Języki formalne i techniki translacji 2013

Rozwiązania kartkówek z lematu o pompowaniu i lematu Ogdena (języki bezkontekstowe)

Autor: Anna Lauks-Dutka

Treść zadania: Pokaż, że następujące języki nie są bezkontekstowe:

- $L_1 = \{w : w \in \{a, b, c\}^* \wedge |w|_a \neq |w|_b \neq |w|_c\}$
- $L_2 = \{w : w \in \{a, b, c\}^* \wedge |w|_a = |w|_b = |w|_c\}$
- $L_3 = \{w : w \in \{a, b, c\}^* \wedge |w|_a = |w|_c > |w|_b\}$
- $L_4 = \{www : w \in \{0, 1\}^*\}$
- $L_5 = \{ww^Rw : w \in \{0, 1\}^*\}$, gdzie w^R oznacza odwrócenie kolejności liter w słowie w
- $L_6 = \{a^n c^k b^n : n \neq k\}$

Rozwiązanie: Przydadzą się:

Lemat o pompowaniu

Niech L język bezkontekstowy. Wtedy istnieje stała n taka, że jeśli $z \in L$ oraz $|z| \geqslant n$, to można podzielić z na $z = uvwxy$ takie, że:

1. $|vx| \geqslant 1$
2. $|vwx| \leqslant n$
3. $\forall_{i \in N} z' = uv^iwx^iy \in L$

Lemat Ogdena

Niech L język bezkontekstowy. Wtedy istnieje stała n taka, że jeśli $z \in L$ oraz $|z| \geqslant n$ i oznaczymy w z n lub więcej pozycji jako wyróżnione, to można podzielić z na $z = uvwxy$ takie, że:

1. v i x zawierają łącznie co najmniej jedną wyróżnioną pozycję
2. vwx zawiera co najwyżej n wyróżnionych pozycji
3. $\forall_{i \in N} z' = uv^iwx^iy \in L$

Ad. $\mathbf{L_1 = \{w : w \in \{a, b, c\}^* \wedge |w|_a \neq |w|_b \neq |w|_c\}}$

Niech n stała z **lematu Ogdena**. Niech $m > n$. Wybieramy słowo $z = a^{m+m!}b^m c^{m+m!}$ i oznaczamy m liter b jako wyróżnione. Słowo $z \in L_1$ i $|z| = 3m + 2m! > n$. Musimy rozpatrzyć wszystkie możliwe podziały tego słowa na $z = uvwxy$ takie, że: v i x zawierają łącznie co najmniej jedną wyróżnioną pozycję oraz vwx zawiera co najwyżej n wyróżnionych pozycji. Zatem:

- nie możemy pompować samego a ani samego c (brak wyróżnionych),
- nie możemy pompować jednocześnie a oraz c (pomiędzy nimi jest więcej niż n wyróżnionych liter).

Pozostają nam do rozpatrzenia podziały, w których:

- pompujemy tylko b (liczba litera a oraz c się nie zmienia). Wtedy dla $i = \frac{m!}{p} + 1$, gdzie $p = |vx|_b$ dostaniemy:

  $$|z'|_b = |uv^iwx^iy|_b = |z|_b + (i-1)|vx|_b = m + (i-1)p = m + \left(\frac{m!}{p} + 1 - 1\right)p = m + m! = |z'|_a,$$

  więc $z' \notin L_1$
- pompujemy a oraz b (liczba liter c nie ulega zmianie). Wtedy dla $i = \frac{m!}{p} + 1$, gdzie $p = |vx|_b$ dostaniemy:

  $$|z'|_b = |uv^iwx^iy|_b = |z|_b + (i-1)|vx|_b = m + (i-1)p = m + \left(\frac{m!}{p} + 1 - 1\right)p = m + m! = |z'|_c,$$

  więc $z' \notin L_1$
- pompujemy b oraz c (liczba liter a nie ulega zmianie). Wtedy dla $i = \frac{m!}{p} + 1$, gdzie $p = |vx|_b$ dostaniemy:

  $$|z'|_b = |uv^iwx^iy|_b = |z|_b + (i-1)|vx|_b = m + (i-1)p = m + \left(\frac{m!}{p} + 1 - 1\right)p = m + m! = |z'|_a,$$

  więc $z' \notin L_1$

Uwagi końcowe:

