Практикум работы на ЭВМ.

Задание 4. Синтаксический анализ простых формальных языков.

1 Примеры языковых программ

В каждом из ниже приводимых языков принимаются следующие обозначения:

- < буква > любая буква латинского алфавита;
- < цифра > любая десятичная цифра;
- < иден > идентификатор длины не более m (m = 2, 3 или 4);
- число > строка из пяти цифр, перед которой стоит знак «плюс» или«минус» (рассматриваются только целые числа);
 - < код > последовательность из 8 нулей и единиц;
 - < символ > символ, допускаемый в записи программы на языке;

<языковая программа> – последовательность операторов языка, разделенных точкой с запятой (;) и не имеющих пробелов; первый, встречающийся пробел, в записи программы, трактуется как ее конец. (Примечание. Вопрос об использовании пробелов в программе может быть оговорен отдельно. Тогда в качестве признака конца программы можно использовать точку.)

Перечень общих правил, проверка из них предлагается в качестве упрощенной постановки задачи *синтаксического контроля*:

- а) использование в программе только допустимых символов;
- б) длина идентификатора не превышает заданного числа m;
- в) число содержит ровно пять цифр;
- г) после точки с запятой может быть либо буква, либо пробел;
- д) после двоеточия может быть только знак равенства;
- е) количество открывающих скобок должно совпадать с количеством закрывающих скобок;
- ж) если идентификатор входит в правую часть некоторого оператора, то он должен входить и в левую часть какого-нибудь предшествующего оператора.

Дополнительные правила указываются в конкретных языках.

1.1 Язык арифметических операций

```
< символ > : := < буква > | < цифра > | < знак > | < разделитель > < знак > : := + | - | * | | < разделитель > : := ; | : | | =
```

< переменная > : := < иден > | < число >

< оператор> : = < иден > := < переменная > | < иден > := < переменная >

Выполнение оператора: вычислить значение по правой части оператора присваивания (справа от : =) и присвоить его идентификатору левой части. Знаки операций понимаются обычным способом (но деление нацело).

Пример языковой программы:

$$X := +00109$$
; $Y := X - -15700$; $A1 := X / Y$; $X := A1 + +14810$

Дополнительные правила:

- з) в правой части оператора не более трех знаков операций;
- и) после знака операции следует либо буква, либо цифра, либо знак «плюс» или «минус».
 - 1.2 Язык булевских операций

```
< символ > : = < буква > | < код > | < знак > | < разделитель> | < цифра >
```

$$< 3$$
HaK $> : := \lor | \rightarrow | + | * | \land$

< разделитель > : := ; | : | | =

< код > – последовательность из 8 нулей и единиц

< оператор > : = < иден > := < переменная > | < иден > := < переменная >

Выполнение оператора: вычислить код по правой части (справа от :=) и присвоить его идентификатору левой части. Вычисление кода производится взятием поразрядной операции над двумя кодами (либо простым присваиванием без вычисления). Операции обозначены: " \checkmark " – дизъюнкция, " * " – конъюнкция, " * " – сложение по модулю 2, " * " – импликация, " * " – операция Шеффера (1*1=0 и a*b=1, если a или b равно 0).

Пример языковой программы:

$$A1 := 11001010$$
; $B18 := 010111111$; $X := A1 * B18$; $Y := B18 - A1$; $A1 := X + 11110011$;

Дополнительные правила:

- з) длина кода равна 8 знакам 0 и 1;
- и) в правой части либо ни одного знака, либо один знак операции.

1.3 Язык операций над массивами

$$<$$
 символ $>$: := $<$ буква $>$ | $<$ цифра $>$ | $<$ знак $>$ | $<$ разделитель $>$ $<$ знак $>$: := $+$ | $-$ | $*$ | $/$

$$<$$
 разделитель $>$: := ; $|:| = |(|)| |,$

$$<$$
 массив $>$: $=$ $<$ иден $>$ $|$ $(<$ число $>$, . . . , $<$ число $>$ $)$

Примечание. Количество r чисел в массиве фиксируется в языке (r=3, 4 или 5).

