Определение. Конечное, непустое множество Σ — **алфавит**

Пример.

- $\{0, 1\}$
- $\{a, b, c \dots z\}$
- Unicode

$$\Sigma^* := \bigcup_{k=0}^{\infty} \Sigma^k$$

Определение. Конкатенация:

$$\Sigma^k \times \Sigma^l \to \Sigma^{k+l}$$

Это аддитивная операция с нейтральным элементом.

Определение. Язык над алфавитом Σ^* — подмножество Σ^*

Σ^* — бесконечное счётное множество

Множество всех языков 2^{Σ^*}, т.к. каждый из элементов Σ^* либо включен, либо нет. Это множество несчетно.

Пример.

1. $A = \{w |$ в w четное число нулей$\} \subset \{0, 1\}^*$

   $01011 \in A \quad 000 \notin A$

2. $Pal = \{w | w -$ палиндром$\} \subset \{0, 1\}^*$

   $010 \in Pal \quad 0000 \in Pal \quad 01 \notin Pal$

Языки надо задавать формально, а не что-то рода *"язык палиндромов"*. Есть два способа это делать:

1. **Распознавание**: есть черный ящик, который на вход получает слово и выдает булево значение — принадлежит слово искомому языку или нет.
2. **Конструирование**: система правил диктует то, как устроены слова в искомом языке.

Пример. Распознавание ПСП:

Тут код.

Пример. Конструирование ПСП:

ε — ПСП

A — ПСП $\Rightarrow (A)$ — ПСП

A, B — ПСП $\Rightarrow AB$ — ПСП

Автоматы распознают принадлежность слова языку.

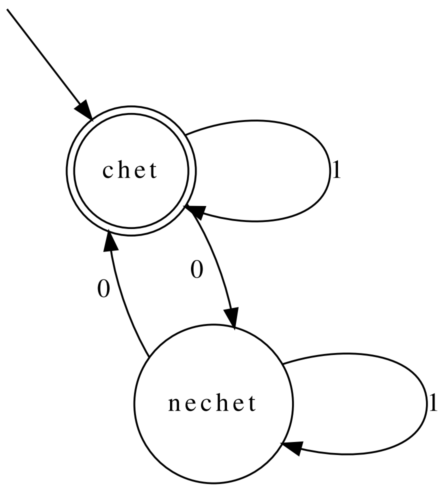

Рис. 1: Автомат для $\{w |$ в w четное число нулей$\}$

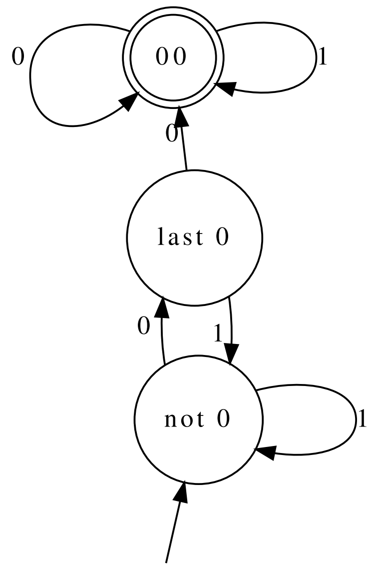

Рис. 2: Автомат $\{w |$ в w содержит два нуля подряд$\}$

Определение. **Детерменированный конечный автомат *(ДКА)***:

1. Состояния, обозначаемые кругами
2. Переходы, обозначаемые ребрами между состояниями и помечаемые символом алфавита. Из любого состояния есть ровно один переход по каждому символу алфавита.
3. Начальное состояние, обозначаемое входящей стрелкой из никуда
4. Допускающие состояния, обозначаемые кругом внутри себя.

Переименуем состояния “not 0” $\to A$, “last 0” $\to B$, “00” $\to C$.

Опишем математически ДКА:

Определение. ДКА $AV = \langle \Sigma, Q, S \in Q, T \subset Q, \delta : Q \times \Sigma \to Q \rangle$

Мгновенное описание AV — пара из состояния A и оставшихся от исходного слова строки. Это объект из множества $Q \times \Sigma^*$, например $\langle A, 11011001 \rangle$.

На мгновенных описаниях можно задать отношение “переходит за 1 шаг”, обозначим его $\vdash$. Формально оно задается следующим образом:

$\langle p, \alpha \rangle \vdash \langle q, \beta \rangle$, если:

1. $\alpha = c\beta$

2. $q = \delta(p, c)$

Тогда для для слова S верно следующее:

$$\langle S, x\rangle \vdash \langle u_1, x_1\rangle \vdash \langle u_2, x_2\rangle \vdash \ldots \vdash \langle u_l, \varepsilon\rangle$$

$$AV \text{ допускает } x \Leftrightarrow u_L \in T$$

Эта запись неудобна, поэтому обозначим транзитивное замыкание $\vdash$ как $\vdash^*$, это — отношение “переходит ≥ 0 шагов”. Тогда получается следующее:

$$AV \text{ допускает } x \Leftrightarrow \langle s, x\rangle \vdash^* \langle t, \varepsilon\rangle$$

Заметим, что все ДКА — счётное множество, а языки - несчётное $\Rightarrow$ не любой язык можно описать с помощью ДКА. Назовем все языки, которые можно получить с помощью ДКА **автоматными**.

Определение. Базовые регулярные языки:

- $\varnothing$
- $\{\varepsilon\}$
- $\{c_i\} \quad \Sigma = \{c_1 \ldots c_z\}$

Обозначим множество базовых регулярных языков над $\Sigma = \{0, 1\}$ как $R_0 = \{\{\}, \{\varepsilon\}, \{0\}, \{1\}\}$

Зададим три операции над этими языками:

1. Объединение: $A \cup B$
2. Конкатенация: $\{ab \mid a \in A, b \in B\}$
3. Замыкание Клини: $A^* = \bigcup\limits_{i=0}^{\infty} A^i$

$$R_1 := \{M \mid M \in R_0 \text{ или } M = AB, A \in R_0, B \in R_0, M = A{\cap}B, A \in R_0, B \in R_0 \text{ или } M = A^*, A \in R_0\}$$

$$R_1 = \{\{\}, \{\varepsilon\}, \{0\}, \{1\}, \{\varepsilon, 0\}, \{\varepsilon, 1\}, \{00\}, \{01\}, \{10\}, \{11\}, \{0\}^*, \{1\}^*\}$$

Определение. Регулярные языки:

$$Reg = \bigcup_{i=0}^{\infty} R_i$$