Домашнее Задание по ТРЯПу №7

Павливский Сергей Алексеевич , 873

24.10.2019

Задание 1.

Постройте конечный автомат по грамматике G :
S $\rightarrow$ abaA | abB | ε, A $\rightarrow$ aB | aa, B $\rightarrow$ bA | aS

Решение

Пусть имеется праволинейная грамматика. Построим по ней конечный детерминированный автомат. Введём специальное допускающее состояние ok. Множеством состояний автомата будет множество нетерминалов грамматики вместе с состоянием ok (Q=N$\bigcup$ok). Для правил вида A$\rightarrow$aB определим функцию перехода в автомате как δ(A,a)=B. Для правил вида A$\rightarrow$a определим функцию перехода в автомате как δ(A,a)=ok.

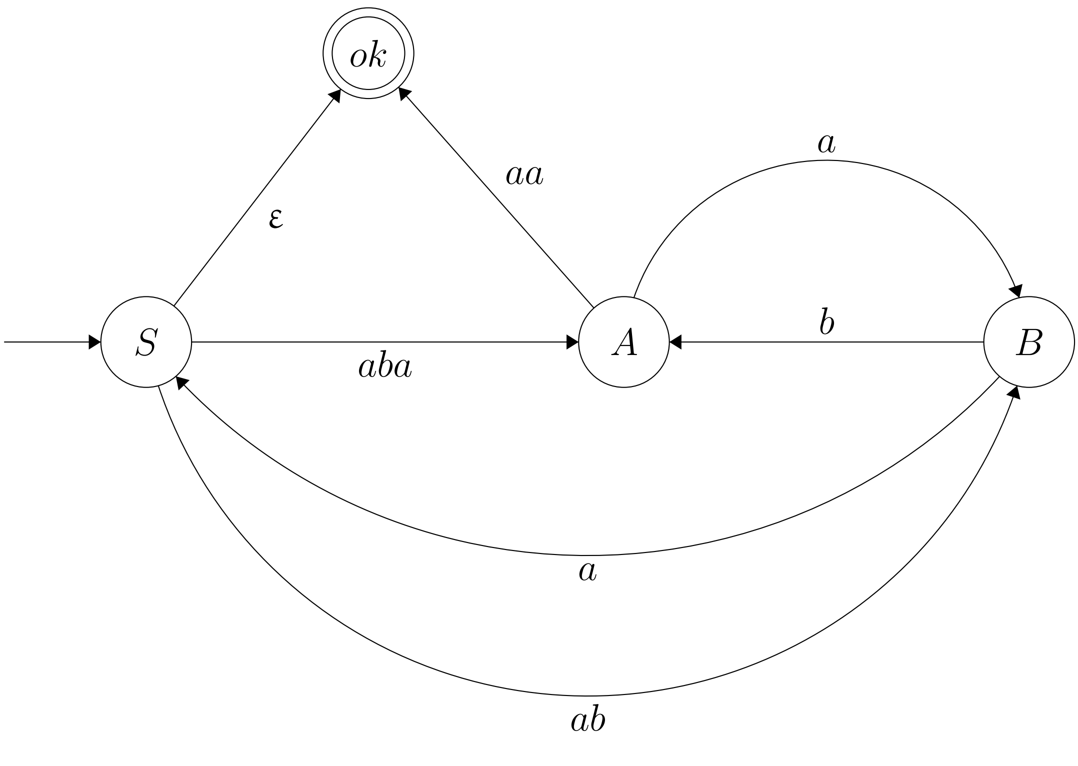
Док-во корректности алгоритма :

Докажем, что если слово выводится в грамматике, то оно допускается автоматом. Рассмотрим последовательность применений правил, дающую слово a длины k. Для каждого правила вида A$\rightarrow$aB в автомате существует переход из состояния A в состояние B по символу a. Таким образом, если после k-1 применения правил мы можем получить строку вида a$c^{-1}B$, то в автомате имеется соответствующая последовательность переходов $\langle$S,a$\rangle \vdash^{k-1} \langle$B,c$\rangle$, а поскольку можно вывести a, то хотя бы для одной строки такого вида существует правило B$\rightarrow$ c, а значит в автомате есть переход $\langle$B,c$\rangle \vdash \langle$ok,ε $\rangle$. Таким образом автомат допускает слово a.