< оператор> : := < иден > := < массив > | < иден > := < массив > < знак > < массив >

Bыполнение операторов: идентификатор обозначает массив из r чисел. Оператор вида

$$<$$
 иден $> := (<$ число $>, ..., <$ число $>)$

выполняется путем присваивания i - му элементу массива i - го числа. Например, $X:=(3,\,8,\,-19)$ означает серию операторов присваивания $X_{[1]}:=3$, $X_{[2]}:=8$, $X_{[3]}:=-19$. Выполнение оператора вида

сводится к копированию массива. Например, $A1:= \mathsf{B}$ означает $A1_{[1]}:= \mathsf{B}_{[1]};$ $A1_{[2]}:= \mathsf{B}_{[2]};$ $A1_{[3]}:= \mathsf{B}_{[3]}.$

Наконец, выполнение оператора, правая часть которого содержит операцию над двумя массивами, сводится к поэлементной операции между парами соответствующих элементов двух массивов. Например, X3 = A / B означает $X3_{[1]} := A_{[1]} / B_{[1]}$, $X3_{[2]} := A_{[2]} / B_{[2]}$, $X3_{[3]} := A_{[3]} / B_{[3]}$ (при r = 3).

Пример языковой программы:

$$X1 := (+00001, -00010, -00156); A1 := X1 * (+00005, -12079, +44172);$$

 $X1 := A1 - X1;$

Дополнительные правила:

- 3) внутри круглых скобок ровно r-1 запятая;
- и) внутри круглых скобок могут быть только цифры, запятая и знаки «плюс» и «минус».
 - 1.4 Язык операций над элементами массивов

$$<$$
 символ $>$: $=$ $<$ буква $>$ $|$ $<$ цифра $>$ $|$ $<$ знак $>$ $|$ $<$ разделитель $>$

$$<$$
 3HaK $>$: := $\lor |\lor|$ *

$$<$$
 разделитель $>$: := ; | : | = | (|)| | , | -

$$<$$
 массив $>$: $=$ ($<$ число $>$, . . . , $<$ число $>$).

Количество r чисел в массиве фиксируется (r = 3, 4, или 5).

< оператор> : := < иден1 > := < массив > | < иден2> := < иден1> < знак > < иден1>

Bыполнение операторов. Идентификатор из буквы и цифры обозначает массив из r чисел. Первая форма оператора — это присвоение элементам массива конкретных чисел. Например,

$$X8 := (-36001, +00102, -03041)$$

означает присвоение X8 [1] := -35001, X8 [2] := +00102, X8 [3] := -03041.

Идентификатор их двух букв обозначает простую переменную (ему присваивается одно число). Вторая форма оператора выполняется так: вычисляется число по элементам двух массивов и присваивает его простой переменной. Операции задаются следующим образом:

$$X1 \vee Y1 = \sum_{i=1}^{r} (X1[i] - Y1[i])^2$$
 - квадрат расстояния между векторами;

 $X1 \wedge Y1 = \max_i (|X1[i] - Y1[i]|)$ - максимальное отклонение между векторами;

X1 * Y1 =
$$\sum_{i=1}^{n} X1[i] \times Y1[i]$$
 - скалярное произведение векторов.

Пример языковой программы:

$$C2 := (+01020, -00108, -00079); X8 := (-35001, +00102, -03041);$$

$$AB := C2 * X8; BA := X8 \wedge C2; XY := X8 \vee C2;$$

Дополнительные правила:

- з) правая часть оператора либо не содержит идентификаторов, либо содержит 2 в виде комбинации буквы и цифры;
- и) внутри круглых скобок точно r чисел, отделенных друг от друга запятыми.
 - 1.5 Язык операций над элементами булевских массивов

$$<$$
 символ $> ::= <$ буква $> |<$ код $> |<$ знак $> |<$ разделитель $> |<$ цифра $>$

