Alphabet $\to$ finite set

$A = \{a, b\}$ $\quad\quad \{a, b, c\}$ $\quad\quad \{0, 1\}$

Elements of the alphabet $\boxed{\text{letters}}$

For $A = \{a, b\}$ the letters are a and b

A word is just a sequence of letters

$u = ababba$

$u$ is a word of length 6 (6 letters)

$A = \{a, b, c\}$

$v = cabaabc$

Product (concatenation) of two words

$u = abba$ $\quad v = bbaab$ $\quad uv = abbabbaab$

This product is associative

abra (cad abra) = (abra cad) abra = abracadabra

Is this product commutative ?

$$uv \overset{??}{=} vu$$

$u = ab$ $\quad v = ba$ $\quad uv = abba$ $\quad vu = baab$

Special word : empty word (no letter)

Notation $1$

$$\boxed{1u = u = u1}$$

$1$ is the neutral element for the product

Length of a word $u$ = number of letters

Notation $|u|$

$|abracadabra| = 11$ $\qquad |a_1 \cdots a_n| = n$

$|1| = 0$ $\qquad$ The empty word is the unique word of length 0

Powers

$u^1 = u \qquad u^2 = uu \qquad u^3 = uuu \quad \cdots$

Convention $\boxed{u^0 = 1}$ $\qquad u^{n+1} = u^n u = u u^n$

Exercise $\quad u^p u^q = u^{p+q}$

<u>Fact</u> $\quad |uv| = |u| + |v|$

$|u^n| = n|u|$

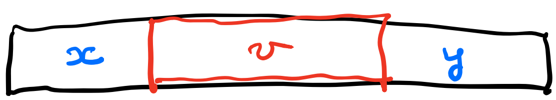
Reversal of a word $\quad u = a_1 \cdots a_n$

$\mapsto \quad \tilde{u} = a_n \cdots a_1$

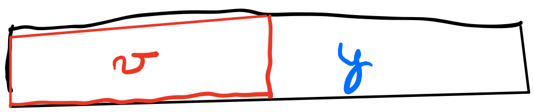
If $u = abca \qquad \tilde{u} = acba$

Factors, prefixes and suffixes of a word $u$

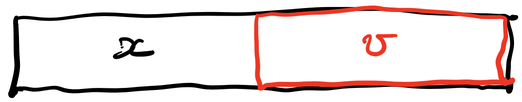
$v$ is a $\boxed{\text{factor}}$ (or infix) of $u$ if there exist some words $x$ and $y$ such that $u = xvy$

$u=$ 


$v$ is a $\boxed{\text{prefix}}$ of $u$ if there exists some word $y$ such that $u = vy$ $\qquad$ (left factor)

$u=$ 


$v$ is a $\boxed{\text{suffix}}$ of $u$ if there exists some word $x$ such that $u = xv$ $\qquad$ (right factor)

$u =$ 


Exercise let $u = abaab$

Give some factors of $u$

$1$, $ba$, $aa$, $baa$, $baab$, $abaab$, ...

Some prefixes of $u$

$aba$, $a$, $ab$, $1$

Some suffixes of $u$

$ab$, $baab$

Languages. A language is a set of words

If $A = \{a, b\}$

$$L_1 = \{aba, babaa\} \leftarrow \text{finite language}$$

$$L_2 = \{a^n b a^n \mid n \geq 0\}$$

$$= \{b, aba, a^2ba^2, \dots\}$$

Operations on languages

$\boxed{\text{Boolean operations}}$ : union, intersection, complement, difference

$$L_1 \cup L_2 = \{u \in A^* \mid u \in L_1 \text{ or } u \in L_2\}$$

$$L_1 \cap L_2 = \{u \in A^* \mid u \in L_1 \text{ and } u \in L_2\}$$

complement of $L$ $$L^c = \{u \in A^* \mid u \notin L\}$$

$$L_1 \setminus L_2 = \{u \in A^* \mid u \in L_1 \text{ and } u \notin L_2\}$$

Product of two languages

$$L_1 L_2 = \{ u_1 u_2 \mid u_1 \in L_1, u_2 \in L_2 \}$$

$$\{ab, aba\}\{1, a\} = \{ab, aba, abaa\}$$

The product of languages is associative

$$(L_1 L_2) L_3 = L_1 (L_2 L_3)$$

Is there a neutral element (identity) for the product of languages? Is there a language $E$ such that, for all languages $L$

$$LE = EL = L \quad ?$$

$E = \{1\}$ is the unique solution.