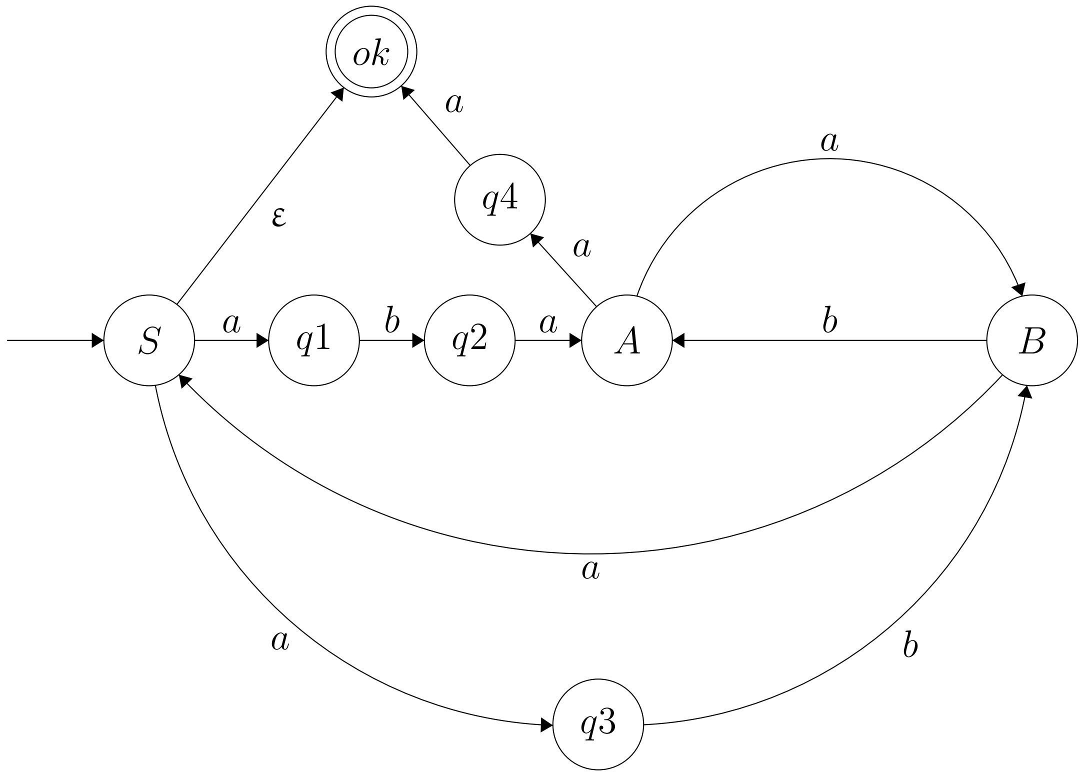
Докажем, что если слово допускается автоматом, то его можно вывести в грамматике. Рассмотрим слово a длины k. Рассмотрим какую-либо последовательность переходов автомата, допускающую данное слово $\langle$S,a$\rangle \vdash^{k} \langle$ok,ε $\rangle$. Для каждого одношагового перехода в автомате существует соответствующее правило в грамматике. Значит для подпоследовательности переходов из k-1 шага $\langle$S,ε $\rangle \vdash^{k-1} \langle$U,c$\rangle$ существует соответствующая последовательность применений правил S$\Rightarrow^{k-1}$ ac^{-1}U. Для последнего перехода в автомате $\langle$U,c$\rangle \vdash \langle$ok,ε $\rangle$ существует правило U$\Rightarrow$c. Таким образом, существует последовательность применений правил грамматики, выводящая слово a.

Источник : neerc.ifmo.ru

Тогда построим по данному алгоритму КА :



Если разбить переходы по строчкам на переходы по символам , то получится :



Получили НКА , который аналогичен изначальной праволинейной грамматике :

Слово принадлежит языку , задаваемому грамматикой , если оно выводимо из правил ее вывода ; слово принадлежит языку , задаваемого НКА , если существует путь из начальной вершины в принимающую по буквам этого слова ; состояние A соответствует тому , что слово в состоянии S имеет вид abaA ; состояние B соответствует тому , что слово в состоянии A имеет вид aB, а также что слово в состоянии S имеет вид abB ; состояние ok соответствует тому , что слово в состоянии S имеет OK , а также , что слово в состоянии A имеет вид aaOK . Методом перебора все состояния покрывают все возможные правила перехода из исходной грамматики .

