Лексические соглашения языка МИКРОЛИСП.

Программа записывается в форме текстового файла. Текст состоит из *атомов*, которые отделяются друг от друга круглыми *скобками* и/или *пробельными* символами. В тексте программы по правилам диалекта Scheme могут быть записаны *комментарии*. Комментарий эквивалентен пробелу.

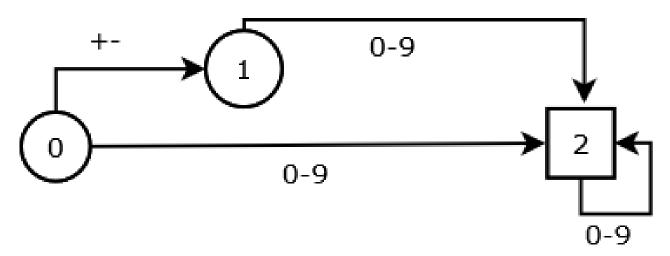
Приведенные выше соглашения позволяют определить лексические классы – токены первого уровня (в левой колонке указано символическое обозначение токена).

```
знак (
 )
                знак )
 #
                обозначает маркер конца текста
                программы (явно НЕ ЗАПИСЫВАЕТСЯ).
 $str
                строка. Атомы этого типа можно не
                отделять пробелами от соседних атомов,
                например, 5"abc"6"def""ghi"
 ?
                ошибка - неопознанный токен
Атомы(кроме строк) подразделяются на токены второго
уровня:
 $int
                целое
 $dec
                вещественное
```

```
$dec вещественное
$id идентификатор
$idq идентификатор предиката
$bool константы #t и #f
операторы:
```

```
+ знак +
- знак -
* знак *
/ знак /
< знак <
= знак =
> знак >
<= цепочка <=
>= цепочка >=
```

Токен \$int формируют целые числа.



Токен \$dec формируют вещественные числа. При этом цепочки, входящие в токен \$int, токену \$dec не принадлежат.

Варианты шаблонов токена \$dec приведены в задании лабораторной работы №4.

Токен \$id формируется по специальным правилам, существенно ограниченным, по сравнению с диалектом Scheme.

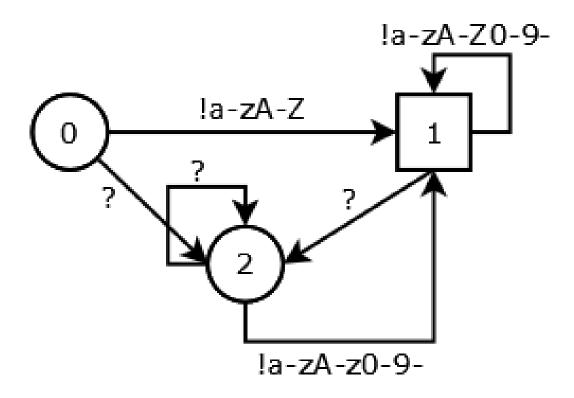
Идентификатор состоит из латинских букв, цифр, знаков

'-'(дефис), '!' и '?'. Он не может **Начинаться** с цифры и/или знака '-', а также не может

ОКАНЧИВАТЬСЯ знаком '?'.

На токен накладываются дополнительные ограничения, запрещающие использовать некоторые комбинации символов.

Токен \$id без дополнительных ограничений:



(c) Golikov K.P. 2021

Идентификатор предиката \$idq состоит из латинских букв, цифр, знаков '-'(дефис), '!' и '?'. Он не может начинаться с цифры и **ДОЛЖЕН**

ОКАНЧИВАТЬСЯ знаком \?'.

Идентификатор предиката **МОЖЕТ** начинаться со знака '-'.

На токен накладываются дополнительные ограничения, запрещающие использовать некоторые комбинации символов.

Варианты дополнительных ограничений для \$id и \$idq приведены в задании лабораторной работы №5.