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

4.1. Niech $L \subseteq \Sigma^*$. Wybierz prawdziwe stwierdzenia:

[TAK] **(A)** Oznaczmy przez $Enlarge(L) = \{x \in \Sigma^* \mid \exists y \in \Sigma^*, xy \in L\}$. Jeśli L jest językiem regularnym, to także *Enlarge(L)* jest językiem regularnym.

Jeżeli $xy \in L$ oraz L jest regularny, to oznacza, że istnieje automat skończony akceptujący xy.

Enlarge(L) jest to zbiór wszystkich przedrostków wszystkich słów z języka L.

To oznacza, że aby otrzymać automat akceptujący *Enlarge(L)* należy zmienić wszystkie stany pośrednie automatu L na stany akceptujące.

[NIE] **(B)** Oznaczmy przez $Flip(L) = \{ww^R \mid w \in L\}$. Jeśli L jest językiem regularnym, to także *Flip(L)* jest językiem regularnym.

Niech L = **a*b**, wtedy $Flip(L) = \{a^n bba^n \mid n \geq 0\} = M$.

Przypuśćmy, że M jest językiem regularnym i niech k będzie stałą z lematu o pompowaniu języków regularnych.

Rozważamy słowo $w = a^k bba^k = xuz$. Słowo w ma długość równą $2k+2 > k$.

Wówczas u może zawierać od jednej do maksymalnie k liter a i występować przed b (przypadek (1)) lub u może zawierać od jednej do maksymalnie dwóch liter b (przypadek (2)), lub u może zawierać od jednej do k liter a i występować po b (przypadek (3)).

Wybranie u w inny sposób spowoduje, że przy rozrastaniu u^i symbole zaczną występować w złej kolejności, np. *aabaab…*

Rozważamy łańcuch xu^2z. W przypadku (1) ciąg liter a występujący przed literami b będzie zawierał co najmniej $k+1$ i co najwyżej $2k$ liter, podczas gdy ciąg liter a występujący po literach b wciąż ma długość k, a więc wtedy xu^2z nie należy do języka. Analogicznie rozważamy przypadek (3). W przypadku (2) łańcuch xu^2z będzie zawierał od trzech do czterech liter b, a więc nie będzie należał do języka M.

Zatem xu^2z w żadnym możliwym przypadku nie należy do języka M, a więc język ten nie może być regularny.

[NIE] **(C)** Oznaczmy przez $PartPart(L) = \{xz \mid |x| = |z| \wedge \exists y \in \Sigma^*, xyz \in L\}$. Jeśli L jest językiem regularnym, to także *PartPart(L)* jest językiem regularnym.

Niech L = **a*#b***, a $PartPart(L) \cap$ **a*b*** $= \{a^n b^n \mid n \geq 0\} = M$. M nie jest regularny z lematu o pompowaniu dla języków regularnych (dowód analogiczny jak w (B), był na wykładzie). Języki regularne są zamknięte ze względu na przecięcie, a więc *PartPart(L)* również nie jest regularny.

4.2. Niech $L \subseteq \Sigma^*$ będzie językiem. Wybierz prawdziwe stwierdzenia:

[TAK] **(A)** Oznaczamy przez $Len(L) = \{x \mid \exists y \in L,\ |x|=|y|\}$. Jeśli L jest językiem regularnym, to także $Len(L)$ jest językiem regularnym.

Jeżeli L jest językiem regularnym to istnieje *DFA* akceptujący dany język. Dla $Len(L)$ istnieje *NFA*, który akceptuje takie słowo, ze względu na warunek o takiej samej długości słowa dla obu języków, jest w stanie dojść do stanu akceptującego co *DFA* dla języka L i zaakceptować język $Len(L)$. *NFA* dla $Len(L)$ powstaje poprzez dopisanie wszystkich symboli z alfabetu Σ nad każdą z krawędzi *DFA* dla L.

