Домашняя работа по ТРЯП №7

Автор - Айвазов Денис из 671 группы

18 октября 2017

1 Решить уравнения с рег. коэффициентами

1.1 $X = ((110)^* + 111^*)X$

В книге "Теория синтаксического анализа, перевода и компиляции"А. Ахо и Дж. Ульмана был представлен алгоритм решения подобных задач. Там использовалось утверждение "Пусть уравнение имеет вид $X = \alpha X + \beta \Rightarrow$, 1) если $\varepsilon \notin \alpha$, тогда $\alpha^*\beta$ — единственное решение 2) если $\varepsilon \in \alpha$, тогда $\alpha^*(\beta + L)$ — решение для любых L".
В нашем случае $\alpha = (110)^* + 111^*$ а $\beta = \emptyset$, и причем $\varepsilon \in \alpha$ Тогда выполняется 2 пункт утверждения и путем разложения по Тейлору(было на семинаре) мы получим $X = ((110)^* + 111^*) \cdot L$;
Таким образом частное решение и решение минимальное по включению $X = \emptyset$; Все решения задаются $X = ((110)^* + 111^*) \cdot L$; $\forall L$;

1.2 $X = (00 + 01 + 10 + 11)X + (0 + 1 + \varepsilon)$

По утверждению из предыдущего пункта в данном случае $\alpha = (00 + 01+10+11), \beta = (0+1+\varepsilon)$. Но теперь выполняется случай 1 т.к. $\varepsilon \notin \alpha$. И тогда $\exists! X = \alpha^*\beta = (00+01+10+11)^*+(0+1+\varepsilon)$ - это и есть единственное решение. А т.к. оно единственное, то оно совпадает с минимальным

1.3 Система с регулярными коэффициентами

$$\left\{\begin{array}{lcl} Q_0 & = & 0Q_0 + 1Q_1 + \varepsilon \\ Q_1 & = & 1Q_0 + 0Q_2 \\ Q_2 & = & 0Q_1 + 1Q_2 \end{array}\right. = \left\{\begin{array}{rcl} (\varepsilon - 0)Q_0 & = & 1Q_1 + \varepsilon \\ Q_1 & = & 1Q_0 + 0Q_2 \\ (\varepsilon - 1)Q_2 & = & 0Q_1 \end{array}\right. =$$

$$= \left\{\begin{array}{lcl} Q_0 & = & 0^*(1Q_1 + \varepsilon) \\ Q_1 & = & 1Q_0 + 0Q_2 \\ Q_2 & = & 1^*0Q_1 \end{array}\right. = \left\{\begin{array}{lcl} Q_0 & = & 0^*(1Q_1 + \varepsilon) \\ Q_1 & = & 10^*(1Q_1 + \varepsilon) + 01^*0Q_1 \\ Q_2 & = & 1^*0Q_1 \end{array}\right. \Rightarrow$$

$\Rightarrow (\varepsilon - (10^*1 + 01^*0))Q_1 = 10^* \Rightarrow Q_1 = (10^*1 + 01^*0)^*10^*$
$Q_2 = 1^*0Q_1 = 1^*0(10^*1 + 01^*0)^*10^*$
$Q_0 = 0^*(1Q_1 + \varepsilon) = 0^*1(10^*1 + 01^*0)^*10^* + 0^*$

Ответ: $\left\{\begin{array}{lcl} Q_0 & = & 0^*(1Q_1 + \varepsilon) = 0^*1(10^*1 + 01^*0)^*10^* + 0^* \\ Q_1 & = & (10^*1 + 01^*0)^*10^* \\ Q_2 & = & 1^*0Q_1 = 1^*0(10^*1 + 01^*0)^*10^* \end{array}\right.$ Это (единственное) как частно, так и общее решение, которое является наименьшей неподвижной точкой.

2 Верно ли, что для любой линейной грамматики $G, L(G) \in REG$?

Т.е. верно ли, что любая линейная грамматика порождает регулярный язык. Ответ - нет.
Приведем в контрпример нерегулярный язык, который уже стал классическим примером: $L = \{a^n b^n \mid n \geq 0\}$. Он нерегулярен и это было доказано на семинаре применением теоремы Майхилла-Нероуда. Сами классы эквивалентности это классы $a^i b^j \forall i, j \in \mathbb{N}_0$. Они различны т.к. для мы имеем множества a^i и a^j, которые различает слово b^i приписыванием его первое слово принимается, а второе - нет.