- bez silni się nie obędzie! Przykładowe **złe słowo**: $z = a^n b^{n+1} c^{n+2}$. Pompując nawet tylko b trzeba dobrać i tak, żeby „dobić” liczbę liter b do liczby liter a lub c. Problem polega na tym, że jak $|z|_c = |z|_b + 1 = |z|_a + 2$, to jedyny przypadek, w którym dobijemy w z' liczbę liter b do a lub c to taki, w którym do v lub x wpadnie nam dokładnie jedno b i weźmiemy $i = 0$ lub $i = 2$. Jeżeli do v wpadnie więcej liter b, to zawsze można pompować i słowo dalej będzie należało do języka!
- w języku ważna jest liczba liter a nie ich kolejność, więc do v i x mogą wpadać bloki złożone z różnych rodzajów liter, przeploty liter po napompowaniu nie mają znaczenia,
- analogiczny do powyższego dowód **nie przejdzie dla słowa** $z = a^{n+n!} b^n c^{n+n!}$. Wówczas, zgodnie z lematem, trzeba rozpatrzyć też przypadek, w którym pompujemy a, b oraz c. Długość vwx w lemacie Ogdena nie ma znaczenia! Ważna jest tylko liczba liter wyróżnionych - przy rozpatrywanych podziałach trzeba uwzględnić wszystkie, w których do vwx wpada co najwyżej n wyróżnionych liter. Żeby nie można było pompować a oraz c jednocześnie trzeba, konstruując z, oddzielić je co najmniej $n+1$ literami wyróżnionymi. Można więc wziąć słowo $z = a^{n+1+n!} b^{n+1} c^{n+1+n!}$ lub $z = a^{m+m!} b^m c^{m+m!}$ dla $m > n$ i koniecznie oznaczyć co najmniej $n+1$ liter b jako wyróżnione (oznaczanie jako wyróżnione a lub c w tych słowach kończy się tym, że też istnieją przypadki, w których wszystkie rodzaje liter można pompować),
- dowód **nie przejdzie z lematu o pompowaniu** (problem z przypadkami, w których trzeba zmniejszać liczbę liter z $m + m!$ do m, niewykonalne).

Ad. $\mathbf{L_2} = \{\mathbf{w} : \mathbf{w} \in \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}^* \wedge |\mathbf{w}|_\mathbf{a} = |\mathbf{w}|_\mathbf{b} = |\mathbf{w}|_\mathbf{c}\}$

Niech n stała z **lematu o pompowaniu**. Wybieramy słowo $z = a^n b^n c^n$. Słowo $z \in L_2$ i $|z| = 3n \geqslant n$. Musimy rozpatrzyć wszystkie możliwe podziały tego słowa na $z = uvwxy$ takie, że: $|vx| \geqslant 1$ oraz $|vwx| \leqslant n$. Możliwe przypadki:

- pompujemy tylko a,
- pompujemy a oraz b,
- pompujemy tylko b,
- pompujemy b oraz c,

- pompujemy tylko c.

W każdym z tych przypadków dla dowolnego $i \neq 1$ zmienimy w $z' = uv^iwx^iy$ liczbę jednego bądź dwóch rodzajów liter, nie zmieniając liczby liter trzeciego rodzaju, więc $z' \notin L_2$.

Nie możemy jednocześnie pompować a i c (czyli odpadają nam przypadki, w których pompujemy a i c oraz a, b i c), ponieważ a i c są od siebie zbyt odległe.

Uwagi końcowe:

- w języku ważna jest liczba liter a nie ich kolejność, więc do v i x mogą wpadać bloki złożone z różnych rodzajów liter, przeploty liter po napompowaniu nie mają znaczenia,
- lemat Ogdena to uogólnienie lematu o pompowaniu, więc dowód z wykorzystaniem lematu Ogdena przejdzie dla tego samego słowa pod warunkiem, że oznaczymy wszystkie litery jako wyróżnione. Biorąc to samo słowo z i oznaczając np. **tylko** b jako wyróżnione dostajemy przypadek, w którym wszystkie rodzaje liter można pompować, wtedy dla „wrednego" podziału ($|vx|_a = |vx|_b = |vx|_c$) zawsze możemy pompować i słowo dalej należy do języka.