< 3Hak>::=!|%|◊

< разделитель > : := ; | : | =

< код> – последовательность из 8 нулей и единиц

```
< иден1> : := < буква > < буква > < иден2 > : := < буква> < цифра > < оператор > : := < иден1> := < код > | < иден2 > := < знак> < код > | < иден2> := < знак> < иден1>
```

Выполнение операторов. Идентификатор из пары букв обозначает массив из 8 булевских значений (1 и 0). Оператор первого вида — это присвоение идентификатору конкретного набора нулей и единиц. Например,

$$AY := 10010110.$$

Идентификатор из буквы и цифры обозначает простую переменную, которой присваивается целое число. Вторая и третья формы оператора требует вычисления значения (целого числа) по коду, задаваемому явно или идентификатором, и присвоения этого значения левой части оператора. Например, X1 := !10100100; X2 := %10001001; или $X3 := \lozenge$ АҮ. Операции трактуются следующим образом:

- ! количество единиц в коде;
- % разность между числом единиц и числом нулей в коде;
- \Diamond целое число, двоичное представление которого образует код.

Дополнительные правила:

- з) после знака должен быть код из восьми нулей и единиц или пара букв, после которых знак «точка с запятой»;
 - и) в левой части каждого оператора буква и цифра, либо 2 буквы.
 - 1.6 Язык условных выражений

```
< символ > : := < буква >| < цифра >| < знак >| < разделитель >
< знак> : := + | - | * | /
< разделитель > : := ; | : | = | < | | [ | ]
< перем> : := < иден >| < число >
< оператор > : := < иден> := < условие > [ < параметр >] [< параметр>]
< условие > : := < перем> < < перем> | < перем> = < перем>
< параметр >: := < перем> < знак> < перем>
```

Выполнение операторов. Вычисляется <условие>; если оно истинно, то идентификатору левой части присваивается значение первого параметра, в противном случае — значение второго параметра из правой части. Знаки арифметических операций понимаются обычным образом (но деление нацело).

Пример языковой программы:

$$X := +00100 < A1 [-10203] [X+-35010]; Y := A=B [X+Y] [X/Y];$$

Дополнительные правила:

- з) в правой части оператора должен быть либо символ < , либо
- = ; и) в правой части оператора две пары квадратных скобок.
- 1.7 Язык арифметических выражений

```
< символ > : := < буква >| < цифра >| < знак >| < разделитель >
< знак> : := + | - | * | /
< разделитель > : := ; | : | = | | <
< перем> : := < иден >| < число >
< выражение > : := < перем> | < выражение > < знак > < перем>
< оператор > : := < иден> := < выражение>
```

Выражение оператора. Выражение представляет собой алгебраическое выражение из идентификаторов, чисел и разделяющих их знаков арифметических операций. Количество знаков арифметических операций в выражении ограничивается величиной р (например, p = 4, 5 или 6).

Пример языковой программы:

$$X := -01096$$
; $Y := +00785$; $A := +00002 + X / -00012*Y$.

Дополнительные правила:

- з) после разделителя = следует либо буква, либо знак + или -;
- и) в правой части оператора не более 2p + k < 3наков>.
- к) Арифметические операции выполняются последовательно слева направо без учета приоритетности. Например, выражение X + Y 2*X равносильно в скобочной записи (((X + Y) X). Вычисленное значение выражения присваивается идентификатору левой части.
- л) Арифметические операции выполняются последовательно справа налево без учета приоритетности. Например, выражение X + Y 2*X равносильно в скобочной записи (X + (Y (2*X))). Вычисленное значение выражения присваивается идентификатору левой части.
- м) Арифметические операции выполняются с учетом приоритетности операций. Вычисленное значение выражения присваивается идентификатору левой части.

