**Języki, automaty i obliczenia**
**kolokwium – przykładowe rozwiązania zadań**
**13 maja 2015**

**Zad. 1.** Dwa języki uważamy za *przemiennie równoważne*, jeśli dla każdego słowa $w$ należącego do jednego z języków, drugi język zawiera słowo różniące się od $w$ tylko kolejnością liter (ale nie liczbą wystąpień liter). Wyrażenia regularne są przemiennie równoważne gdy ich języki są. Napisz wyrażenie regularne o głębokości gwiazdkowej 1 (tzn. bez zagnieżdżenia gwiazdki) przemiennie równoważne wyrażeniu

$$a(a(bc)^* + b)^*.$$

**Rozwiązanie.** Skoro kolejność liter jest nieistotna, to równoważnie możemy rozpatrywać wyrażenie

$$a\,(a(bc)^*)^*\,b^*.$$

Podzielmy język $L$ opisany przez to wyrażenie na dwa rozłączne podzbiory $L_1$ i $L_2$. Niech $L_1$ zawiera słowa z $L$ nie zawierające $c$. Zatem wyrażenie regularne dla $L_1$ otrzymamy przez usunięcie wewnętrznej gwiazdki:

$$L_1 \;=\; a\,a^*\,b^* \;=\; a^+\,b^*.$$

$L_2$ zawiera słowa z $L$, które zawierają przynajmniej jedną literę $c$. Zatem do „wygenerowania" tych słów potrzebna jest przynajmniej jednokrotna iteracja wewnętrznej gwiazdki. Skoro kolejność liter jest nieistotna to możemy wykonać wszystkie iteracje wewnętrznej gwiazdki „przy pierwszej okazji", otrzymując przemiennie równoważne wyrażenie:

$$a\,a(bc)^+\,a^*b^*,$$

które można dalej uprościć do:

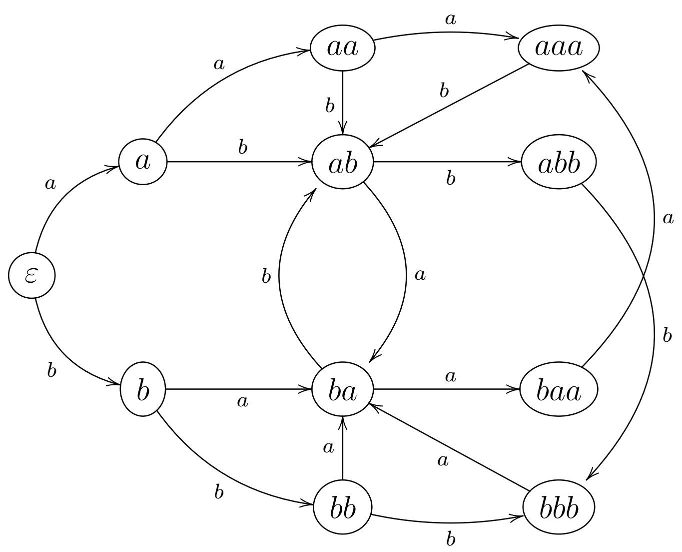
$$a\,a^+\,(bc)^+\,b^*.$$

Wyrażenie regularne równoważne przemiennie wyrażeniu z zadania to:

$$a^+\,b^* \;+\; a\,a^+\,(bc)^+\,b^*, \qquad \text{albo} \qquad a^+\,b^*\,(a(bc)^+ \;+\; \varepsilon).$$

**Zad. 2.** Narysuj minimalny automat deterministyczny dla języka tych słów nad alfabetem $\{a, b\}$, które nie zawierają żadnego palindromu długości 4. Zabronione są skrzyżowania krawędzi na rysunku.

**Rozwiązanie.** Wszystkie stany na rysunku są akceptujące, jeden stan nieakceptujący (śmietnik) pominięty wraz z wszystkimi krawędziami do niego prowadzącymi.




Automatu tego *nie da się narysować bez skrzyżowań krawędzi*, gdyż graf nieskierowany otrzymany przez wymazanie kierunku strzałek nie jest planarny.

**Zad. 3.** Rozważmy język $L$ słów nad alfabetem $\{a, b\}$, w których najdłuższy ciąg kolejnych liter $a$ jest ściśle dłuższy niż najdłuższy ciąg kolejnych liter $b$. Na przykład słowo

$$bbaaaabbbababbb = b^2a^4b^3a^1b^1a^1b^3$$

należy do języka, bo ciąg $a^4$ jest ściśle dłuższy niż każdy z czterech ciągów $b^2, b^3, b^1, b^3$. Czy język $L$ jest bezkontekstowy? Jeśli tak to podaj gramatykę bezkontekstową generującą ten języki i uzasadnij jej poprawność, a jeśli nie to udowodnij, że nie jest.