[NIE] **(B)** Oznaczamy przez $DropMiddle(L) = \{xy \mid |x|=|y| \wedge \exists a \in \Sigma\ xay \in L\}$. Jeśli L jest językiem regularnym, to także $DropMiddle\ (L)$ jest językiem regularnym.

Kontrprzykład:
Niech $L=$**a*# b***
$M = DropMiddle(L) \cap \mathbf{a^*b^*} = \{\ a^n b^n \mid n \geq 0\}$
Ponieważ M nie jest regularny (było na wykładzie) to $DropMiddle(L)$ też nie jest regularny, bo języki regularne są zamknięte ze względu na przecięcie.

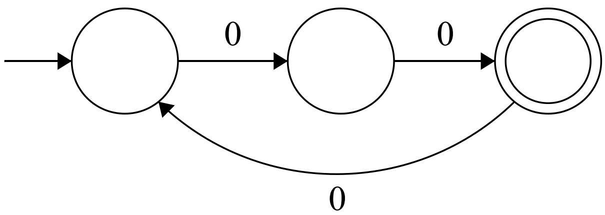
[NIE] **(C)** Oznaczamy przez $DropMiddlePart(L)=\{xz \mid \exists y \in \Sigma^*,\ |x|=|y|=|z|,\ xyz \in L\}$. Jeśli L jest językiem regularnym, to także $DropMiddlePart(L)$ jest językiem regularnym.

Kontrprzykład:
Niech $L=$**a*# b***
$M = DropMiddlePart(L) \cap \mathbf{a^*b^*} = \{a^n b^n \mid n \geq 0\}$
Ponieważ M nie jest regularny (było na wykładzie) to $DropMiddlePart(L)$ też nie jest regularny, bo języki regularne są zamknięte ze względu na przecięcie.

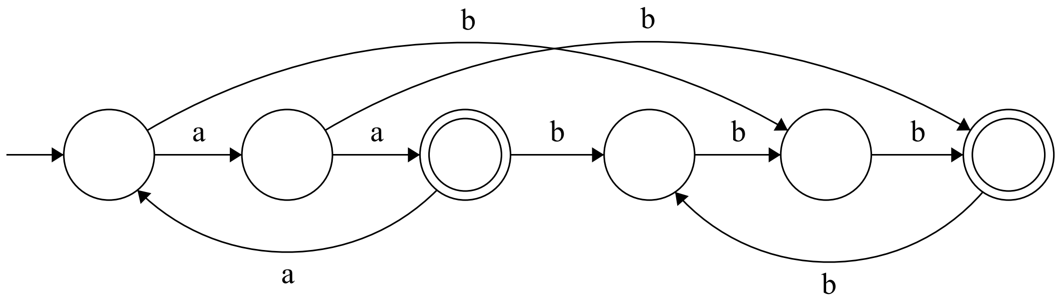
4.3. Wybierz prawdziwe stwierdzenia:

[TAK] **(A)** Dany jest język L nad alfabetem $\Sigma = \{0\}$. Oznaczamy $Plus(L) = \{a^m b^n \mid 0^{m+n} \in L\}$. Jeśli język L jest regularny, to także $Plus(L)$ jest językiem regularnym.

Dla każdego z języków regularnych L nad alfabetem $\Sigma = \{0\}$ istnieje deterministyczny automat skończony akceptujący go. Każdy z automatów akceptujących język L można zmodyfikować, tak aby akceptował on język *Plus(L)*. W tym celu przejścia dla symbolu *0* powinny zostać zastąpione przejściami dla symbolu *a*, następnie należy stworzyć bliźniaczy automat jednak z przejściami dla symboli *b.* Ze stanów automatu dla symboli *a* należy poprowadzić przejścia przechodzące do tych stanów w automacie dla symboli *b*, do których prowadziły przejścia w automacie dla symboli *a*, co zapewni, że sumaryczna długość ciągów symboli *a* i *b* będzie równa długości słów języka L i że symbole *a* będą poprzedzały symbole *b*. Np. L = **00(000)*** jest językiem regularnym. DFA dla L:



DFA dla *Plus(L):*



[NIE] **(B)** Dany jest język L nad alfabetem $\Sigma = \{0\}$. Oznaczamy $Minus(L) = \{a^m b^n \mid 0^{m-n} \in L\}$. Jeśli język L jest regularny, to także *Minus(L)* jest językiem regularnym.