Ad. $\mathbf{L_3 = \{w : w \in \{a, b, c\}^* \wedge |w|_a = |w|_c > |w|_b\}}$

Niech n stała z **lematu o pompowaniu**. Wybieramy słowo $z = a^{n+1}b^nc^{n+1}$. Słowo $z \in L_3$ i $|z| = 3n+2 \geqslant n$. Musimy rozpatrzyć wszystkie możliwe podziały tego słowa na $z = uvwxy$ takie, że: $|vx| \geqslant 1$ oraz $|vwx| \leqslant n$. Możliwe przypadki:

- pompujemy tylko a albo pompujemy a i b. Wtedy dla dowolnego $i \neq 1$ zmienimy w $z' = uv^iwx^iy$ liczbę liter a nie zmieniając liczby liter c, więc $z' \notin L_3$,
- pompujemy tylko c albo pompujemy b i c. Wtedy dla dowolnego $i \neq 1$ zmienimy w $z' = uv^iwx^iy$ liczbę liter c nie zmieniając liczby liter a, więc $z' \notin L_3$,
- pompujemy tylko b. Wtedy dla $i = 2$ dostaniemy:

$$|z'|_b = |uv^iwx^iy|_b = |z|_b + |vx|_b = n + |vx|_b \geqslant n + 1 = |z'|_a,$$

  więc $z' \notin L_3$

Nie możemy jednocześnie pompować a i c (czyli odpadają nam przypadki, w których pompujemy a i c oraz a, b i c), ponieważ a i c są od siebie zbyt odległe.

Uwagi końcowe:

- w języku ważna jest liczba liter a nie ich kolejność, więc do v i x mogą wpadać bloki złożone z różnych rodzajów liter, przeploty liter po napompowaniu nie mają znaczenia,
- inne **dobre słowo**: $z = a^{n+1}c^{n+1}b^n$. Przypadki: pompujemy samo a albo samo c albo samo b, albo a i c (wtedy musimy wziąć $i = 0$) albo c i b.

Ad. $\mathbf{L_4 = \{www : w \in \{0, 1\}^*\}}$

Niech n stała z **lematu o pompowaniu**. Wybieramy słowo $z = 0^n1^n0^n1^n0^n1^n$. Słowo $z \in L_4$ i $|z| = 6n \geqslant n$. Musimy rozpatrzyć wszystkie możliwe podziały tego słowa na $z = uvwxy$ takie, że: $|vx| \geqslant 1$ oraz $|vwx| \leqslant n$. Możliwe przypadki:

- pompujemy tylko zera. Uwaga: bloki zer w z są oddzielone od siebie n jedynkami. Ponieważ $|vwx| \leqslant n$, możemy pompować tylko jeden z trzech bloków zer! Zatem dla dowolnego $i \neq 1$ w $z' = uv^iwx^iy$ zaburzymy strukturę tylko jednego z trzech bloków zer. Takie słowo nie może mieć już dłużej postaci „www", więc $z' \notin L_4$

- pompujemy tylko jedynki. Analogicznie jak wyżej bloki jedynek w z są oddzielone od siebie n zerami. Ponieważ $|vwx| \leqslant n$, możemy pompować tylko jeden z trzech bloków jedynek. Zatem dla dowolnego $i \neq 1$ w $z' = uv^iwx^iy$ zaburzymy strukturę tylko jednego z trzech bloków jedynek. Takie słowo nie może mieć już dłużej postaci „www”, więc $z' \notin L_4$
- pompujemy zera i jedynki:
  - v albo x to blok 0^+1^+. Ponieważ $|vwx| \leqslant n$, to albo v albo x może być takiej postaci oraz możemy zahaczyć o jedną granicę zer i jedynek z częściami v, x, które pompujemy. Wtedy dla dowolnego $i \neq 1$ w $z' = uv^iwx^iy$ zaburzymy strukturę „www” pomiędzy pierwszym i drugim w, nie mając szans na analogiczną zmianę pomiędzy drugim i trzecim w albo na odwrót, więc $z' \notin L_4$
  - $v = 0^k, x = 1^l$, gdzie $k \geqslant 1, l \geqslant 1, k + l \leqslant n$. Analogicznie, ponieważ $|vwx| \leqslant n$, to możemy zahaczyć tylko o jedną granicę zer i jedynek z częściami v, x, które pompujemy. Wtedy dla dowolnego $i \neq 1$ w $z' = uv^iwx^iy$ zaburzymy strukturę „www” pomiędzy pierwszym i drugim w, nie mając szans na analogiczną zmianę pomiędzy drugim i trzecim w albo na odwrót, więc $z' \notin L_4$

