Rozwiązania zadań testowych autorstwa studentów: Jakub Czapiga, Piotr Lubczyński, Wojciech Musiał, Marcin Czerenko, Rafał Zemła, Jan Malina, Oleksandr Masliukivskyi, Karol Bieleń, Andrzej Kępka

3.1. Czy poniższe języki są bezkontekstowe?

[NIE] **(A)** $L_1 = \{a^i b^j c^k \mid i < j < k\}$

Podejrzewamy, że język ten nie jest bezkontekstowy, ponieważ posiada on silną zależność pomiędzy długościami swoich słów składowych. Korzystamy więc z lematu o pompowaniu języków bezkontekstowych. Oznaczamy przez n stałą z lematu. Pompujemy słowo $w = a^n b^{n+1} c^{n+2}$. *Po* sprawdzeniu wszystkich możliwych wariantów faktoryzacji słowa w dochodzimy do wniosku, że język ten nie jest bezkontekstowy, ponieważ pompowanie zarówno w górę jak i w dół nie daje nam słów należących do języka L_1.

[TAK] **(B)** $L_2 = \{a, b, c\}^* - L_1$

Przewidujemy, że jest to język bezkontekstowy. Język L_2 jest dopełnieniem języka L_1 z punktu (A). Po zaprzeczeniu warunku z języka L_1, możemy przedstawić język L_2 jako sumę trzech języków:

$$L_2 = \{a^i b^j c^k \mid i \geq j\} \cup \{a^i b^j c^k \mid j \geq k\} \cup (\{a, b, c\}^* - \mathbf{a^*b^*c^*})$$

Dwa pierwsze języki są bezkontekstowe. Przykładowa gramatyka dla $\{a^i b^j c^k \mid i \geq j\}$ to:

$S \to AC$
$A \to aAb \mid B$
$B \to aB \mid \varepsilon$
$C \to cC \mid \varepsilon$

Trzeci język jest językiem regularnym, gdyż różnica teoriomnogościowa dwóch języków regularnych też jest językiem regularnym. Cała suma jest więc językiem bezkontekstowym.

3.2. Czy poniższe języki są bezkontekstowe?

[NIE] **(A)** $L_1 = \{x \in \{a,b,c\}^* \mid |x|_a \neq |x|_b \wedge |x|_b \neq |x|_c \wedge |x|_c \neq |x|_a\}$

gdzie $|x|_a$ oznacza liczbę wystąpień symbolu a w słowie w.

Skorzystamy z faktu, że języki bezkontekstowe są zamknięte ze względu na przecięcie z językami regularnymi oraz z lematu Ogdena.

Niech język $R = \mathbf{a^*b^*c^*}$. R jest językiem regularnym (gramatyka jest trywialna). Jeśli L_1 byłby bezkontekstowy to: $M = L_1 \cap R = \{a^n b^m c^k \mid n \neq m,\ m \neq k,\ k \neq n\}$ też byłby bezkontekstowy, a nie jest.

Przykład ten jest omawiany na wykładzie, przytoczę tutaj dowód:

$M = \{a^n b^m c^k \mid n \neq m,\ m \neq k,\ k \neq n\}$. Przypuśćmy, że M jest

bezkontekstowy. Niech k będzie stałą z lematu Ogdena. Weźmy pod uwagę łańcuch $w = a^k b^{k+k!} c^{k+2k!}$. Założmy, że wyróżniamy pozycje symboli a, niech rozkład $w = xuyvz$ spełnia warunki lematu Ogdena. Jeżeli u lub v zawiera dwa różne symbole, to xu^2yv^2z nie należy do M. (Dla przykładu, jeśli u składa się z symboli a i b to xu^2yv^2z zawiera symbol b poprzedzający symbol a.) Jednak przynajmniej jedno spośród u i v musi zawierać symbole a, ponieważ tylko te symbole występują na wyróżnionych pozycjach. Zatem jeśli $v \in \{b\}^*$ lub $v \in \{c\}^*$, to u musi należeć do $\{a\}^+$. Jeżeli $v \in \{a\}^+$ to u musi należeć do $\{a\}^*$, gdyż inaczej jakiś symbol b lub c poprzedziłby symbol a. Rozważmy szczegółowo przypadek, gdy $v \in \{b\}^*$, a $u \in \{a\}^+$. (Pozostałe przypadki traktowane są w podobny sposób.) Niech $p = |u|$. Wtedy $1 \le p \le k$, czyli p dzieli $k!$. Niech q będzie liczbą całkowitą, taką, że $pq = k!$.
Wtedy $w' = xu^{2q+1}yv^{2q+1}z \in M$. Jednakże $u^{2q+1} = a^{p(2q+1)} = a^{2pq+p} = a^{2k!+p}$. Ponieważ xyz zawiera dokładnie $k - p$ symboli a, to w' zawiera $2k! + p + (k-p)$ czyli $2k! + k$ symboli a, czyli tyle samo, co symboli c, stąd $w' \notin M$ – sprzeczność. Podobna sprzeczność pojawia się w przypadku, gdy $v \in \{c\}^*$ lub $v \in \{a\}^+$. Zatem M nie jest językiem bezkontekstowym, więc także L_1 nie jest językiem bezkontekstowym.

[TAK] **(B)** $L_2 = \{x \in \{a,b,c\}^* \mid |x|_a \neq |x|_b \lor |x|_b \neq |x|_c \lor |x|_c \neq |x|_a \}$

Ten język można zinterpretować jako sumę teoriomnogościową trzech innych języków: pierwszego, w którym $|x|_a \neq |x|_b$, drugiego, w którym $|x|_b \neq |x|_c$ i ostatniego, w którym $|x|_c \neq |x|_a$. Każdy z nich jest bezkontekstowy – przykładowa gramatyka dla pierwszego:

$S \to aSb \mid bSa \mid abS \mid baS \mid Sba \mid Sab \mid aA \mid bB \mid cS \mid Sc$
$A \to aA \mid Aa \mid cA \mid Ac \mid \epsilon$
$B \to bB \mid Bb \mid cB \mid Bc \mid \epsilon$

Bardzo łatwo przerobić tę gramatykę tak, aby spełniała warunki dla drugiego i trzeciego języka, a więc je pominę.
Ostatecznie wniosek jest taki, że język L_1 jest bezkontekstowy, ponieważ języki bezkontekstowe są zamknięte ze względu na sumę teoriomnogościową.
Alternatywnie można stworzyć niedeterministyczny automat ze stosem- jak pamiętamy, wystarczy, gdy choć jedna „droga" prowadziła do akceptacji, żeby automat akceptował słowo.
Tak samo, jak z gramatyką powyżej, ten problem można podzielić na trzy podproblemy. Taki automat może w sposób niedeterministyczny „zgadnąć", który z nich doprowadzi do akceptacji; powiedzmy, że $|x|_a \neq |x|_b$. Wówczas wystarczy, że gdy automat przeczyta z wejścia literę a, doda ją na stos, zaś gdy b - zdejmie jedno a ze stosu. Jeśli zajdzie sytuacja, że napotkane zostanie b przy pustym stosie, trzeba odwrócić sytuację - b wrzuca na stos,

a zdejmuje. I tutaj niespodzianka, akceptacja następuje przez przeczytanie całego wejścia i NIEPUSTY stos.