Niech L = **0***. Wtedy $Minus(L) = \{a^m b^n \mid m,n \geq 0,\ m \geq n\}$. Stosując lemat o rozrastaniu; niech k będzie stałą z lematu. Rozważamy słowo $w = a^k b^k = xuz$. Słowo w ma długość $2k > k$. Część u może zawierać od jednej do k liter *a*. Jakkolwiek nie pompowalibyśmy słowa *w,* zawsze xu^0z będzie miało mniej liter *a* niż liter *b.* Więc *Minus(L)* nie jest regularny.

[NIE] **(C)** Dany jest skończony alfabet Σ z ustalonym porządkiem całkowitym. Dla słowa $x \in \Sigma^*$ oznaczamy przez *sort(x)* słowo otrzymane przez posortowanie liter w porządku rosnącym. Na przykład jeśli $a \prec b \prec c$ to *sort(abacbaa) = aaaabbc*. Dla języka $L \subseteq \Sigma^*$ oznaczamy $Sort(L) = \{sort(x) \mid x \in L\}$. Jeśli język L jest regularny, to także *Sort(L)* jest językiem regularnym.

Niech L = **(ab)***. Wtedy $Sort(L) = \{a^n b^n \mid n \geq 0\}$, a wiemy że to nie jest język regularny z lematu o pompowaniu języków regularnych (było na wykładzie).

4.4. Czy poniższe języki są regularne?

[NIE] **(A)** $L_1 = \{1^k y \mid y \in \{0,1\}^*, y \text{ zawiera co najwyżej } k \text{ jedynek dla } k \geq 1\}$

Ten przypadek rozpatrujemy podobnie jak podpunkt (C). Przypuszczamy dla dowodu nie wprost, że język ten jest regularny i niech *n* będzie stałą z lematu o rozrastaniu języków regularnych. Rozważamy słowo $w = 1^n 0 1^n = xyz$. Zgodnie z twierdzeniem lematu o pompowaniu $|xy| \leq n$. Stąd też wynika, że *y* będzie zbudowane z jedynek (co najmniej jednej). Rzecz jasna chodzi tu o jedynki z pierwszej części słowa. W sytuacji, gdy napompujemy słowo xy^0z, gdzie *y* zawiera dowolną liczbę jedynek z zakresu *1* do *n* to otrzymane słowo nie będzie należeć do języka L_3, ponieważ po lewej stronie (względem *0*) będzie mniej jedynek niż po prawej stronie. Wynika z tego, że język nie jest regularny.

[TAK] **(B)** $L_2 = \{1^k y \mid y \in \{0,1\}^*, y \text{ zawiera co najmniej } k \text{ jedynek dla } k \geq 1\}$

Tak, ponieważ możemy dla tego języka skonstruować wyrażenie regularne: $\mathbf{10}^*\mathbf{1}(\mathbf{1}|\mathbf{0})^*$

[NIE] **(C)** $L_3 = \{1^k 0 y \mid y \in \{0,1\}^*, y \text{ zawiera co najwyżej } k \text{ jedynek dla } k \geq 1\}$

