Algoritmuselmélet
3. heti feladatsor megoldása

1. Tekintsük az alábbi L nyelvet:

$$L = \{a^n b^m \mid n, m \geq 1, n \neq 3m\}$$

Lássa be, hogy ez az L nyelv nem reguláris.

Megoldás

Tegyük fel indirekt, hogy a nyelv reguláris, azaz van olyan M determinisztikus, teljes véges automata, ami a nyelv szavaira elfogadó, a nyelven kívüli szavakra pedig elutasító állapotba jut.

Jelölje N ezen automata állapotainak számát és tekintsük az $a^3, a^6, a^9, \ldots, a^{3N}, a^{3(N+1)}$ szavakat. Mivel ezek száma $N+1$, a skatulya elv miatt biztosan van köztük két olyan különböző szó, a^{3i} és a^{3j} $(i \neq j)$ melyeket a kezdőállapotból beolvasva ugyanabba az (X-szel jelölt) állapotba jutunk. Tekintsük most az $a^{3i}b^i$ szót, ami nem eleme a nyelvnek, ezért ezt elolvasva egy nem elfogadó Y állapotba jut az M automata.

Tekintsük továbbá az $a^{3j}b^i$ szót is. Mivel a^{3j} olvasása után ugyanott vagyunk, mint a^{3i} olvasása után, nevezetesen X-ben, innentől pedig ugyanaz történik, mint $a^{3i}b^i$ esetén, ezért az $a^{3j}b^i$ szó hatására is az Y nem elfogadó állapotba jutunk, ami rossz, mert az $a^{3j}b^i$ szót el kéne fogadnunk, hiszen $3j \neq 3i$.

Így ellentmondásra jutottunk, vagyis a kezdeti feltevésünk volt hibás, nem létezhez ez az M automata.

2. A $\{0, 1\}$ ábécé feletti L_1 nyelv azokból a szavakból áll, amikben nincsen `00` részszó. (Tehát például `0101`, `11`, ε benne vannak a nyelvben, de `10011` nincsen). Adjon reguláris kifejezést erre a nyelvre és magyarázza el, hogy a kifejezés miért éppen ezt a nyelvet írja le.
Ezt úgy lehet például megtenni, hogy elmagyarázzuk, hogy a kifejezés melyik része miről gondoskodik, a kifejezés miért ír le minden jó szót (itt azt ellenőrizzük le, hogy minden lehetőségre gondoltunk), illetve hogy hogyan értük el, hogy más szavakat ne írjon le (mi gondoskodik arról, hogy ne legyen két 0 egymás mellett).

Megoldás

Megmutatjuk, hogy az alábbi reguláris kifejezés jó lesz:

$$1^* \cdot (011^*)^* \cdot (0 + \varepsilon)$$

Az első tag gondoskodik arról, hogy kezdhessük a szót `1`-sel, a középső tag szeparálja el a `0`-kat egymástól `1`-ekkel, az utolsó tag pedig gondoskodik arról, hogy a szó végződhessen `0`-ra is.

Ebben a kifejezésben minden `0`-t követ vagy legalább egy darab `1` vagy a `0` után vége van a szónak, ezért `0` után sose jöhet másik nulla, vagyis `00` nincs a leírt szavakban.

Azt kell még belátnunk, hogy minden olyan szót, amiben nincsen `00` leírtunk. Ehhez vegyük észre, hogy ezek a szavak alapvetően ötfélék lehetnek:

(a) Nincs bennük `0`: ezt leírjuk akkor, amikor a kifejezés 2. tagjában, $(011^*)^*$-ben egyszer sem vesszük (011^*)-t és a harmadik tagban ε-t veszünk.

(b) Van bennük `0`, nincs bennük `00` és a szó elején `1` áll. Ezek a szavak felbonthatók a reguláris kifejezésnek megfelelően: a szó eleji `1`-ket leírja 1^*, utána pedig minden egyes `0`-val kezdődő blokkot (kivéve az esetleg levő szó végi egy darab 0-t) leír a középső tagban levő (011^*) egy-egy példánya. Ha a szó végén nem `0` van, akkor a harmadik tagból az ε-t, különben pedig a `0`-t választjuk.

(c) Van bennük `0`, nincs bennük `00` és a szó elején `0` áll. Hasonlóan az előző esethez minden egyes `0`-val kezdődő blokkot (kivéve az esetleg levő szó végi egy darab 0-t) leír a középső tagban levő (011^*) egy-egy példánya. Ha a szó végén nem `0` van, akkor a harmadik tagból az ε-t, különben pedig a `0`-t választjuk itt is.