Задание 2.

Является ли грамматика G из предыдущей задачи однозначной?

Решение

Нет , неоднозначна . Слово abbaa выводится из последовательности переходов S $\to$ abB $\to$ abbA $\to$ abbaa , а также из последовательности S $\to$ abB $\to$ abbA $\to$ abbaB $\to$ abbaaS $\to$ abbaa . Так как для одного слова два различных дерева вывода , то по определению грамматика неоднозначна .

Задание 3.

Язык L задан КСГ: S $\to$ aSa | aSb | bSa | bSb | a.

1. Является ли L регулярным языком?
2. Является ли дополнение L регулярным языком?

Решение

Правила вывода устроены так , что они заменяют центральный элемент слова на элемент с символом с каждой из сторон (возможны пустым) . Но если замененный элемент был центральным , то есть количество с каждой из сторон от него было одинаково , то и новый элемент будет также центральным , так как добавляемых вместе с ним элементов одинаковое количество с каждой из сторон от него . Тогда данная КСГ порождает любое слово вида : w_1aw_2 , где w_1 = $(a|b)^*$, w_2 = $(a|b)^*$, $|w_1|$ = $|w_2|$. Докажем его нерегулярность по лемме о накачке . Для любого p возьмем слово вида b^pab^p . Тогда y - это последовательные элементы из первого b^p , а тогда при i = 2 получится слово b^kab^p, k > p (т.к. |y| > 0) $\notin$ L . Противоречие . Значит L $\notin$ Reg , а так как регулярность замкнута относительно операции дополнения , то и $\overline{L} \notin$ Reg.

Задание 4.

Построить для следующих языков, заданных над алфавитом $\{a, b\}$, КС-грамматики:

а) PAL = $\{ \mathrm{w} \mid \mathrm{w} = \mathrm{w}^R \}$;

б) L = $\{ a^n b^m \mid \mathrm{n} \leqslant \mathrm{m} \leqslant 2\mathrm{n} \}$;

в) $\Sigma^* \setminus \{ a^n b^n \mid \mathrm{n} \geqslant 0\}$.

Решение

а)

S $\to$ aAa | bAb | a | b

A $\to$ aAa | bAb | a | b | ε

Докажем по индукции по длине слова верность данной КС грамматики .

База : длина слова 1 - возможные палиндромы a, b , которые получаются по правилам вывода S $\to$ a и S $\to$ b соответственно .

Переход : пусть мы получили все слова длины от 1 до k . Докажем , что мы можем получить все слова длины k+1 . Слово слово длины k+1 может иметь вид wa$w_1$aw^R или wb$w_1$bw^R , где $w_1 \in$ a, b, ε ; w - произвольное слово такое , что $|\mathrm{wa}w_1\mathrm{aw}^R| = |\mathrm{wb}w_1\mathrm{bw}^R| = \mathrm{k} + 1$. Но данные слова получаются из слова вида wAw^R переходами wAw^R $\to$ waAaw^R $\to$ wa$w_1$aw^R , или wAw^R $\to$ wbAbw^R $\to$ wb$w_1$bw^R соответственно , что возможно , так как множество значений w_1 является подмножеством возможных значений переходов A . A слово $|\mathrm{wAw}^R| < \mathrm{k} + 1$, а значит по предпололожению индукции выводимо . Значит и любое $\forall$ w : $|\mathrm{w}| = \mathrm{k} + 1$ выводимо , а значит , по индукции , все слова из PAL выводимы из данной грамматики . Значит язык PAL принадлежит языку , порождаемому грамматикой . Обратное включение очевидно следует из доказательства первого включения . Значит языки равны ч . т . д .

