Formális nyelvek és automaták

Nagy Sára gyakorlatai alapján
Készítette: Nagy Krisztián

1. gyakorlat

Bevezető a kurzushoz

„ Az emberi munka eredménye mindig valamilyen társadalmi igényt elégít ki. Ez a megállapítás igaz a tudományos élet területén is. Míg az elvont tudományok esetében az eredmény és az azt kiváltó társadalmi igény megfeleltetése gyakran nem triviális feladat, az alkalmazott tudományok művelői általában jól tudják, milyen gyakorlati céllal végzik kutató munkájukat.

A matematikai nyelvészetre gondolva ez a megfeleltetés teljesen egyértelmű.

A második világháború alatt és az azt követő években a kutatási és fejlesztési munkákban forradalmi változások következtek be. A műszaki és természettudományokban ugrásszerűen megnőtt a magasan kvalifikált szakemberek száma, ráadásul sokkal nagyobb hányaduk kezdett a kutatás és fejlesztés területén dolgozni, mint az azt megelőző időkben. Ez a változás természetesen az új műszaki és tudományos eredmények, vele együtt pedig a tudományos publiká-ciók számának hirtelen növekedésével járt.

Korábban egy szakember viszonylag könnyen figyelemmel kísérhette azt a mintegy 5-6 folyóiratot, ahol tartalmas cikkeket remélhetett. A problémát tovább nehezítette, hogy a publikációk egy része, elsősorban az angolul beszélők számára „szokatlan” nyelven volt írva, mint az orosz, japán vagy akár a magyar.

Az információhoz való hozzáférés igénye hirtelen megnövelte a fordí-tandó cikkek számát olyannyira, hogy ezt a lényegében egysíkú, majdhogynem mechanikus munkát sem fordítói kapacitással, sem pénzzel nem lehetett győzni.

Ugyanakkor az egyik, éppen abban az időben, az 50-es évek elején kiteljesedő, és az információrobbanáshoz nem kis mértékben hozzájáruló új tudomány, a számítástechnika azt állította, hogy célja a gépies szellemi munka kiváltása. Magától értetődő volt tehát az a törekvés, hogy a fordítást, ezt a gépies szellemi munkát egy számítástechnikai eszköz, a számítógép végezze.

Minthogy azonban a számítógép csak szabatosan megfogalmazott feladat megoldására képes, szükség volt a fordítás mint feladat formális leírására. Ez volt tehát az a kiváltó társadalmi igény, amely az 50-es években szinte a semmiből egy új tudomány a matematikai nyelvészet megalkotását eredményezte. Az új tudomány születési évének az 1956 esztendőt tekinthetjük. Ekkor publikálta ugyanis a szakma atyja, és mindmáig egyik legnagyobb egyénisége Noam Chomsky munkásságának első eredményeit.”

/Forrás: Bach István – Formális nyelvek Bevezetés Typotex kiadó; Budapest 2002/
/A könyv az alábbi linken elérhető: http://mek.niif.hu/05000/05099/05099.pdf /

A tárgy így igen fontos szerepet játszik a matematikai nyelv megalkotásában, továbbá a fordítóprogramok világában.

Alapfogalmak:

/A pontos definíciók és részletes leírások megtalálhatók Hunyadvári László tanár úr honlapján az alábbi linken: http://aszt.inf.elte.hu/~hunlaci/defi-1.pdf /

Gyakorlati szempontból:

Legyen X ábécé.
$a \in X$ ekkor betű.

$X_i = \{a, b\}$ egy olyan ábécé, amely az a és a b betűket tartalmazza.
Jelölje u a szavakat.
Ekkor például a fenti ábécé esetén $bab,\ baba,\ aaa, abbbb$ az X_i ábécé betűiből előállítható szavak.

Vegyük az $u = abba$ szót. Jelölje $l(u)$ a szavak hosszát. A mi esetünkben most $l(u) = 4$, mivel az $u = abba$ szóban 4 betű található.

