Definition

An alphabet is a finite set Σ of symbols. A word is a finite sequence of symbols, including the empty sequence which is denoted by λ. The set of words over Σ is denoted by Σ^*. The length of w, denoted as $|w|$, is the number of the symbols in it. The concatenation of two words is denoted by writing them next to each other.

Definition (Language)

A (formal) language over Σ is a subset of Σ^*.

Example: The set of all grammatically correct Polish sentences is a language over the Polish alphabet extended with . , ! etc.

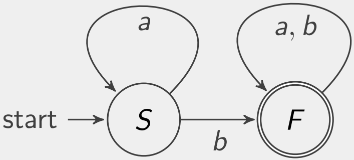
Definition (Deterministic Finite Automaton, *DFA*)

A deterministic finite automaton over the finite alphabet Σ is a tuple $\mathcal{A} = \langle Q, S, F, \delta \rangle$, where Q is the nonempty finite set of states, $S \in Q$ is the starting state, $F \subseteq Q$ is the set of halting states, and $\delta\ Q \times \Sigma \to Q$ is the transition function.

Any word $w \in \Sigma^*$ acts on $\mathcal{A}$ as follows. Start from the initial state S, and apply the transition function to the actual state on the next symbol (from left to right) of w to get the next state. The word w is accepted by $\mathcal{A}$ if w moves the automaton into a final state (i.e. an element of F). The language generated or accepted by $\mathcal{A}$ is

$$L(\mathcal{A}) = \{w \in \Sigma^* \ : \ w \text{ is accepted by } \mathcal{A}\}.$$

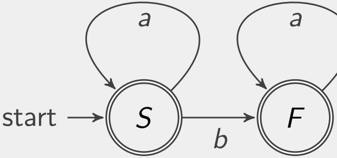
Examples



Accepts the language

$$a^* b(a|b)^* = \text{the set of strings that have at least one } b$$

Examples

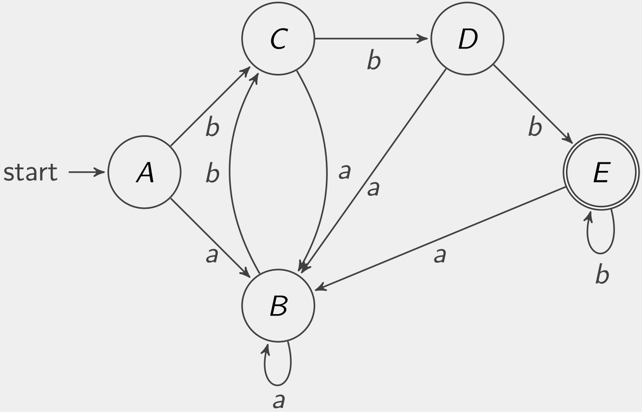


Accepts the language

$$a^*|a^*ba^* = \text{the set of strings that have at most one } b$$

Problem: where to transit from F with input b?

Examples

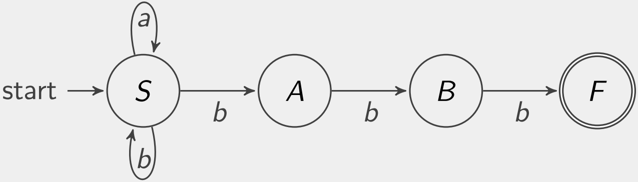


Accepts the language

$$(a|b)^* bbb = \text{The set of strings that end in 3 consecutive } b\text{'s}$$

Examples

Detour: Non-deterministic automaton



Accepts the language

$$(a|b)^*bbb = \text{The set of strings that end in 3 consecutive } b\text{'s}$$