):

Рассмотр вывод $S' \Rightarrow^* \gamma\alpha \Rightarrow^* w$ и первое появление γ в нём.



основы формы

основа формы не лежит в α целиком.

На предыдущем шаге применено $A \to \beta_1 \cdot \beta_2$

$$S' \Rightarrow^* \gamma' A u \Rightarrow \underbrace{\gamma' \beta_1}_{\gamma} \beta_2 u \Rightarrow^* w \Rightarrow \overset{\text{допустимый}}{[A \to \beta_1 \cdot \beta_2]}$$

для γ по опред $\beta_1 \neq \lambda \Rightarrow [A \to \beta_1 \cdot \beta_2]$ базисный

$\gamma = \lambda$ - допустимый базисный пункт i_0 ∎

Лемма 2

Пункт $[B \to \cdot \beta]$ допустимый для АП $\gamma \Leftrightarrow$ $\exists$ путь, помечен λ, в $[B \to \cdot \beta]$ из некоторого базисного пункта, допустим для γ.

Док-во (Леммы):

I (⇐) Пусть $[A \to \beta_1 \cdot B\beta_2]$ допустим для $\gamma \Rightarrow \exists$ вывод.

$S' \Rightarrow^* \gamma' \beta_1 B \beta_2 u \Rightarrow^* \gamma B v u \Rightarrow^* w \Rightarrow \exists$ вывод $S' \Rightarrow^*$
$\Rightarrow^* \gamma B v u \Rightarrow \gamma \beta v u \Rightarrow^* w' \Rightarrow [B \to \cdot \beta]$ допустимый для γ.

II (⇒) Пункт $[B \to \cdot \beta]$ допустимый для АП γ'

$\exists$ вывод.

$S' \Rightarrow^* \gamma B u \Rightarrow \gamma \beta u \Rightarrow^* w$

Расм первое появлении γ в этом выводе. По 11 утв соответ некотор базисн. пункт $[A \to \beta_1 \cdot \beta_2]$, допуст для γ

$S' \Rightarrow^* \gamma' A v \Rightarrow \gamma' \beta_1 \beta_2 v \Rightarrow \gamma \beta_2 v \underset{\text{0 или более шагов}}{\Rightarrow^*} \gamma B u \Rightarrow^* w$

если $\beta_2 = B\beta_2'$, то наш базисный пункт - $[A \to \beta_1' \cdot B\beta_2']$ по опред I_G есть λ-переход в пункт $[B \to \cdot \beta]$ если $\beta_2 \Rightarrow^+ B\beta_2'$, то β_2 начинается с нетерм C, $C \Rightarrow^+ B\alpha$ (т.к. вывод правый)

$\beta_2 = C\alpha_1 \qquad C \Rightarrow B_1 u_1$

$B_1 \Rightarrow B_2 u_2$

$\dots$

$B_n \Rightarrow B u$

Вернёмся к док-ву теоремы

$\exists$ правила $(C \to B_1 u_1) \dots (B_n \to B u) \in P$

в I_G есть переход $[C \to \cdot B_1 u_1] \xrightarrow{\lambda} [B_1 \to \cdot B_2 u_2]$

$[B_n \to \cdot B u] \xrightarrow{\lambda} [B \to \cdot \beta]$ ▥

Вернёмся к док-ву теорем:

I $(\Leftarrow)$

Индукция по $|\gamma|$

БИ $\gamma = \lambda$

$i_0 = [S' \to \cdot S]$ допустим для $\gamma = \lambda$, по

Л2 все пункты достиж по λ-переход из

i_0, допустимы для $\gamma = \lambda$

ШИ $\gamma = \bar{\gamma} X$

рассм. базисные пункт для γ (Л1): $[A \to \beta_1 X \cdot \beta_2]$

$\exists$ переход

$[A \to \beta_1 \cdot X \beta_2] \xrightarrow{X} [A \to \beta_1 X \cdot \beta_2]$ путь до

$[A \to \beta_1 \cdot X \beta_2]$ помечен $\bar{\gamma}$, для него выполн П.И.

$$S' \Rightarrow^* \bar{\gamma}' A u \Rightarrow \overbrace{\underbrace{\bar{\gamma}' \beta_1}_{\bar{\gamma}} X}^{\gamma} \cdot \beta_2 u \Rightarrow^* w \Rightarrow$$

$\Rightarrow [A \to \beta_1 X \cdot \beta_2]$ допустим для γ

Конечный пункт (достижим по пути, помеч γ)

либо совп с $[A \to \beta_1 X \cdot \beta_2]$, либо дост из него по λ-перех => по Л2 ок допуст. для γ.

II (=>)

индукция по $|\gamma|$

Б.И. $\gamma = \lambda$ Произвольный пункт, допуст для γ: $[A \to \beta_1 \cdot \beta_2] = [A \to \cdot \beta_2]$

Ш.И: $\gamma = \bar{\gamma} X$

по Л1 $\exists$ базисн пункт, допустим для γ: $[A \to \beta_1 X \cdot \beta_2]$

$S' \Rightarrow^* \bar{\gamma}' A u \Rightarrow \overbrace{\underbrace{\bar{\gamma}' \beta_1}_{\bar{\gamma}} X}^{\bar{\gamma} X} \beta_2 u \Rightarrow^* w \Rightarrow [A \to \beta_1 \cdot X \beta_2]$ допустим для $\bar{\gamma}$ и для него вып ПИ.

$\exists$ путь из i_0 в $[A \to \beta_1 \cdot X \beta_2]$, помечен. $\bar{\gamma}$ => добавим к пути переход $[A \to \beta_1 \cdot X \beta_2] \xrightarrow{X} [A \to \beta_1 X \cdot \beta_2]$, получим путь, помеч γ

по Л2 $\forall$ небазисного пункта допустим для γ $\exists$ путь, помечен λ в данный пункт из базисного пункта. ▨

Следствие 1

Язык, распознаваемый автоматом $\mathcal{I}_G$, совпадает со множеством всех активных префиксов грамматики G.

Все состояния автомата $\mathcal{I}_G$ являются заключительными. Если γ — активный префикс, то по теореме в автомате найдётся путь из начального пункта i_0, помеченный цепочкой γ, то есть автомат распознаёт γ. Обратно, если в автомате существует путь из i_0 в i, помеченный цепочкой γ, то γ является активным префиксом, т.к. по теореме i — допустимый для него пункт.

Хз нужно ли следующие следствия рассказывать:

Следствие 2

Множество пунктов, достежимых из $\mathcal{I}_0$ по пути помеченного γ, совпадают с множеством допустимых пунктов для активного префикса γ.

Следствие 3

Если в состоянии A_G есть пункты $[A \to \alpha_1 X \alpha_2]$ и $[B \to \beta_1 Y \beta_2]$, то X=Y.