Üres szó: ε . Az üres szó hossza: $l(\varepsilon) = 0$
$0 \leq l(u) < \infty$

Műveletek szavakkal:

- Konkatenáció

  $u = u_1 u_2 \dots u_k \quad l(u) = k$
  $v = v_1 v_2 \dots v_s \quad l(v) = s$, ekkor
  $uv = u_1 u_2 \dots u_k v_1 v_2 \dots v_s \quad l(uv) = k + s$

  $(u'u'')u''' = u'(u''u''')$ asszociatív
  $\varepsilon u = u = u\varepsilon$ ε semleges elem

- Hatványozás
  (Erősebb a konkatenációnál)

  $u^2 = uu$
  $u^3 = uuu$
  $(ab)^2 = abab$
  $a^2 b^2 = aabb$
  $ab^3 = abbb$
  $u^1 = u$
  $u^0 = \varepsilon$

- Szavak fordítottja
  Jelölése: u^{-1}
  $u = u_1 u_2 \dots u_{n-1} u_n\ esetén\ u^{-1} = u_n u_{n-1} \dots u_2 u_1$

Nyelv:

$L \subseteq X^*$ A nyelv az ábécé lezártjának egy részhalmaza (L kontinuum számosságú)
X^* - összes szó (megszámlálhatóan végtelen)
$L_1 = \{ab, a\}$ $|L_1| = 2$
$L_2 = \{au \mid u \in X^*\} = \{a, a^2, ab, a^3, a^2b, aba, ab^2 \ldots\}$ $|L_1| = \infty$

$L_1 \subset L_2$ ✔
$L_1 \cap L_2 = L_1$
$L_1 \backslash L_2 = \emptyset$ $|L_1 \backslash L_2| = 0$

$|\emptyset| = 0$ üres nyelv
$L_3 = \{\varepsilon\}$ $|L_3| = 1$

Műveletek nyelvekkel:

- Unio

  $L' \cup L'' = \{u \mid u \in L' vagy\ u \in L''\}$
  $\emptyset \cup L = L = L \cup \emptyset$,

- Konkatenáció

  $L'L'' = \{u'u'' \mid u' \in L' \text{és}\ u'' \in L''\}$
  Példák:
  $L_1 = \{ab, a\}$ $L_2 = \{bb, a\}$
  $$\left.\begin{array}{l} L_1L_2 = \{abbb, abb, aba, aa\} \\ L_2L_1 = \{bbab, bba, aab, aa\} \end{array}\right\} \rightarrow \text{nem kommutatív}$$
  $L_1L_2 \cap L_2L_1 = \{aa\}$
  $(L'L'')L''' = L'(L''L''')$ $asszociatív$
  $(L' \cup L'')L''' = L'L''' \cup L''L''' \rightarrow igaz\ a\ disztributivitás\ (unioval)$
  $(L' \cap L'')L''' \neq L'L''' \cap L''L'''$
  $\{\varepsilon\}L = L = L\{\varepsilon\}$
  $\emptyset L = \emptyset$

- Hatványozás

  $L^2 = LL$ $L^3 = LLL$ $L^n = L^{n-1}L = LL^{n-1}$
  $L^1 = L$
  $L^0 = \{\varepsilon\}$

- Lezárás

  $L^* = L^0 \cup L^1 \cup L^2 \cup \ldots = \cup_{i=0}^{\infty} L^i$
  $\left(L^+ = L^1 \cup L^2 \cup \ldots = \cup_{i=1}^{\infty} L^i \quad pozitív\ lezárás\right)$
  $X^* = \{a,b\}^* = \{\varepsilon\} \cup \{a,b\} \cup \{aa,ab,ba,bb\} \cup \{három\ hosszúságú\ szavak\} \cup \ldots$
  Példa:
  Határozzuk meg, hogy az adott szó eleme-e $L^* = \{b^2, a\}^*$ -nak!
  - $abba \in L^*$, mert felbontható az alábbi módon: $a|b^2|a$
  - $baba \notin L^*$, mert nem lehet felbontani L^* -beli szavakra

Reguláris műveletek nyelveken:

- unio
- konkatenáció
- lezárás

Reguláris nyelvek családja:

- elemi reguláris nyelvek $\emptyset, \{\varepsilon\}, \{a\} \quad a \in U, ahol\ U\ az\ univerzális\ jelkészletet\ jelenti$
- reguláris nyelv az elemi reguláris nyelvekből a reguláris műveletek véges sokszori alkalmazásával előállíthatók

Példák:
$L_1 = \{bb,a\} = \{b\}\{b\} \cup \{a\}$
$X^* = \{a,b\}^* \quad L_2 = \left\{au \mid u \in X^*\right\} = \{a\}(\{a\} \cup \{b\})^*$

Példa nem reguláris nyelvre:
$L = \{a^n b^n | \, n \geq 0\} = \left\{\varepsilon, ab, a^2 b^2, a^3 b^3, \ldots\right\}$

Egyéb ajánlott irodalom: Hunyadvári László – Manhertz Tamás Automaták és Formális Nyelvek című könyve.
Elérhető: http://aszt.inf.elte.hu/~hunlaci/book.pdf