

## Формальные языки

**Задача 1.** Найти описание лексического синтаксиса вашего второго самого любимого языка программирования. В нем найти описание следующих языков:

Для сокращения записи, определим следующие синтаксические конструкции для регулярных выражений:

$$A+ \stackrel{\text{def}}{=} AA^*$$

$$A? \stackrel{\text{def}}{=} A|\epsilon$$

$$[a_1 - a_n] \stackrel{\text{def}}{=} a_1|a_2|\dots|a_n$$

Если мы захотим использовать один из специальных символов '+', '\*', '?', '|', '(', ')', '[', ']' в качестве одного из символов нашего языка, то мы будем экранировать этот символ обратным слешем: '\+', '\\*', ...

- язык идентификаторов;

Описание: <https://doc.rust-lang.org/grammar.html#identifiers>

Регулярное выражение:  $([a-z][A-Z]|\_)([0-9][a-z][A-Z]|\_)^*$

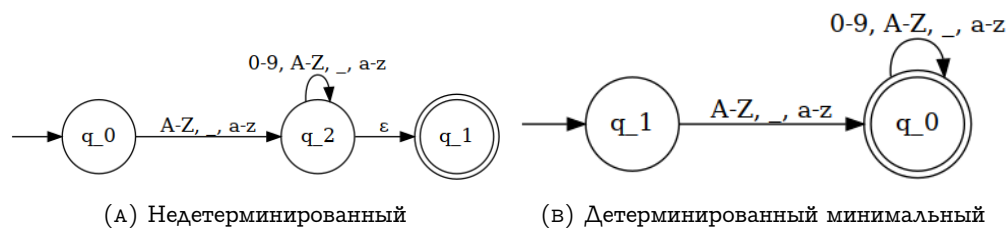
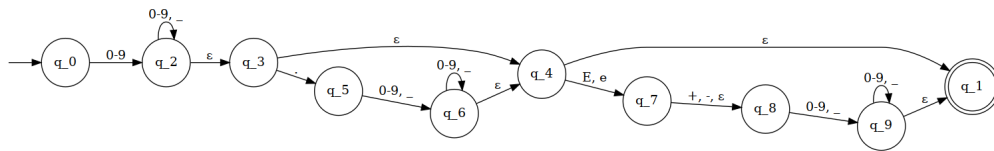
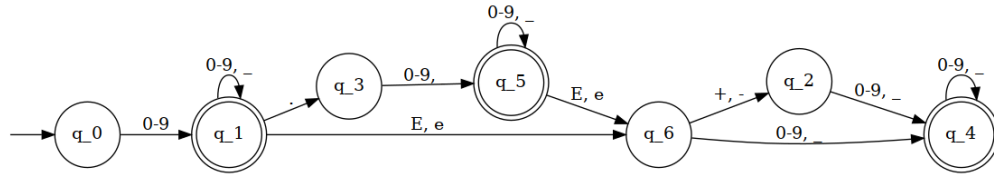


Рис. 1. Конечные автоматы для языка идентификаторов

- язык каких-нибудь чисел (целых в десятичной счисления, с плавающей точкой, любых других).  
Описание: <https://doc.rust-lang.org/grammar.html#number-literals>  
Опишем только десятичные числа с плавающей точкой (прим. 42, 12.34, 1.79E2, 100\_000\_000 и т.д.)  
Регулярное выражение:  $[0-9]([0-9]|\_)*\cdot([0-9]|\_+)?((e|E)(-|\+)?([0-9]|\_+))?$



(А) Недетерминированный



(В) Детерминированный минимальный

Рис. 2. Конечные автоматы для языка идентификаторов

- язык ключевых слов (можно какого-нибудь их конечного подмножества);  
Описание: <https://doc.rust-lang.org/grammar.html#keywords>  
Регулярное выражение:

```
_|abstract|alignof|as|become|box|  
break|const|continue|crate|  
do|else|enum|extern|false|  
final|fn|for|if|impl|  
in|let|loop|macro|match|  
mod|move|mut|offsetof|override|  
priv|proc|pub|pure|ref|  
return|Self|self|sizeof|static|  
struct|super|trait|true|type|  
typeof|unsafe|unsized|use|virtual|  
where|while|yield
```