б)

S $\to$ aSb|aSbb|ε

Каждое слово выводимое по правилам грамматики - это конечный набор переходов согласно правилам вывода . Пусть для вывода некоторого слова w по данному правилу было использовано k переходов (переход по ε не будем считать , т.к. он финальный) . Тогда количество букв a в слове w равно k . Минимальное количество букв b в слове w - это если каждый раз мы добавляли минимально возможное при переходе число букв b , т.е. 1 , то есть миниимальное количество букв b в w равно k . Аналогично , максимальное равно 2b . Значит $n \leqslant m \leqslant 2n$. Почему перебираются все слова такого вида ? Для любого требуемого количества t букв a мы можем сделать t переходов aSbb, опять же без учета перехода по ε , и , как ранее было сказано , получить t букв a в начале (так как по индукции легко видно , что после каждого перехода перед S в слове стоят только буквы a , а после буквы b) , и 2t букв b , которых с гарантией хватит на любое слово , а далее , если 2t > количества букв b равного q в требуемом слове , то последовательно заменяем по одному переходу aSbb на aSb , до тех пор , пока количество букв b в полученном слове $\neq$ q (q всегда достижимо , так как за одну замену мы уменьшаем количество букв b на 1 , а так как мы начинаем из верхней границы множества допустимых значений q , а элементы множества расположены на расстоянии 1 , равное шагу , с которым мы меняем количество букв b , то мы проходим все допустимые значения q ; то есть если слово корректное , то мы гарантированно достигнем уменьшением количества букв b на 1 значения q) . После данной последовательности преобразований слова с максимальным количеством букв b , получаемого после t преобразований , мы получаем любое слово удовлетворяющее условию с количеством букв a равным t . Тогда , т.к. в любом слове , которое может быть необходимо вывести конечное число букв a , то по вышеуказанному алгоритму это делается за (количество букв a в требуемом слове) + 1 шагов (в конце еще преобразуем оставшееся S в ε) . Значит , мы построили алгоритм вывода любого слова из языка по правилам данной грамматики , что и требовалось . Значит язык L принадлежит языку порождаемому грамматикой . Обратное включение очевидно из доказательства первого включения . Значит языки равны ч . т . д .

в)

$S \to bA|Aa|aSb$

A → aA | bA | ε

Включение в одну сторону очевидно : если у нас переходит переход из S в bA или Aa , то мы переходим из слова вида a^kSb^k (видно из правил : единственный переход в S не переходящий в bA или Aa , где из A нельзя перейти в S - это aSb , который сохраняет количество a и b с обеих сторон от S одинаковым , и не меняет то , что слева от S только a , а справа только b) . После перехода из такого слова в a^kbAb^k , выводимое слово гарантированно $\notin$ { $a^n b^n$| n $\geqslant$ 0} , так как количество b в нем уже > количества a , а при добавлении еще одного a оно уже окажется справа от b , то есть будет нарушено словие порядка . Аналогично , a^kAb^k $\notin$ { $a^n b^n$| n $\geqslant$ 0} . Значит язык , порождаемый грамматикой , не содержит слов $\in$ { $a^n b^n$| n $\geqslant$ 0} , а значит он $\in$ Σ^* \ { $a^n b^n$| n $\geqslant$ 0} .

В другую сторону : любое слово $\in$ Σ^* \ { $a^n b^n$| n $\geqslant$ 0} может быть представлено в виде a^twb^t , где w - некоторое слово $\neq a^q b^q$. Тогда первыми t переходами вида S → aSb грамматика порождает требуемую оболочку a^tSb^t вокруг S . А далее любое возможно слово w выражется переходами по правилам грамматики , так как возможны несколько вариантов : если первая буква w = b , то делается переход S → bA , а дальше для каждой последующей буквы слова w берется переход с такой же буквой в начале , либо aA либо bA , а когда буквы закончатся , A заменяется на ε ; если же первая буква слова w - это a , то последняя буква слова w может быть только a , иначе бы первая буква была a, последняя b , и они должны были бы быть включены в оболочку . Тогда делается 2 перехода S → Aa , помещая a в конец w , а потом Aa → aAa , помещая другое a в начало слова , а дальше , аналогично случаю с началом слова w на b , для каждой встреченной буквы w берем соответствующий переход aA или bA , а когда буквы закончатся делаем переход A → ε . Значит каждое слово w $\in$ $\in$ Σ^* \ { $a^n b^n$| n $\geqslant$ 0} также $\in$ языку , порождаемому грамматикой , то есть доказано включение в другую сторону . Значит языки равны ч . т . д .