1.8 Язык булевских выражений

```
< символ > : := < буква > | < код > | < цифра > | < знак> | < разделитель> < код > — последовательность из восьми нулей и единиц < знак> : := < | < | + | \rightarrow < разделитель > : := ; | : | | = < перем> : := < иден > | < код > < выражение > : := < перем> | < выражение > < знак > < перем>
```

Выполнение оператора. Выражение состоит из идентификаторов, кодов и разделяющих их знаков булевских операций. Операции – поразрядные – выполняются в выражении слева направо (другой вариант задания – справа налево). Булевские операции обозначены: " \checkmark " – дизъюнкция, " \land " – конъюнкция, "+" – сложение по модулю 2, " \rightarrow " – импликация. Вычисленное значение выражения (восьмиразрядный код) присваивается идентификатору левой части оператора. Количество знаков операций выражения ограничено величиной р (например, р = 4, 5 или 6).

Пример языковой программы:

```
A := 10011001; B := 01110111-A + 10110011; A := B \land A \lor B;
Дополнительные правила:
```

- и) каждый код состоит точно из 8 нулей и единиц; количество знаков операций в правой части оператора не ограничено, операции выполняются слева направо без учета приоритетности.
- к) каждый код состоит точно из 8 нулей и единиц; количество знаков операций в правой части оператора не более р, операции выполняются справа налево без учета приоритетности.
- л) каждый код состоит точно из г нулей и единиц (определяется константой); количество знаков операций в правой части оператора не ограничено, операции выполняются слева направо с учетом приоритетности (сначала выполняется операция ∧, потом все остальные операции в порядке их следования).

1.9. Язык арифметических выражений

```
< символ > : = < буква > | < цифра > | < знак > | < разделитель >
```

Выполнение оператора. Выражение представляет собой алгебраическое выражение из идентификаторов, чисел и разделяющих их знаков арифметических операций.

Пример языковой программы:

$$X := -01096$$
; $Y := +00785$; $A := +00002 ++$; $X := A--$.

1.10 Язык условных выражений

Выполнение операторов. Вычисляется <условие>; если оно истинно, то идентификатору левой части присваивается значение первого выражения, в противном случае — значение второго выражения из правой части оператора. Знаки арифметических операций понимаются обычным образом (целочисленные операции).

Пример языковой программы:

$$X = +00100 < A1 ?-10203: X + -35010; Y = A == B? X + Y:X/Y;$$

 $MIN = A < B ? A : B;$

1.11 Язык операций над массивами

```
< массив >: = < иден > | (< число >, . . . , < число > )
```

Примечание. Количество r чисел в массиве фиксируется в языке (r=3, 4 или 5).

$$<$$
 выражение $>$: $=$ $<$ массив $>$ $<$ знак $>$ $<$ выражение $>$

Bыполнение операторов: идентификатор обозначает массив из r чисел. Оператор вида

$$<$$
 иден $> := (<$ число $>, ..., <$ число $>)$

Например, X:=(3,8,-19) означает серию операторов присваивания $X_{[1]}:=3,$ $X_{[2]}:=8,$ $X_{[3]}:=-19.$ Выполнение оператора вида

сводится к копированию массива и выполняется путем присваивания i - му элементу массива i - го числа.

. Например,
$$A1 := \mathbf{F}$$
 означает $A1_{[1]} := \mathbf{F}_{[1]}$; $A1_{[2]} := \mathbf{F}_{[2]}$; $A1_{[3]} := \mathbf{F}_{[3]}$.

Наконец, выполнение оператора, правая часть которого содержит операцию над двумя массивами, сводится к поэлементной операции между парами соответствующих элементов двух массивов. Например, X3 = A / B означает $X3_{[1]} := A_{[1]} / B_{[1]}$, $X3_{[2]} := A_{[2]} / B_{[2]}$, $X3_{[3]} := A_{[3]} / B_{[3]}$ (при r = 3).