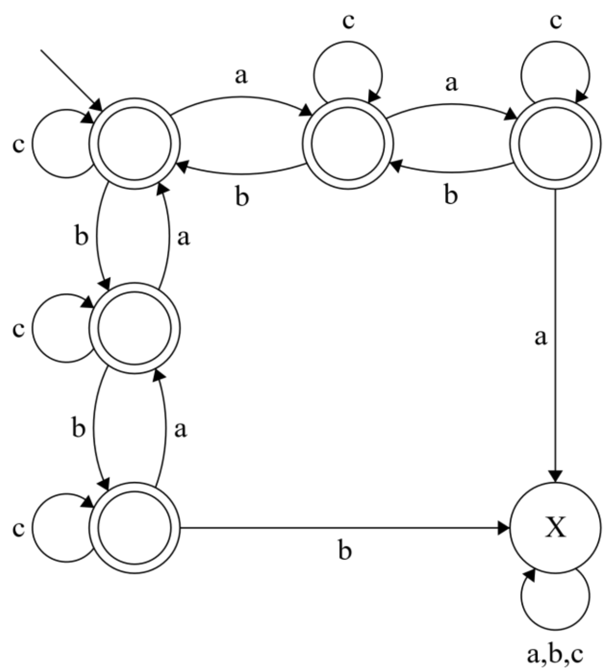
Przypuśćmy dla dowodu nie wprost, że język ten jest regularny i niech *n* będzie stałą z lematu o rozrastaniu języków regularnych. Rozważamy słowo $w = 1^n 0 1^n = xyz$. Zgodnie z twierdzeniem lematu o pompowaniu $|xy| \leq n$. Stąd też wynika, że *y* będzie zbudowane z jedynek (co najmniej jednej). Rzecz jasna chodzi tu o jedynki z pierwszej części słowa. W sytuacji, gdy napompujemy słowo xy^0z, gdzie *y* zawiera dowolną liczbę jedynek z zakresu *1* do *n* to otrzymane słowo nie będzie należeć do języka L_3, ponieważ po lewej stronie (względem *0*) będzie mniej jedynek niż po prawej stronie. Wynika z tego, że język nie jest regularny.

[NIE] **(D)** $L_4 = \{1^k 0 y \mid y \in \{0,1\}^*, y \text{ zawiera co najmniej } k \text{ jedynek dla } k \geq 1\}$

Przypuśćmy dla dowodu nie wprost, że język ten jest regularny i niech *n* będzie stałą z lematu o rozrastaniu języków regularnych. Rozważamy słowo $w = 1^n 0 1^{n+1} = xyz$. Zgodnie z twierdzeniem lematu o pompowaniu $|xy| \leq n$. Stąd też wynika, że *y* będzie miało same jedynki. Rzecz jasna chodzi tu o jedynki z pierwszej części słowa. Jeżeli napompujemy *y* z pojedynczym symbolem *1* liczbą większą lub równą 3 (uzyskamy z napompowanego *y* 3 lub więcej symboli *1*) dojdzie do sytuacji, że liczba jedynek po lewej stronie będzie większa od liczby jedynek po prawej stronie (względem *0*). W efekcie otrzymamy słowo nie należące do języka. Stąd wynika, że L_4 nie jest językiem regularnym.

4.5. Czy poniższe języki są regularne?
($|x|_a$ oznacza liczbę symboli *a* w słowie *x*)

[TAK] **(A)** $L_1 = \{w \in \{a, b, c\}^* \mid$ dla każdego prefiksu *x* słowa *w* zachodzi $| \, |x|_b - |x|_a| \leq 2\}$



Istnieje deterministyczny automat skończony, który akceptuje język L_1, więc język L_1 jest regularny.

[NIE] **(B)** $L_2 = \{w \in \{a, b, c\}^* \mid$ dla każdego prefiksu *x* słowa *w* zachodzi $| \, |x|_b - |x|_a| \leq 2 \vee | \, |x|_c - |x|_a| \leq 2\}$