3.3. Czy poniższe języki są bezkontekstowe?

[TAK] **(A)** $L_1 = \{a^m b^n c^p d^q \mid m+p=n+q;\ m,n,p,q > 0\}$

Język ten jest rozpoznawany przez automat ze stosem. Przy czytaniu każdego znaku na stos odkładany lub zdejmowany jest specjalny symbol. Automat działa w dwóch trybach, kontrolowanych przez stany. Najpierw jest w trybie *PLUS*, w którym czytając symbole *a* oraz *c* odkłada symbol stosowy na stos, zaś wczytując symbole *b* oraz *d* zdejmuje jeden symbol ze stosu. Jeśliby się okazało, że przy wczytywaniu symbolu *b* lub *d* stos jest już pusty, to należy zmienić tryb na *MINUS* i zacząć odkładać symbol stosowy przy wczytywaniu *b* oraz *d* zaś zdejmować jeden symbol ze stosu przy wczytywaniu *c*. Gdyby znów stos stał się pusty, zaś wejście jeszcze nie byłoby doczytane do końca, to należy zmienić tryb z powrotem na *PLUS*. Poza tym, w każdym trybie, każdej literce przypisany jest stan, który kontroluje, aby nie nastąpił przeplot różnych liter oraz, aby zachowana była odpowiednia kolejność liter. Zmiana litery w słowie wejściowym wywołuje zmianę stanu. Automat akceptuje słowo wejściowe w przypadku, gdy po przeczytaniu całego słowa stos jest pusty. Z faktu iż istnieje taki automat, wynika, że język ten jest bezkontekstowy.

[TAK] **(B)** $L_2 = \{a^m b^n c^p d^q \mid m+q=n+p;\ m,n,p,q>0\}$

Analogicznie jak w podpunkcie (A), z tym, że w trybie *PLUS* składamy na stos symbol stosowy przy czytaniu *a* oraz *d*, zaś zdejmujemy symbol stosowy przy czytaniu *b* oraz *c*, odpowiednio postępujemy odwrotnie w trybie *MINUS*.

[TAK] **(C)** $L_3 = \{a^m b^n c^p d^q \mid m+n=p+q;\ m,n,p,q>0\}$

Analogicznie jak w podpunkcie (A) z tą różnicą, że nie ma potrzeby stosowania dwóch trybów. Dla liter *a* i *b* na stos odkładany jest symbol specjalny. Natomiast dla liter *c* i *d* ze stosu ściągany jest symbol specjalny.

3.4. Czy poniższe języki są bezkontekstowe:

[NIE] **(A)** $L_1 = \{(a^n b)^m \mid m, n \geq 1\}$

Korzystamy z lematu o pompowaniu dla języków bezkontekstowych:

Niech *n* – stała z lematu. Pompujemy słowo $w = a^n b a^n b = xuyvz$

Rozważamy przypadek: $u = \varepsilon$, $y = \varepsilon$, $v = a^k b a^q$. Ponieważ $|uyv| \leq n$, więc $k+q \leq n-1$. Teraz badamy
$xu^2yv^2z = a^{n-k}(a^k b a^q)(a^k b a^q)a^{n-q}b = a^n b a^{k+q} b a^n b \notin L$,
gdyż $k+q \leq n-1 < n$.

Pozostałe przypadki są jeszcze prostsze. Z lematu wynika więc, że język nie jest bezkontekstowy.

[TAK] **(B)** $L_2 = \{a,b\}^* - L_1$

Konstruujemy niedeterministyczny automat ze stosem. Jeżeli słowo zaczyna się od b, to należy do języka. Badamy słowo $w = a^{i_1} b a^{i_2} b \ldots a^{i_n} b$. Słowo to może należeć do języka lub nie. Niedeterministycznie wybieramy dwie liczby j oraz k. Badamy, czy zachodzi $i_j = i_k$. Jeżeli tak – słowo odrzucamy, jeśli równość nie zachodzi – akceptujemy. Gdyby ostatnią literą słowa była a – akceptujemy.

3.5. Czy poniższe języki są bezkontekstowe?

[NIE] **(A)** $L_1 = \{ r\#s\#t \mid r, s, t \in \{a,b\}^*, |r| = |s| = |t| \}$

- korzystając z lematu o pompowaniu dla języków bezkontekstowych, rozważamy słowo $w = a^n\#a^n\#a^n = xuyvz$, gdzie n jest stałą z lematu.

Wówczas uv może zawierać jedynie litery a, aby przy pompowaniu nie pojawiły się w słowie więcej niż dwa znaki #, podczas gdy część y może zawierać maksymalnie jeden znak #, aby był spełniony warunek $|uyv| < n$.

Z tego wynika, że pompując to słowo, zawsze przynajmniej jeden ciąg liter a nie będzie zwiększany, a więc warunek $|r| = |s| = |t|$ nie będzie spełniony, czyli słowo nie będzie należało do L_1, z czego wynika, że nie jest on bezkontekstowy.