The product of languages is not commutative

$$\{a\}\{b\} = \{ab\} \qquad \{b\}\{a\} = \{ba\}$$

The product is distributive over union

$$L(L_1 \cup L_2) = LL_1 \cup LL_2$$

$$(L_1 \cup L_2) L = L_1 L \cup L_2 L$$

What happens with the intersection ?

$$\{b\}\{a, aa\} = \{ba, baa\}$$

$$\{ba\}\{a, aa\} = \{baa, baaa\}$$

$$\{b\}\{a, aa\} \cap \{ba\}\{a, aa\} = \{baa\}$$

$$(\underbrace{\{b\} \cap \{ba\}}_{\emptyset})\{a, aa\} = \emptyset$$

$$\emptyset L = \{u_1 u_2 \mid u_1 \in \emptyset,\ u_2 \in L\} = \emptyset$$

The product is not distributive over intersection

### Powers of a language

$$L^1 = L \qquad L^2 = LL \qquad L^3 = LLL \quad \cdots$$

Convention $\quad L^0 = \{1\} \qquad\qquad L^{n+1} = L^n L$

<u>Example</u> $\qquad L = \{a, ab, ba\}$

$L^2 = \{aa, aab, aba, baa, abab, abba, baab, baba\}$

$L^3 =$

### Star of a language

$$L^* = \bigcup_{n \geq 0} L^n = L^0 \cup L^1 \cup L^2 \cup \cdots$$
$$= \{1\} \cup L \cup L^2 \cup \cdots$$

$$L^+ = \bigcup_{n > 0} L^n = L \cup L^2 \cup L^3 \cup \cdots$$

Example $\quad L = \{a, ab\}$

$L^* = \{1, a, aa, ab, aaa, aab, aba, aaaa, aaab, \ldots\}$

Important, but confusing formulas

$$\emptyset^* = \{1\} \qquad\qquad \emptyset^+ = \emptyset$$
$$\{1\}^* = \{1\} \qquad\qquad \{1\}^+ = \{1\}$$

### Rational languages

(regular languages)

Definition $\quad$ The set of rational languages of $A^*$ is the smallest set $Rat(A^*)$ of languages such is the smallest set Rat(A*) of languages such that:

(1) $\emptyset$ is a rational language

(2) For each letter $a \in A$, $\{a\}$ is a rational language

(3) Rational languages are closed under the operations of finite union, product and star

$$L_1, L_2 \in Rat(A^*) \Rightarrow \begin{array}{l} L_1 \cup L_2 \in Rat(A^*) \\ L_1 L_2 \in Rat(A^*) \end{array}$$

$$L \in Rat(A^*) \Rightarrow L^* \in Rat(A^*)$$

<u>Prop</u> Every finite language is rational.

$\{a_1 a_2 \cdots a_n\}$ $\qquad\qquad$ $\{abac\}$

$$\{abac\} = \{a\}\{b\}\{a\}\{c\}$$

are rational languages

$$\{a_1 a_2 \cdots a_n\} = \{a_1\}\{a_2\} \cdots \{a_n\}$$

If $L$ is finite, say $L = \{u_1, u_2, \dots, u_k\}$ where $u_1, \dots, u_k \in A^*$

$$L = \{u_1\} \cup \{u_2\} \cup \cdots \cup \{u_k\}$$

Examples

(1) $\{a, ab, ba\}^*$ is rational

(2) $A^* aba A^*$ "

