**Лабораторная работа №7**

***Лексический анализ. Задания.***

Для заданного варианта лексики:

* построить диаграмму состояний и переходов распознающего КА, обратить внимание на количество возвращаемых символов в каждом заключительном состоянии и на логику распознавания префиксов и суффиксов (начала и окончания) лексем;
* проверить правильность работы КА при помощи [Java-приложения](http://ermak.cs.nstu.ru/trans/Trans51.htm) [http://ermak.cs.nstu.ru/trans/exe.jpg](http://ermak.cs.nstu.ru/trans/Examples/LEX.jar)(приложение можно скачать либо взять в папке для ИВТ файл LEX.jar).

<http://ermak.cs.nstu.ru/trans/Trans51.htm> - использовать для дополнения к заданию.

1.  **00111100** – цепочка, содержащая четное число подряд идущий единиц и нулей. *Подсказка*: *автомат имеет состояния, соответствующие получению четного 0, нечетного 0, четной 1 и нечетной 1.* Проверка лексической правильности **всей** строки.

*2.*   Цепочка символов **a,b,c**в любых сочетаниях. Лексема заканчивается цепочкой одинаковых символов. Например: **abcbbb|abcc|abababaaaa|abb.***Подсказка: состояния запоминают, какой был последний символ, и что этот символ повторялся.*

*3.*   Лексемы вида **+01 | +10 | + | 11010010 –**после + лексема может содержать **01** или **10**, любое другое сочетание (**11** или **00**) считается началом другой лексемы, перед которой идет одиночный +. Например: **+|+01|01110+|11010|+10** *Подсказка: в некоторых заключительных состояниях необходим возврат 2 символов.*

*4.*   Строка символов **a,b,c,**в которой любая пара одинаковых символов разделяется символом «-»,например **a-a-abc-c-cb-b-bcaca-a.** Проверка лексической правильности **всей** строки. *Подсказка: состояния запоминают, какой был последний символ, и что за ним был символ «-».*

5.    Строка символов **a,b,c,**в которой любая пара различных символов разделяется символом «-»,например **aaa-b-ccc-bbb-c-a-c-aa.** Проверка лексической правильности **всей** строки.

*6.*   **((…)..(..)((..))) –**строка, содержащая скобки. Если в строке есть последовательность одинаковых скобок, то между ними не должно быть других символов. Например, **(aaa(**aaa)) – строка с ошибкой. Проверка лексической правильности **всей** строки. *Подсказка: состояние запоминает, какая была последняя скобка, и что за ней был хотя бы один символ.*

**7.**   Строка символов **a,b,**цепочка одинаковых символов считается лексемой, причем, если за ней следует такая же цепочка, то она не включается в лексему, иначе к ней присоединяется еще один символ. Например: **aaa|bb|aaab|aab|aac|ababab|aa (!!!!).**

*8.*    Строка содержит символы **a** и **\*.**Лексемами считаются цепочки из любого числа символов **a,**а также нечетного числа символов **\****Подсказка: при контроле четности просмотренных символов****\*****использовать возврат 2 символов.*

*9.*   Строка цифр **1,2,3** в любой последовательности. Выделяется лексемы **12, 23, 31**, а также любые другие последовательности, например **12/31/23/23/1333/23/132.***Подсказка: в начальном состоянии отслеживается получение пар***12, 23, 31***, которые выделяются как лексемы. Для всех других последовательностей производится переход в новые состояния, в которых появление этих же пар приводит к распознаванию предыдущей лексемы и  возврату 2 символов.*

*10.* Строка содержит пробелы-разделители, цепочки символов **a**, точки и многоточия (три точки). Лексемами считаются цепочки символов **a**, точки и троеточия, например **aaa|...|.|aaa|.|.|aaa|...|aaa**  *Подсказка: использовать возврат 2 символов.*

**11.** Строка символов **a,b,**лексемами считаются цепочки, в которых каждый символ повторяется подряд более одного раза. Одиночные символы считаются отдельными лексемами, например: **aaabbbaabbbbb|a|bbaa|b|a|bbb.***Подсказка: использовать возврат 2 символов.***(!!!!)**

12. **a+++a-aaa+a--+a**- переменные из символов **a**, операции постинкремента и постдекремента, сложения и вычитания  - подряд могут идти 1 или 3 знака, в последнем случае – первые два одинаковые (автомат проверяет правильность **всей** цепочки).

**13.** Строка из символов **a,b,c**содержит лексемы **ab,bc,ca,** цепочки из остальных сочетаний распознаются как отдельные лексемы (**ab,bc,ca**выступают как их ограничители), например: **aaa|bc|acba|ca|bbbbbabbba|bc.***Подсказка: использовать возврат 2 символов.***(!!!!)**

14.  Лексемы**- (…..)**, **((…..))**, **(((….)))**– одинарные, двойные или тройные парные скобки, содержащие внутри любую цепочку символов. Меньшее количество закрывающихся скобок не считается концом лексемы.

**15.**  Строка содержит символы **a** и **\*.**Лексемы содержат четное число \* или **a,**последний нечетный символ считается отдельной лексемой, например: **\*\*\*\*\*\*|a|\*\*|aa|a|\*|aaaaaa|\*\*\*\*|a|\*\*|\*|a|\*\*\*\*|a.**

**16.**  Строка содержит символы **a** и **\*.**Лексемами считаются сочетания **\*a** и **a\*,**а также остальные цепочки, состоящие только из символов одного вида, например: **aa|a\*|a\*|\*\*\*|\*a|aa|a\*|a\*|\*\*|\*a|aaaa.**