[TAK] **(B)** $L_2 = \{ rst \mid r, s, t \in \{a,b\}^*, |r| = |s| = |t| \}$

- ponieważ można stworzyć gramatykę bezkontekstową generującą ten język:

$S \to aA \mid bA \mid \varepsilon$

$A \to aB \mid bB$

$B \to aS \mid bS$

3.6. Czy poniższe języki są bezkontekstowe?

[NIE] **(A)** $L_1 = \{x \in \{a, b\}^* \mid |x|_a = |x|_b^2\}$ gdzie $|x|_a$ oznacza liczbę wystąpień symbolu a w słowie x.

$L_1 \cap$ **a*b*** = $\{a^{i^2}b^i \mid i \geq 0\}$. Jeśli $\{a^{i^2}b^i \mid i \geq 0\}$ nie jest bezkontekstowy, to L_1 też, gdyż języki bezkontekstowe są zamknięte względem przecięcia z językami regularnymi. Dla $\{a^{i^2}b^i \mid i \geq 0\}$ stosujemy lemat o rozrastaniu dla języków bezkontekstowych. Niech n będzie stałą z lematu. Próbujemy znaleźć faktoryzację $xuyvz$ dla słowa $w = a^{n^2}b^n$. W najbardziej ewentualnie sprzyjającym przypadku $v = b$, $u = a^{n-1}$, czyli $xu^2yv^2z = a^{n^2+(n-1)}b^{n+1}$, ale $n^2+n-1 < (n+1)^2$ czyli $xu^2yv^2z \notin \{a^{i^2}b^i \mid i \geq 0\}$, a więc $\{a^{i^2}b^i \mid i \geq 0\}$ nie jest bezkontekstowy, stąd także L_1 nie jest bezkontekstowy.

[TAK] **(B)** $L_2 = \{x \in \{a, b\}^* \mid |x|_a = |x|_b\}$

Można stworzyć automat ze stosem, który przy pustym stosie składa bieżący czytany symbol na stos. Teraz jeśli symbol na stosie jest identyczny z bieżącym symbolem czytanym, to dokłada przeczytany symbol do stosu, zaś czytając symbol przeciwny temu na wierzchołku stosu, zdejmuje jeden symbol ze stosu. Akceptacja następuje przez pusty stos i puste wejście.

3.7. Czy poniższe języki są bezkontekstowe?

[TAK] **(A)** $L_1 = \{x \in \{a,b,c\}^* \mid ||x|_b - |x|_a| \leq 2 \vee ||x|_c - |x|_a| \leq 2\}$, gdzie $|x|_a$ oznacza liczbę symboli a w słowie x.

Automat ze stosem w niedeterministyczny sposób wybiera, którą nierówność będzie badał. Jeśli wybierze nierówność $-2 \leq |x|_a - |x|_b \leq 2$, to ignoruje czytane litery c. Jeśli stos jest pusty, to składa czytaną literę a lub b na stosie. Jeśli czytana litera jest identyczna z leżącą na wierzchołku stosu, to dopisuje tę literę do stosu. Jeśli czytana litera a lub b jest inna, niż leżąca na wierzchołku stosu, to zdejmuje tę literę z wierzchołka stosu. Jeśli wejście jest już puste, bo zostało przeczytane do końca, to stos powinien być pusty lub zawierać jedną lub dwie litery. To można sprawdzić używając dwóch dodatkowych stanów i zdejmując po jednej literze zmieniając przy tym stan. Jeśli potwierdzimy zachodzenie sprawdzanej nierówności, akceptujemy, jeśli nie odrzucamy. Jeśli na początku zostanie wybrana druga nierówność postępujemy podobnie, tym razem ignorując czytane litery b.