$= \{u \in A^*$ such that $aba$ is a factor of $u\}$

(3) $ab A^* = \{u \in A^* \mid ab$ is a prefix of $u\}$

(4) Set of words of even length is rational

pair

$$= \{ u \in A^* \mid |u| = 2n \text{ for some } n \geq 0 \}$$

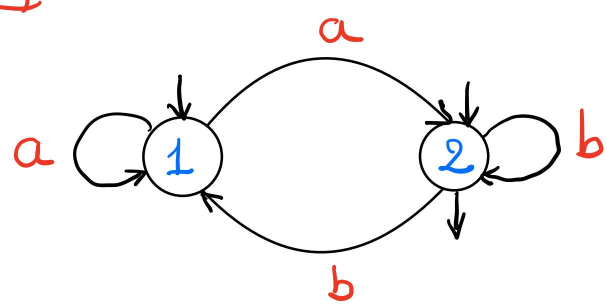
Suppose $A = \{a, b\}$

then $A^2 = \{aa, ab, ba, bb\}$

$(A^2)^* = \{$ words of even length $\}$

---

## Automata




Set of States : $\{1, 2\}$

Alphabet $A = \{a, b\}$

Initial states : $\{1, 2\}$

Final states : $\{2\}$

Transitions : $\{ (1, a, 1), (1, a, 2), (2, b, 1), (2, b, 2) \}$

$a \circlearrowleft 1$ $\quad 1 \xrightarrow{a} 2 \quad 2 \xrightarrow{b} 1 \quad 2 \circlearrowright b$

Consecutive transitions

$$p \xrightarrow{a} q \xrightarrow{c} r$$

$$1 \xrightarrow{a} 2 \xrightarrow{b} 2$$

**Path** : a set of consecutive transitions

$$p_0 \xrightarrow{a_1} p_1 \xrightarrow{a_2} p_2 \rightarrow \cdots \xrightarrow{a_k} p_k$$

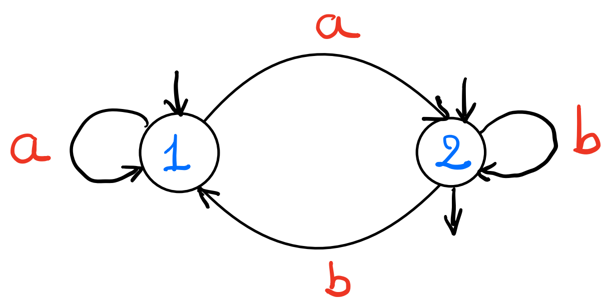
$p_0$ is the origin of the path

end of the path

$a_1 a_2 \cdots a_k$ is the $\boxed{\text{label}}$ of the path

Special path: the empty path around some state $q$

$q \xrightarrow{1} q$ no transition




Initial path = path starting in an initial state
Final path = path ending in a final state
successful path = path both initial and final.

$\to 1 \xrightarrow{a} 1 \xrightarrow{a} 2 \xrightarrow{b} 2 \xrightarrow{b} 1 \xrightarrow{a} 2 \xrightarrow{b} 2 \to$

$\to 1 \xrightarrow{a} 1 \xrightarrow{a} 2 \xrightarrow{b} 2 \xrightarrow{b} 1 \xrightarrow{a} 2 \xrightarrow{b} 1$

The language { recognised / defined / accepted } by the automaton is the set of labels of some successful path.

Here the word aabbab is recognised

Exercices $A = \{a, b\}$

$$A^{*} A = A^{*}$$

$$(L^*)^* = L^* \qquad (L^c)^c = L$$

$$(a^*b)^* =$$

$$A^* \supseteq A^*A^* = \{ uv \mid u \in A^*, v \in A^* \}$$
$$\supseteq \{ 1v \mid v \in A^* \} = A^*$$

$$a^*b = \{1, a, a^2, \ldots \} b$$
$$= \{b, ab, a^2b, \ldots \}$$

$$(a^*b)^* = 1 \cup A^*b$$

$$(a^*b)^* = \{ a^{n_1} b a^{n_2} b \ldots a^{n_k} b, \; n_1 \geq 0, n_2 \geq 0, \ldots, n_k \geq 0, k \geq 0 \}$$

$$abbbab = a^1 b a^0 b a^0 b a^1 b$$