Uwagi końcowe:

- analogiczny dowód przechodzi już dla słowa $z = 0^n10^n10^n1$. Nie mamy szans na pompowanie wszystkich trzech jedynek albo trzech bloków zer jednocześnie,
- trzeba słowo dobrać tak, żeby granice pomiędzy trzema w były sztywne. Przykładowe złe słowo: $z = (01)^n(01)^n(01)^n$. Dowód nie przechodzi, kiedy vx to wielokrotności bloku $(01)^3$. Wtedy dla dowolnego i możemy pompować i słowo dalej należy do języka! Z podobnej przyczyny złym słowem jest $z = 1^n1^n1^n$ czy $z = 0^n0^n0^n$.

Ad. $\mathbf{L_5 = \{ww^Rw : w \in \{0, 1\}^*\}}$

Niech n stała z **lematu o pompowaniu**. Wybieramy słowo $z = 0^n1^n1^n0^n0^n1^n$. Słowo $z \in L_5$ i $|z| = 6n \geqslant n$. Musimy rozpatrzyć wszystkie możliwe podziały tego słowa na $z = uvwxy$ takie, że: $|vx| \geqslant 1$ oraz $|vwx| \leqslant n$. Możliwe przypadki:

- pompujemy tylko zera. Uwaga: mamy dwa bloki zer w z, które są oddzielone od siebie $2n$ jedynkami. Ponieważ $|vwx| \leqslant n$, możemy pompować tylko jeden z nich! Zatem dla dowolnego $i \neq 1$ w $z' = uv^iwx^iy$ zmienimy liczbę zer tylko w jednym bloku, więc zaburzymy strukturę pomiędzy pierwszym a ostatnim w w „ww^Rw”. Takie słowo nie może mieć już dłużej postaci „ww^Rw”, więc $z' \notin L_5$
- pompujemy tylko jedynki. Analogicznie jak wyżej mamy dwa bloki jedynek w z, które są oddzielone od siebie $2n$ zerami. Ponieważ $|vwx| \leqslant n$, możemy pompować tylko jeden z nich! Zatem dla dowolnego $i \neq 1$ w $z' = uv^iwx^iy$ zmienimy liczbę jedynek tylko w jednym bloku, więc zaburzymy strukturę pomiędzy pierwszym a ostatnim w w „ww^Rw”. Takie słowo nie może mieć już dłużej postaci „ww^Rw”, więc $z' \notin L_5$
- pompujemy zera i jedynki:
  - v albo x to blok 0^+1^+. Ponieważ $|vwx| \leqslant n$, to albo v albo x może być takiej postaci oraz możemy zahaczyć o jedną granicę zer i jedynek z częściami v, x, które pompujemy. Wtedy dla dowolnego $i \neq 1$ w $z' = uv^iwx^iy$ zmieniamy „ww^R” nie mając szans na analogiczną zmianę ostatniego „w” w „ww^Rw” albo na odwrót zmieniamy „w^Rw” nie mając szans na analogiczną zmianę pierwszego „w” w „ww^Rw”, więc $z' \notin L_5$
  - $v = 0^k, x = 1^l$, gdzie $k \geqslant 1, l \geqslant 1, k+l \leqslant n$ podobnie jak wyżej. Ponieważ $|vwx| \leqslant n$, to możemy zahaczyć tylko o jedną granicę zer i jedynek z częściami v, x, które pompujemy. Wtedy dla dowolnego $i \neq 1$ w $z' = uv^iwx^iy$ zmienimy „ww^R”nie mając szans na analogiczną zmianę ostatniego „w” w „ww^Rw” albo na odwrót zmieniamy „w^Rw” nie mając szans na analogiczną zmianę pierwszego „w” w „ww^Rw”, więc $z' \notin L_5$

Uwagi końcowe:

- inne istotnie różne od poprzedniego dobre słowo to: $z = (01)^n(10)^n(01)^n$. Granice podziału pomiędzy ww^R i $w^R w$, czyli 11 i 00 są tu dobrze oddzielone (odpowiednio dużą liczbą liter),
- dowód **nie przechodzi** dla słowa $z = 0^n 1 1 0^n 0^n 1$. Bloki zer są oddzielone od siebie tylko dwoma jedynkami, co oznacza, że możemy pompować zera w obu blokach jednocześnie. Istnieją złośliwe podziały (takie, że pompujemy dwa razy więcej zer z drugiego bloku niż z pierwszego), przy których dla dowolnego i możemy pompować! Złymi słowami są również $z = 1^n 1^n 1^n$ czy $z = 0^n 0^n 0^n$.