**Rozwiązanie.** Język $L$ nie jest bezkontekstowy. Aby tego dowieść, użyjemy lematu o pompowaniu. Dla dowolnego $n$, rozważmy słowo

$$w_n \;=\; b^n a^{n+1} b^n \in L$$

i niech $w_n \;=\; uxvyw$ będzie dowolną dekompozycją $w_n$ taką, że $|xy| \geq 1$ oraz $|xvy| \leq n$. Rozważmy dwa przypadki.

Jeśli $a$ nie występuje w $xy$, to obydwa słowa $x, y$ znajdują się po tej samej stronie infiksu $a^{n+1}$, ponieważ $|xvy| \leq n$. Czyli słowo $ux^2vy^2w$ to

$$\text{albo} \qquad b^{n+m}a^{n+1}b^n \qquad \text{albo} \qquad b^n a^{n+1} b^{n+m},$$

gdzie $m = |xy| \geq 1$, w obydwu przypadkach nie należy do $L$.

Jeśli $a$ występuje w $xy$, to $xy \in b^*a^+$ albo symetrycznie $xy \in a^+b^*$, znów dlatego, że $|xvy| \leq n$. Powiedzmy, że $xy = b^m a^{k+1}$. Wtedy słowo $uvw = b^{n-m}a^{n-k}b^n$ nie należy do $L$.

**Zad. 4.** Czy język $L = \{a, b, \$\}^* - M$, gdzie

$$M \;=\; \{\, w\$v \;:\; w, v \in \{a,b\}^* \text{ są palindromami i } w\$v \text{ jest palindromem}\,\}$$

jest językiem bezkontekstowym? Jeśli tak to podaj gramatykę bezkontekstową generującą ten język i uzasadnij jej poprawność, a jeśli nie to udowodnij, że nie jest.

**Rozwiązanie.** Język $L$ jest bezkontekstowy, ponieważ jest sumą następujących języków:

1. $L_{\neq 1}$: słowa, w których liczba wystąpień litery \$ jest różna od 1
2. $L_{\$}$: słowa $w\$v$ takie, że $v \neq w^R$ $(w, v \in \{a,b\}^*)$
3. słowa $w\$v$ takie, że $w$ nie jest palindromem $(w, v \in \{a,b\}^*)$
4. słowa $v\$w$ takie, że $w$ nie jest palindromem $(w, v \in \{a,b\}^*)$

Gramatyka dla $L$ zawiera zatem cztery produkcje dla nieterminala początkowego $S$:

$$S \;\longrightarrow\; S_{\neq 1} \mid N_{\$} \mid N\,\$\,S_0 \mid S_0\,\$\,N.$$

Pozostałe produkcje to:

$$\begin{array}{rcll}
S_0 & \longrightarrow & S_0\,a \mid S_0\,b \mid \varepsilon & \text{czyli } L(S_0) = \{a,b\}^* \\
S_{\neq 1} & \longrightarrow & S_0 \mid S_{\geq 2} & \text{czyli } L(S_{\neq 1}) = L_{\neq 1} \\
S_{\geq 2} & \longrightarrow & S_0\$S_0\$ \mid S_{\geq 2}\,a \mid S_{\geq 2}\,b \mid S_{\geq 2}\,\$ & \\
N & \longrightarrow & a\,N\,a \mid b\,N\,b \mid a\,S_0\,b \mid b\,S_0\,a & \text{czyli } L(N) = \text{nie-palindoromy} \\
N_\$ & \longrightarrow & a\,N_\$\,a \mid b\,N_\$\,b \mid a\,S_1\,b \mid b\,S_1\,a & \text{czyli } L(N_\$) = L_\$ \\
S_1 & \longrightarrow & S_0\,\$\,S_0 &
\end{array}$$

**Zad. 5*.** To samo pytanie co w zadaniu 4, ale dla języka $L = \{a,b\}^* - h(M)$, gdzie $h$ to homomorfizm wyznaczony przez $a \mapsto a$, $b \mapsto b$, $\$ \mapsto \varepsilon$.

**Rozwiązanie.** Język $h(M)$ zawiera słowa postaci $w\,w^R$ takie, że $w$ jest palindromem. Ponieważ każdy palindrom spełnia równość $w = w^R$, mamy

$$h(M) \;=\; \{w\,w \;:\; w \text{ jest palindromem}\}.$$

Zauważmy, że następujące stwierdzenia są równoważne:

- $w$ jest palindromem
- $w\,w$ jest palindromem,

zatem mamy równość

$$h(M) \;=\; \{w\,w \;:\; w\,w \text{ jest palindromem}\} \;=\; \{w\,w : w \in \{a,b\}^*\} \cap \text{palindromy}.$$

Dalej łatwo.