Пример языковой программы:

$$X1 := (+00001, -00010, -00156); A1 := X1 * (+00005, -12079, +44172);$$

 $X1 := A1 - X1 + A1;$

Дополнительные правила:

- з) Операции выполняются слева направо без учета приоритетности.
- и) Операции выполняются справа налево без учета приоритетности.
- К) Операции выполняются с учетом приоритетности.
- 1.12 Язык операций над массивами

$$<$$
 символ $>$: $=$ $<$ буква $>$ $|$ $<$ цифра $>$ $|$ $<$ знак $>$ $|$ $<$ разделитель $>$

$$< 3HaK > : := + | - | * | /$$

$$<$$
 разделитель $>$: $=$; $|:|=|(|)||,$

Примечание. Количество r чисел в массиве фиксируется в языке (r=3, 4 или 5).

< выражение> : := < массив > | (< выражение> < знак > < выражение>)

Bыполнение операторов: идентификатор обозначает массив из r чисел. Оператор вида

Например, X:=(3,8,-19) означает серию операторов присваивания $X_{[1]}:=3,$ $X_{[2]}:=8,$ $X_{[3]}:=-19.$ Выполнение оператора вида

сводится к копированию массива и выполняется путем присваивания i - му элементу массива i - го числа.

. Например,
$$A1 := Б$$
 означает $A1_{[1]} := Б_{[1]}; A1_{[2]} := Б_{[2]}; A1_{[3]} := Б_{[3]}.$

Наконец, выполнение оператора, правая часть которого содержит операцию над двумя выражениями, сводится к вычислению каждого их выражений и затем поэлементной операции между парами соответствующих результатов. Скобки указывают на приоритет выполнения операций. Например, X3 = (A/(5+A)) означает $X3_{[1]} := A_{[1]}/(5_{[1]+}A_{[1]})$, $X3_{[2]} := A_{[2]}/(5_{[2]+}A_{[2]})$, , $A_{[3]}/(5_{[31]+}A_{[1]})$, . Все операции целочисленные.

Пример языковой программы:

$$X1 := (+00001, -00010, -00156);$$
 $A1 := ((X1 * (+00005, -12079, +44172))*X1);$ $X1 := (A1 - (X1+A1));$

Дополнительные правила:

1.13 Язык операций над массивами

```
< символ > : := < буква > | < цифра > | < знак > | < разделитель > < знак > : := + | - | * | / < разделитель > : := ; | : | = | (|) | | ,
```

< число >: := < знак > | (< цифра >, . . . , < цифра >)

Примечание. Количество цифр в числе равно фиксируется в языке (3, 4 или 5).

```
< оператор> : := < иден > := < выражение > 
< выражение > : := <число > | (< выражение> < знак > < выражение>) 
Выполнение операторов: < иден > := (< число >, . . . , < число >)
```

Например, X := +3 означает оператор присваивания. Выполнение оператора < иден > := (< выражение> < знак > < выражение>)

сводится вычислению выражения в скобках и присваиванию переменной полученного результата.

Пример языковой программы:

$$X1 := +00001$$
; $A1 := ((X1 * +00005))*X1)$; $X1 := (A1 - (X1 + A1))$;

Дополнительные правила:

3) количество знаков операций в правой части оператора присваивания не более r (r определяется в программе).

1.14 Язык булевских выражений

```
< символ > : = < буква > | < код > | < цифра > | < знак> | < разделитель>
```

< код > – последовательность из восьми нулей и единиц

```
< 3HaK> : := \lor | \land | + | \rightarrow
```

< разделитель > : := ; | : | | =

< перем> : := < иден > | < код >

< выражение > : = < перем> | (< выражение > < знак > <выражение>)

Выполнение оператора. Булевские операции обозначены: " \checkmark " – дизъюнкция, " \land " – конъюнкция, "+" – сложение по модулю 2, " \rightarrow " – импликация. Вычисленное значение выражения (восьмиразрядный код) присваивается идентификатору левой части оператора. Количество знаков операций выражения ограничено величиной р (например, р = 4, 5 или 6).

Пример языковой программы:

$$A := 10011001$$
; $B := (01110111 + 10110011)$; $A := B \land (A \lor B)$;

Дополнительные правила:

з) каждый код состоит точно из 8 нулей и единиц; количество г знаков операций в правой части оператора ограничено (определяется в программе), операции выполняются слева направо без учета приоритетности.