Ad. $\mathbf{L_6 = \{a^n c^k b^n : n \neq k\}}$

Niech n stała z **lematu Ogdena**. Wybieramy słowo $z = a^n c^{n+n!} b^n$. $z \in L_6$ i $|z| = 3n + n! \geqslant n$. Oznaczmy w nim n liter a jako wyróżnione. Musimy rozpatrzyć wszystkie możliwe podziały tego słowa na $z = uvwxy$ takie, że: v i x zawierają łącznie co najmniej jedną wyróżnioną pozycję oraz vwx zawiera co najwyżej n wyróżnionych pozycji. Zauważmy, że:

- v lub x muszą zawierać przynajmniej jedno a,
- do v i do x nie mogą wpaść bloki złożone z dwóch lub trzech rodzajów liter (a^+b^+ albo $a^+b^+c^+$), ponieważ dla $i = 2$ dostaniemy w $z' = uv^2wx^2y$ przeploty tych liter, więc $z' \notin L_6$.

Pozostają nam do rozpatrzenia następujące podziały:

- pompujemy tylko a, czyli $v = a^k, x = a^l$, gdzie $k + l \geqslant 1$ oraz $k + l \leqslant n$. Wtedy dla dowolnego $i \neq 1$ zmieniamy liczbę liter a w $z' = uv^iwx^iy$ nie zmieniając liczby liter b, więc $z' \notin L_6$
- pompujemy a i c, czyli $v = a^k, x = c^l$, gdzie $k \geqslant 1, l \geqslant 1$. Wtedy podobnie jak wyżej dla dowolnego $i \neq 1$ zmieniamy liczbę liter a w $z' = uv^iwx^iy$ nie zmieniając liczby liter b, więc $z' \notin L_6$
- pompujemy a i b, czyli $v = a^k, x = b^l$, gdzie $k \geqslant 1, l \geqslant 1$. Wtedy dla $i = \frac{n!}{p} + 1$, gdzie $p = |v|_a$ dostaniemy:

$$|z'|_a = |uv^iwx^iy|_a = |z|_a + (i-1)|vx|_a = n + (i-1)|v|_a = n + \left(\frac{n!}{p} + 1 - 1\right)p = n + n! = |z'|_c$$

Uwagi końcowe:

- w ostatnim rozpatrywanym podziale spełniającym warunki lematu zakładamy, że zmieniamy liczbę liter a i b. Musimy w tym wypadku rozpatrzyć między innymi taki, w którym pompujemy taką samą liczbę liter a co b, więc absolutnie nie można się tu posiłkować tym, że dla jakiegoś i dostaniemy różną liczbę liter a, co b,
- inne dobre słowo do lematu Ogdena to $z = a^{n+n!}c^n b^{n+n!}$, ale wtedy koniecznie trzeba oznaczyć n liter c jako wyróżnione,
- dowód nie przejdzie z lematu o pompowaniu (problem z przypadkami, w których trzeba zmniejszać liczbę liter z $n! + n$ do n, niewykonalne),
- bez silni się nie obędzie! Przykładowe złe słowo: $z = a^n c^{n+1} b^n$. Dlaczego analogiczny do powyższego dowód dla tego słowa nie działa? Problem pojawia się, gdy pompujemy jednocześnie a oraz b. Wiemy już, że w tym wypadku trzeba dobrać i tak, żeby „dobić" liczbę liter a (lub b) do liczby liter c (której nie zmieniamy). Problem polega na tym, że jak $|z|_c = |z|_a + 1$ to jedyny przypadek w którym dobijemy w z' liczbę liter a do c to taki, w którym do v wpadnie nam dokładnie jedno a i weźmiemy $i = 2$. Jeżeli do v wpadnie więcej liter a to zawsze można pompować i słowo dalej będzie należało do języka!