Niech $M = L_2 \cap \mathbf{c}^*(\mathbf{ab})^*\mathbf{bbb} = \{ c^i(ab)^j bbb \mid i - 2 \leq j \leq i + 2 \}$
Jeśli L_2 byłby językiem regularnym to *M* też musiałby być regularny, bo języki regularne są zamknięte ze względu na przecięcie. Zakładamy dla celów dowodu z wykorzystaniem lematu o rozrastaniu dla języków regularnych, że M jest regularny. Niech *p* będzie stałą z lematu. Rozważamy słowo $w = c^p(ab)^p bbb = xyz$ gdzie $|w| \geq p,\ |xy| \leq p,\ y \neq \varepsilon$.
Mamy tylko jeden przypadek: *y* zawiera od jednej do *p* liter *c*.
W tym wypadku nasze słowo po napompowaniu ma postać $c^{p+l}(ab)^p bbb$.
Ale $c^{p+l}(ab)^p bbb \notin M$ dla $l > 2,$ ponieważ nieprawda, że $p + l - 2 \leq p$ dla $l > 2$. Wynika z tego, że język *M* nie jest regularny, więc język L_2 również nie jest regularny.

Dlaczego poniższe rozwiązanie jest błędne i nie można przedstawić języka L_2 jako sumy dwóch języków regularnych?

UWAGA! błędne rozwiązanie:
[TAK] **(B)** $L_2 = \{w \in \{a, b, c\}^* \mid$ dla każdego prefiksu x słowa w zachodzi $|\,|x|_b - |x|_a| \leq 2 \;\lor\; |\,|x|_c - |x|_a| \leq 2\}$

> Język L_2 można przedstawić jako sumę teoriomnogościową dwóch języków, jednego z warunkiem $|\,|x|_b - |x|_a| \leq 2$ oraz drugiego z warunkiem $|\,|x|_c - |x|_a| \leq 2$. Każdy z tych języków jest regularny (por. punkt (A)). Wobec tego, z uwagi na to, że języki regularne są zamknięte ze względu na sumę teoriomnogościową, język L_2 jest też regularny.

Rozważmy dwie formuły:
F1: $\forall x\, p(x)$ lub $\forall x\, q(x)$
F2: $\forall x\, [\, p(x)$ lub $q(x)\,]$

Mamy:
F1 => F2 ale w ogólności nie zachodzi F2 => F1.

Na tej samej zasadzie mamy język L_2 zdefiniowany w pkt. 4.5 (B) (odpowiednik formuły F2) oraz język L_x traktowany jako suma dwóch języków (odpowiednik formuły F1).
L_x jest sumą dwóch języków regularnych, a więc sam jest regularny.
Natomiast L_2 jest różny od języka L_x (L_2 jest nadzbiorem języka L_x). Dowód z przecięciem pokazuje, że L_2 nie jest regularny.

Weźmy słowo: $bbbaaa$

$L_2 = \{w \in \{a, b, c\}^* \mid \forall$ prefiksu x słowa w zachodzi $|\,|x|_b - |x|_a| \leq 2 \;\lor\; |\,|x|_c - |x|_a| \leq 2\}$

$L_{x1} = \{w \in \{a, b, c\}^* \mid \forall$ prefiksu x słowa w zachodzi $|\,|x|_b - |x|_a| \leq 2\}$
$L_{x2} = \{w \in \{a, b, c\}^* \mid \forall$ prefiksu x słowa w zachodzi $|\,|x|_c - |x|_a| \leq 2\}$

$L_x = L_{x1} \cup L_{x2}$

Zauważmy że $bbbaaa \notin L_{x1}$, gdyż istnieje prefiks bbb, który nie spełnia warunku z L_{x1} (ma zbyt dużo liter b); $bbbaaa \notin L_{x2}$ gdyż istnieje prefiks $bbbaaa$, który nie spełnia warunku z L_{x2} (ma zbyt mało liter c). Wobec tego całe słowo $bbbaaa \notin L_x$.