[NIE] **(B)** $L_2 = \{ x \in \{a,b,c\}^* \mid ||x|_b - |x|_a| \leq 2 \wedge ||x|_c - |x|_a| \leq 2\}$

Przypuśćmy bowiem, że język L_2 jest bezkontekstowy i niech n będzie stałą z lematu o rozrastaniu. Weźmy pod uwagę łańcuch $w = a^n b^n c^n = xuyvz$ o długości równej $3n \geq n$. Rozpatrujemy przypadek, gdy uv zawiera wyłącznie symbole a. Gdy będziemy zwiększać liczbę symboli a, to napompowane słowo nie będzie należeć do naszego języka, ponieważ różnica liter a w słowie będzie rosła i przekroczy wartość 2 względem pozostałych symboli. Podobna sytuacja wystąpi, gdy uv zawiera wyłącznie symbole b lub wyłącznie symbole c. Przypuśćmy teraz, że u zawiera symbole a i v zawiera symbole b lub u zawiera symbole b i v zawiera symbole c. W trakcie pompowania otrzymamy słowo, które nie będzie spełniało jednego z warunków należenia do L_2. Kolejno bierzemy przypadki, w których u zawiera tylko symbole a i v zawiera symbole a i b albo u zawiera tylko symbole b i v zawiera symbole b i c. Podobnie jak poprzednio w czasie pompowania otrzymamy słowo nie spełniające jednego z warunków należenia do języka, czyli słowo/a nie należące do języka. Analogicznie postępujemy z dwoma ostatnimi przypadkami tj. u zawiera symbole a i b i v zawiera symbole b albo u zawiera symbole b i c i v zawiera symbole c. Tutaj także po napompowaniu otrzymamy słowo nie spełniające jednego z warunków należenia do języka. Po rozpatrzeniu wszystkich możliwości wnioskuję, że język L_2 nie jest językiem bezkontekstowym.

3.8. Czy poniższe języki są bezkontekstowe?

[TAK] **(A)** $L_1 = \{a^n b^m c^{|n-m|} \mid n, m \geq 0\}$

1) Można napisać gramatykę bezkontekstową:

S -> B | AC
A -> aAb | ε
B -> aBc | A
C -> bCc | ε

2) Można utworzyć deterministyczny automat ze stosem, który ma 4 stany: q_a, q_b, q_c *i* q_{toggle}. Automat akceptuje przez pusty stos w stanie q_c.
Zaczynamy w stanie q_a i przy przetwarzaniu a odkładamy na stos pewien symbol x dla każdego przetworzonego a.
Przy pierwszym napotkanym b przechodzimy do stanu q_b, w którym dla każdego przetworzonego b zdejmujemy ze stosu symbol x. W przypadku pustego stosu przełączamy się ze zdejmowania na dokładanie symbolu x na stos, czyli wchodzimy w stan q_{toggle}.
Przy pierwszym napotkanym c przechodzimy do stanu q_c, w którym dla każdego przetworzonego c zdejmujemy ze stosu symbol x. Akceptujemy przy pustym stosie gdy wejście będzie przeczytane do końca.
Jeżeli stos będzie pusty a dalej będziemy na wejściu napotykali c lub całe wejście będzie przeczytane, a stos nie będzie pusty lub w stanie q_b przeczytamy a lub w stanie q_c przeczytamy a lub b to odrzucamy to słowo.
Skoro istnieje deterministyczny automat ze stosem akceptujący język L_1 to jest on bezkontekstowy.

[NIE] **(B)** $L_2 = \left\{a^n b^m c^{\left\lfloor \frac{n}{m} \right\rfloor} \,\middle|\, n, m > 0,\ \left\lfloor \frac{n}{m} \right\rfloor > 0\right\}$

Załóżmy, że język L_2 jest bezkontekstowy i p jest stałą z lematu o pompowaniu. Rozważmy słowo $w = a^{p^2} b^p c^p = uvxyz$, $p > 0$ o długości $2p + p^2 > p$

Słowo $w = a^{p*p} b^p c^p \in L_2$ ponieważ $\frac{p^2}{p} = p$

Rozpatrujemy 5 przypadków:

1) *vy* zawiera tylko litery a, wtedy po napompowaniu otrzymujemy $a^{p^2+l} b^p c^p \notin L_2$ ponieważ $\left\lfloor \frac{p^2+l}{p} \right\rfloor \neq p$ gdy $l \geq p$
2) *vy* zawiera tylko litery b, wtedy po napompowaniu otrzymujemy $a^{p^2} b^{p+l} c^p \notin L_2$ ponieważ $\left\lfloor \frac{p^2}{p+l} \right\rfloor \neq p$ gdy $l > 0$
3) *vy* zawiera tylko litery c, wtedy po napompowaniu otrzymujemy $a^{p^2} b^p c^{p+l} \notin L_2$ ponieważ $\left\lfloor \frac{p^2}{p} \right\rfloor \neq p + l$ gdy $l > 0$
4) *vy* zawiera litery *a* i *b*, bez nadmiernej utraty ogólności można przyjąć, że liczba pompowanych liter *a* jest taka sama, jak liczba pompowanych liter *b*; wtedy po napompowaniu otrzymujemy $a^{p^2+l} b^{p+l} c^p \notin L_2$ ponieważ $\left\lfloor \frac{p^2+l}{p+l} \right\rfloor \neq p$ gdy $l > 0$
5) *vy* zawiera litery *b* i *c*, bez nadmiernej utraty ogólności można przyjąć, że liczba pompowanych liter *b* jest taka sama, jak liczba pompowanych liter *c*; wtedy po napompowaniu otrzymujemy $a^{p^2} b^{p+l} c^{p+l} \notin L_2$ gdyż $\left\lfloor \frac{p^2}{p+l} \right\rfloor \neq p + l$ gdy $l > 0$

W żadnym z powyższych przypadków nie można „napompować” tego słowa, więc język L_2 nie jest bezkontekstowy.

3.9. Czy poniższe języki są bezkontekstowe?

[TAK] **(A)** $L_1 = \{a^n b^m | n, m \geq 0, m \geq n, m - n$ jest parzyste$\}$

Można zbudować automat ze stosem. Odczytując ze słowa wejściowego symbol *a* wkładamy specjalny symbol na stos. W momencie pojawienia się pierwszego symbolu *b* zapamiętujemy ten fakt poprzez zmianę stanu. W tym stanie odczytując symbol *b* ściągamy specjalny symbol ze stosu. Jeśli słowo wejściowe się skończyło a stos nie jest pusty to znaczy, że nie jest spełniony warunek $m \geq n$. Wówczas odrzucamy. Jeśli stos jest pusty to $m \geq n$ jest spełnione. Teraz, jeśli stos jest pusty, a w słowie wejściowym wciąż są znaki *b*, to naprzemiennie (dla kolejnych symboli *b*) wkładamy i zdejmujemy symbol ze stosu, odpowiednio zmieniając stany. Jeśli po odczytaniu wszystkich znaków z wejścia stos jest pusty to $m - n$ jest parzyste, wtedy akceptujemy. Jeśli zaś na stosie pozostał jakiś symbol to $m - n$ nie jest parzyste, wówczas odrzucamy.

[NIE] **(B)** $L_2 = \{a^n b^m | n, m > 0, (n \geq m^2 \vee m \geq n^2)\}$

Stosujemy lemat o pompowaniu. Zakładamy, że

k – stała z lematu, $w = a^k b^{k^2} = xuyvz$

Istnieją 3 możliwości:

1) $uyv = a^i, i \leq k$ – pompując liczba symboli *b* pozostanie niezmienna, więc rozrośnie się tylko jeden człon słowa
2) $uyv = b^i, i \leq k$ – pompując liczba symboli *a* pozostanie niezmienna, więc rozrośnie się tylko jeden człon słowa
3) $uyv = a^i b^j, i + j \leq k$ – to najbardziej prawdopodobnym przypadkiem udanego pompowania jest:
   $u = a, y = \varepsilon, v = b^{k-1}$
   Wtedy $xu^2yv^2z = a^{k-1}a^2b^{2(k-1)}b^{k^2-(k-1)} = a^{k+1}b^{(2k-2)+k^2-(k-1)} = a^{k+1}b^{k^2+k-1} \notin L_2$
   gdyż $k^2 + k - 1 < (k+1)^2$

Wobec powyższego język L_2 nie jest bezkontekstowy.