GLAVA 1

Malo čiste matematike

Nema ništa praktičnije od dobre teorije.
Leonardo Da Vinči

3. Azbuke, riječi, jezici

Azbuka je, prosto, neki konačan neprazan skup A objekata koji se nazivaju simbolima.

Riječ ili string je n-torka simbola iz skupa A. Umjesto da pišemo riječ kao ($a_1, a_2, \ldots, a_n$), uobičajeno je da riječ pisemo kao $a_1 a_2 \ldots a_n$, to jest bez zagrada i zareza. Ako je riječ $u = a_1 a_2 \ldots a_n$, tada kažemo da je n **dužina riječi** u i pišemo $|u| = n$.

Uvešćemo i jednu posebnu riječ dužine 0, koju ćemo zapisivati takođe simbolom 0 (zašto tako radimo biće jasno u Glavi 5.).

Skup svih riječi azbuke A označićemo sa A*.

Bilo koji podskup od A* zvaćemo **jezikom** azbuke A.

Ne pravimo razliku između simbola $a \in$A i riječi dužine 1 koja se sastoji samo od tog simbola.

Ako su $u, v \in$A*, tada pišemo uv da označimo riječ koja se dobija kada iza niza simbola riječi u, doda niz simbola riječi v.

Na primjer, neka je A $= \{a, b, c\}$, $u = bab$ i v = caa, onda je

$$uv = babcaa \quad \text{i} \quad vu = caabab.$$

Jasno je da je, za svako u,

$$u0 = 0u = u, \text{ gdje je 0 prazna riječ,}$$

kao i da za svako u, v, w, važi:

$$u(vw) = (uv)w.$$

Takođe, ako je $uv = uw$ ili $vu = wu$, onda je $v = w$.

Ako je u riječ, i $n > 0$, pišemo

$$u^{[n]} = uu\ldots u \quad (n \ puta).$$

Takođe, pišemo $u^{[0]} = 0$. Koristimo uglaste zagrade da izbjegnemo konfuziju sa stepenovanjem.

Ako je $u \in$A*, pisaćemo u^{R} da označimo riječ u, napisanu unazad, tj. ako je $u = a_1 a_2 \ldots a_n$, onda je $u^{\mathrm{R}} = a_n \ldots a_2 a_1$. Jasno, $0^{\mathrm{R}}=0$, a $(uv)^{\mathrm{R}} = v^{\mathrm{R}} u^{\mathrm{R}}$, gdje su $u, v \in$A*.