Наш язык $L = \{a^n b^n \mid n \geq 0\}$ задается грамматтикой: $\Gamma = G(N, T, P, S)$, где $N = \{S\}$, $T = \{a, b\}$, $P = \{S \to aSb \mid \varepsilon\}$, S=S;

Докажем по индукции по длине слова (2n) сначала, что слова, порождаемые грамматикой принадлежат языку:
База - $S \Rightarrow \varepsilon; S \Rightarrow aSb \Rightarrow ab$. Шаг индукции. Предположим, что слово $\omega \in L(G) : |\omega| = 2n$. Значит оно порождается применением правила 1 n раз, а затем применением правила 2. Значит если мы применяем не n а n+1 раз правило 1, то получим, что слово $0\omega 1 \in L(G)$ и $|\omega| = 2n + 2$ - тоже из языка.
Теперь, докажем обратное включение - $L \subseteq L(G)$. База - пустое слово выводится грамматикой и принадлежит языку. Теперь допустим, что любое слово $\omega \in L : |\omega| = 2n$ порождается грамматикой. Тогда следующее по длине слово из языка (длины 2n+2) будет иметь вид $0\omega 1$ и получаться применением правила 1 на этапе порождения слова ω, а затем применением правила 2.
Отсюда следует включение в обе стороны т.е. наша линейная грамматика порождает нерегулярный язык$L = \{a^n b^n \mid n \geq 0\}$

3 * Показать, что $L^=$ не порождается линейными КСГ

4 Верно ли, что если $F = \{a^n b^n | n \geq 1\} \subseteq L_1 \cap L_2$ то L_1 или L_2 нерегулярный

1. В общем случае это не верно. Приведем контрпример:
   Возьмем $L_1 = a^*b^*, L_2 = a^+b^+; L_1, L_2 \in REG$ т.к. эти языки заданы РВ. Тогда $F = \{a^n b^n | n \geq 1\} \subset a^+b^+ = L_2 = L_1 \cap L_2$. Таким образом мы получили что нерегулярность подмножества пересечения двух языков не гарантирует нерегулярность хотя-бы одного из них.

2. Теперь покажем, что хоть это не обязательно, но может случиться, что языки объединения могли быть нерегулярными(или по крайней мере один из них). Возьмем $L_1 = \{a^n b^n | n \geq 0\}, L_2 = a^+b^+; L_1 \notin REG; L_2 \in REG$. Теперь $F = \{a^n b^n | n \geq 1\} = L_1 \cap L_2$.

5 Постройте КС-грамматики для языков:

5.1 $\{a, b\}^* \backslash PAL$

Вот сама грамматика, заданная аксиомой S, алфавитом $\Sigma = \{a, b\}$, и нетерминалами S, D, O и выводами: $\{S \to aSa|bSb|D; D \to aOb|bOa; O \to aO|bO|\varepsilon\}$. Это работает так: переходы по S значат symmetry - сохраняют симметрию палиндрома. D - значит dissymmetrical - т.е. (единственный раз) нарушающий симметрию переход. O - others, т.е. добавление любых букв внутри, когда мы уже нарушили симметрию снаружи.
Сначала докажем, что наша КСГ распознает все непалиндромы:
$\forall \omega \in L = \alpha x \beta \alpha^R$ - где α - какая-то (возможно пустая) последовательность, сохраняющая "палиндромность". $x, y \in \{a, b\} : x \neq y$. $\beta \in \Sigma^*$. Выводимость следует сразу из названий нетерминалов - осуществляем вывод из S (любой из первых двух, в которых справа тоже есть S) таким образом мы получаем симметричную картину, в которой затем из S делаем O, один раз нарушаем симметрию неравными символами с двух сторон и затем внутрь вклиниваем любую комбинацию символов: $S \Rightarrow \alpha S \alpha^R \Rightarrow \alpha D \alpha^R \Rightarrow \alpha x O y \alpha^R \Rightarrow \alpha x \beta y \alpha^R = \omega \in \{a, b\}^* \backslash PAL$. Ч.т.д.

В обратную сторону, что КСГ не распознает палиндромы. Т.е. покажем вывод от хвоста к голове.
$\forall \omega$ применяем правила вывода $S \to aSa, S \to bSb$ в обратную сторону, пока у нас с двух концов ω одинаковые символы. Затем, если омега закончилось, значит оно было палиндромом. Но мы не можем его вывести из грамматики, т.к. пока у нас в выводе присутствует нетерминал S, который пропадет только в случае, когда найдутся две разных буквы с двух концов(тогда $xSy \to D \to$), а оставшаяся последовательность нам не важна(она может быть любой, но точно будет выводиться по правилам вывода из O), т.к. палиндромность уже нарушена и существует место, где на равном расстоянии от концов стоят разные буквы. Значит слово непалиндром, и оно выводится КСГ. Что и требовалось доказать.