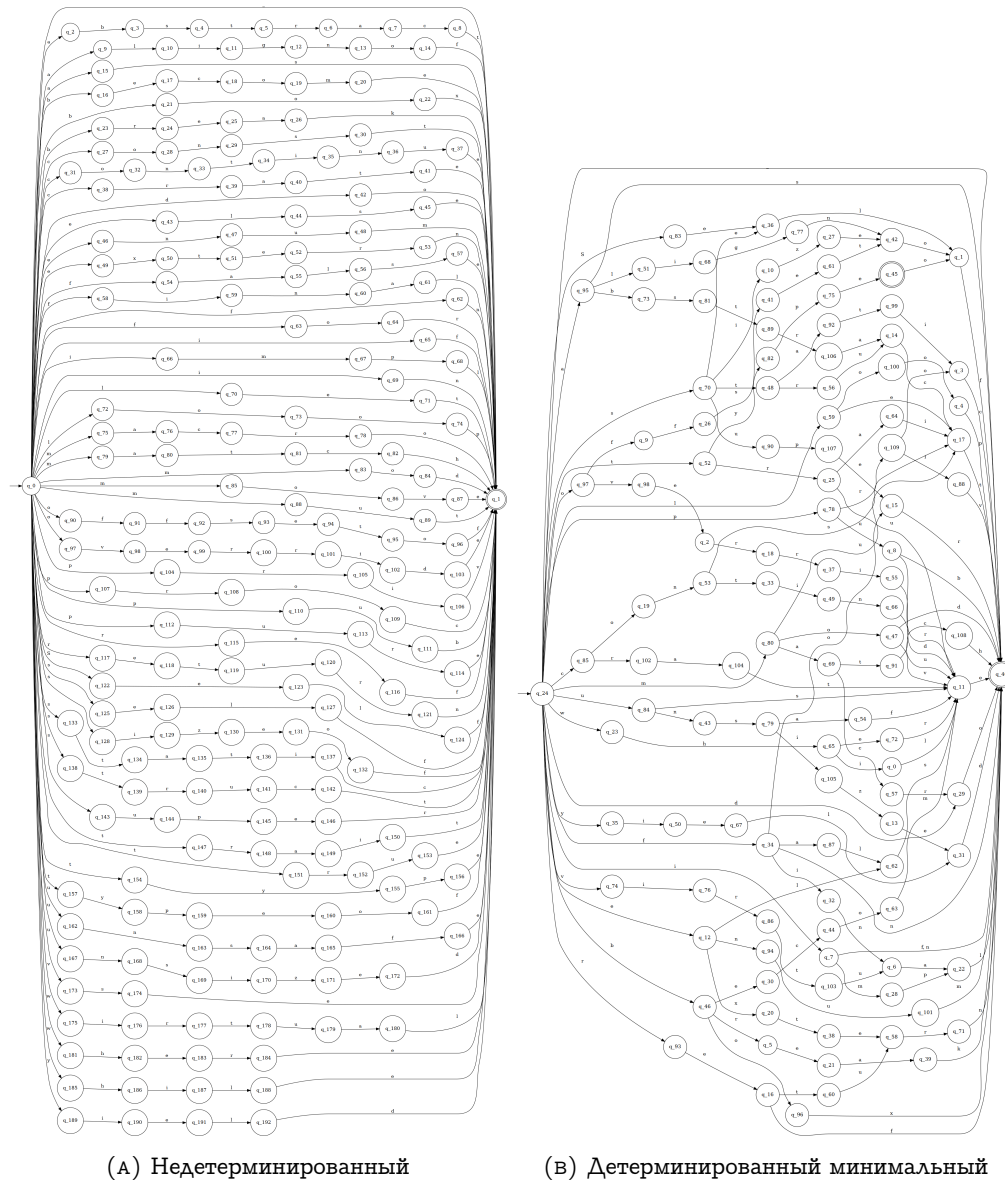


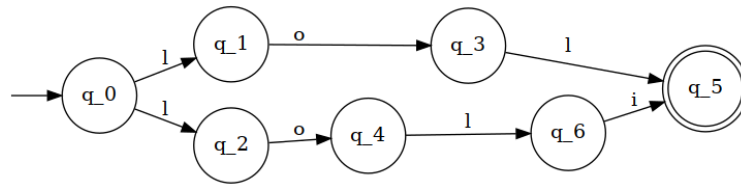
Рис. 3. Конечные автоматы для языка ключевых слов

**Задача 2.** Реализовать алгоритм Томпсона детерминизации конечных автоматов.

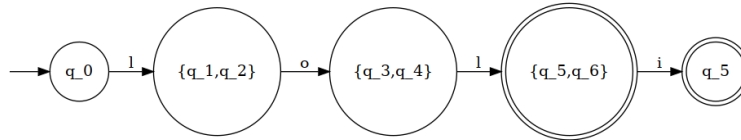
Программная реализация:

<https://github.com/bagar0x60/automata-hw>

Примеры работы скрипта:

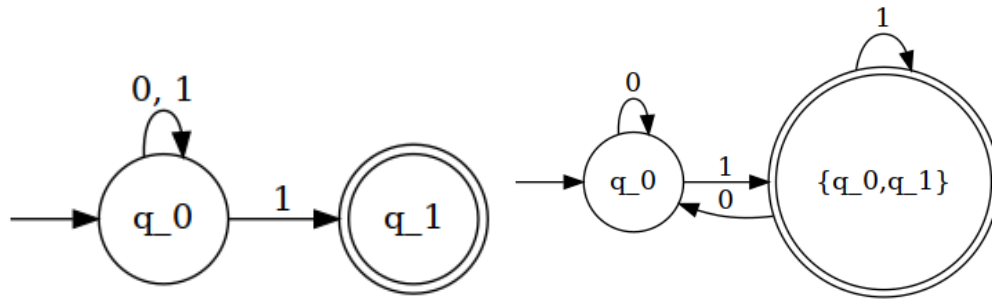


(А) Исходный



(B) Детерминированный

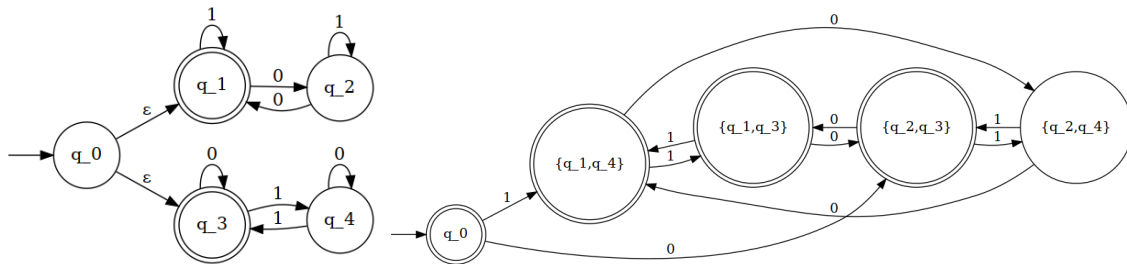
Рис. 4. Пример детерминизации 1



(А) Исходный

(B) Детерминированный

Рис. 5. Пример детерминизации 2



(А) Исходный

(B) Детерминированный

Рис. 6. Пример детерминизации 2

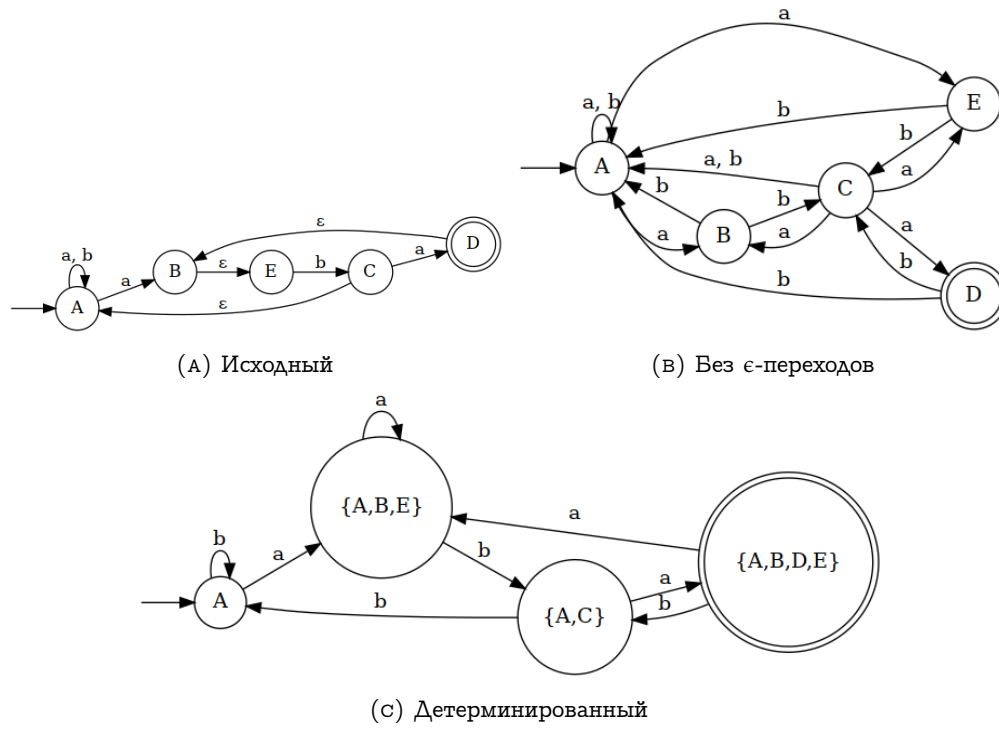


Рис. 7. Пример детерминизации 3