Филиппова Анастасия

Домашнее задание 1

(1) **Постройте грамматику для языка $\{a^n b^m | n > m > 0\}$ и докажите, что она на самом деле порождает этот язык.** Построим следующую грамматику:
$G = (\{S, A\}, \{a, b\}, \{S \to aA, A \to aA | aAb | ab\}, \{S\})$
Докажем по индукции, что данная грамматика порождает язык $\{a^n b^m | n > m > 0\}$, натуральный параметр - количество совершенных переходов:
1) База $k = 2$:
$S \xrightarrow{a} A \to \left\{a^2 b, a^2 Ab, a^2 A\right\}$
2) Пусть на $k = n$ переходе:
$S \xrightarrow{n} \left\{a^n b^m, a^n Ab^m, a^n A\right\}$, где m принимает все значения из отрезка $[2, n-1]$.
Индукционный переход: на n+1 переходе получаютя цепочки:
$S \xrightarrow{n+1} \left\{a^{n+1} b^{m+1}, a^{n+1} Ab^{m+1}, a^{n+1} A\right\}$, где m принимает все значения из отрезка $\in [2, n-1]$.
Рассмотрим n+1 переход:
$a^n Ab^m \to \left\{a^{n+1} b^{m+1}, a^{n+1} Ab^{m+1}, a^{n+1} Ab^m\right\}$
$a^n A \to \left\{a^{n+1} b, a^{n+1} Ab, a^{n+1} A\right\} \Rightarrow$ предположение индукции доказано, то есть данная грамматика действительно пораждает данный язык.

(2) **Докажите, что для любого регулярного множества R можно построить грамматику G, которая будет порождать это множество.**
Регулярное множество R будет описываться регулярным выражением r. Будем доказывать по индукции, натуральным параметром является длина регулярного выражения.
1)База индукции: $n = 0$ - пустое слово:$G = (\{S\}, \{\}, \{S \to \epsilon\}, \{S\})$
2) Пусть для регулярных множеств, которые описываюся регулярным выражением длины k, существует грамматика, которая порождает это множество. Докажем, что это справедливо для выражений длины больше k. Пусть P, Q - регулярные множества, описывающееся регулярными выражениеми $p, q, |p| = k, |q| = l, l <= k$. По предположению индукции $G_1(P) = (\{N_1\}, \{T_1\}, \{P_1\}, \{S_1\}), G_2(Q) = (\{N_2\}, \{T_2\}, \{P_2\}, \{S_2\})$.
$p|q$ описывает множество $P \cup Q$. Построим следующую грамматику: $G = (\{N_1, N_2\}, \{T_1, T_2\}, \{S \to S_1 | S_2\}, \{S\})$. Легко проверить, что такая грамматика порождает множество $P \cup Q$ - длина регулярного вы ражения $|P|q| = k + l + 1$.
Для множества PQ построим $G = (\{N_1, N_2\}, \{T_1, T_2\}, \{S \to S_1 S_2\}, \{S\})$. Очевидно, что она порождает множество PQ, длина регулярного выражения $|pq| = l + k + 1$.
Для множества P^*: зададим грамматику по индукции: для P^1, P^0 - очевидно, для P^h $G = (\{N_1\}, \{T_1\}, \{S \to S_1 ... S_1\}, \{S\})$, где $S_1 ... S_1$ - S_1 n раз.
Таким образом, получаем, что для любого регулярного выражения длиной больше k существует грамматика, которая их порождает. (так как регулярное выражение длины k есть подвыражения меньшей длины, связанные опрециями | или * или конкатенация. Что и требовалось доказать.

(3) **Постройте грамматику для языка $\{a^n b^m | n, m > 0, m \neq n\}$ и докажите, что она на самом деле порождает этот язык.**
Обозначим за $P = \{a^n, a^m\}^*, n > m > 0, Q = \{a^n, b^m\}, 0 < n < m$. Тогда искомый язык $L = P \cup Q$. Для множества P грамматика $G_1 = (\{S_1, A_1\}, \{a, b\}, \{S \to aA, A_1 \to aA_1 | aA_1 b | ab\}, \{S_1\})$, для Q аналгично: $G_2 = (\{S_2, A_2\}, \{a, b\}, \{S_2 \to A_2 b, A_2 \to A_2 b | aA_2 b | ab\}, \{S_2\})$, тогда $G(L) = (\{S, A_1, A_2, S_1, S_2\}, \{a, b\}, \{S \to S_1 | S_2\}, \{S\})$

(4) **Постройте РВ для языка:**
а) Всех всех слов четной длины над алфавитом $\{a, b\}$
б) Всех всех слов, содержащих ровно одно подслово "ab"над алфавитом $\{a, b\}$
в) Всех слов в "camelCase"над алфавитом a, b, A, B (слово начинается с нескольких строчных букв, далее повторяются блоки из одной заглавной и некоторого, возможно нулевого, количества строчных букв)
а) $(aa|bb|ab|ba)^*$
б) $(b|\epsilon)(b^*|a^*)(ab)(a^*|b^*)(a|\epsilon)$
в) $(a|b)(a|b)^*((A|B)(a|b)^*)^*$

(5) **Постройте РВ для языка $L \subseteq \{t, x\}^*$ всех слов, которые начинаются и заканчиваются на "txt"и докажите его корректность.**

$(txt)(t|x)^*(txt)|(txtxt)|(txt)$ - регулярное выражение для данного языка. Докажем, что построенное множество содержит те и только те слова, которые указаны в условии.

Сначала докажем включение $L \subseteq R$, где R - регулярное множество, описываемое данным регулярным выражением. Пусть $w \in L$, тогда возможны 3 случая: $w = txt, w = txtxt$ - очевидно, что $w \in R$, третий вариант: $w = txtutxt$, где $u \in \{x, t\}^*$ Очевидно, что $\forall w \in L \hookrightarrow w \in R \Rightarrow L \subseteq R$.

Докажем обратное включение: пусть $w \in R$. Очевидно, что $w = txtutxt|txtxt|txt$, где $u \in \{t, x\}^* \Rightarrow R \subseteq L$.

Доказали, что $L = R$, то есть данное регулярное выражение корректно.

(6) 1)Есть процедура генерации цепочек языка. Чтобы проверить, принадлежит ли $w \in \{a, b\}^*$ языку, нумеруем сгенерированные цепочки. Далее сравниваем w со всеми цепочками L. Если принадлежит, процедура останавливается, если нет, то, если язык L конечен, то сравнив w с каждой цепочкой, не найдя соответсвие, процедура остановится с ответом нет. Если язык L не является конечным, то процедура никогда не остановится.

2)В обратную сторону: занумеруем слова из $\{a, b^*$ в лексикографическом порядке: $1 - \epsilon, 2 - a, 3 - b, 4 - aa, 5 - ab, 6 - ba, 7 - bb...$Если принадлежит - добавляем в L, если нет, итерируемся далешь по номерую Таким образом перечислим все цепочки L.