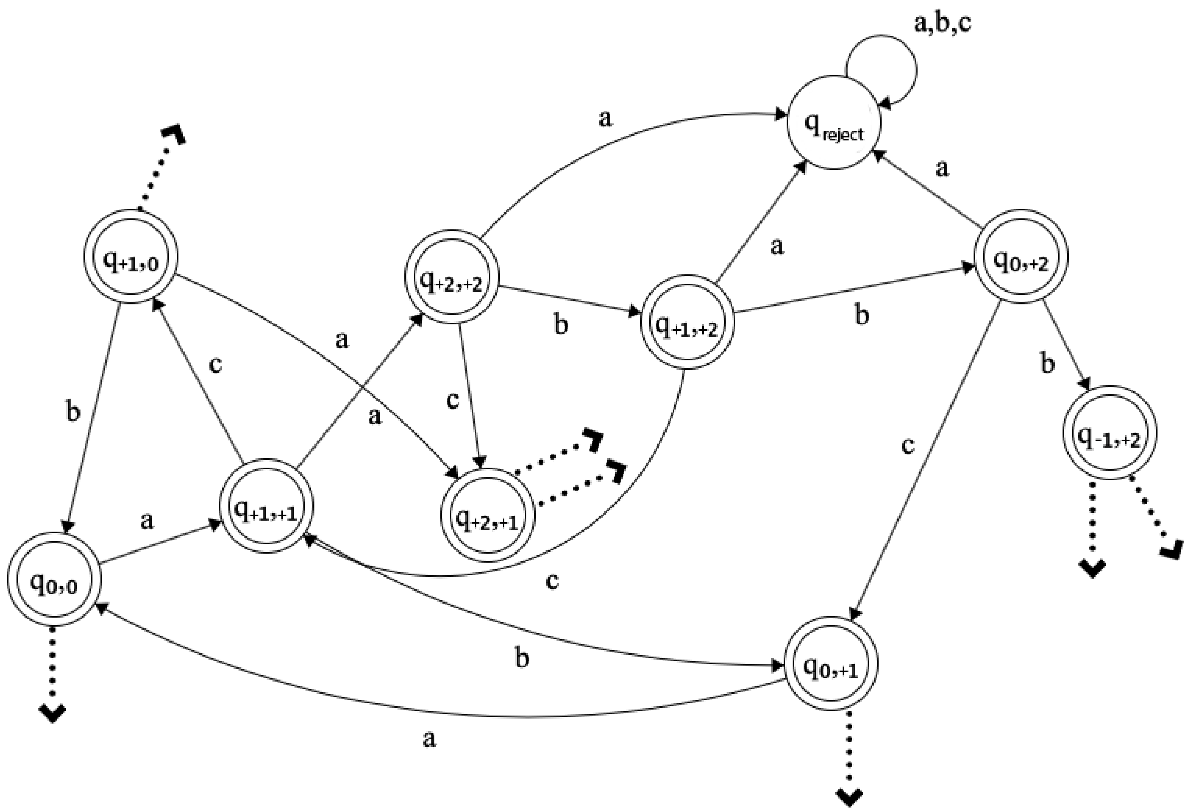
Jednak $bbbaaa \in L_2$, gdyż wprawdzie prefiks bbb nie spełnia pierwszej składowej alternatywy warunku z L_2 (ma zbyt dużo liter b), ale za to spełnia drugą składową (liczba liter a jest taka sama, jak liczba liter c); prefiks $bbbaaa$ nie spełnia drugiej składowej alternatywy warunku z L_2 (ma zbyt mało liter c), ale za to spełnia pierwszą składową (liczba liter b jest taka sama, jak liczba liter a).

Wniosek stąd taki, że język L_x różni się od języka L_2, a z całości rozumowania wynika, że język L_x jest językiem regularnym, zaś język L_2 nie jest regularny.

4.6. Czy poniższe języki są regularne: ($|x|_a$ oznacza liczbę wystąpień symboli a w łańcuchu x)

[TAK] **(A)** $L_1 = \{w \in \{a, b, c\}^* \mid$ dla każdego prefiksu x słowa w zachodzi $||x|_b - |x|_a| \le 2 \land ||x|_c - |x|_a| \le 2\}$

Podejrzewam, że ten język jest regularny. Więc próbuję skonstruować automat skończony. Niech automat ten koduje w stanie różnicę między liczbą symboli a oraz b i różnicę między liczbą symboli a oraz c. Stan $q_{+2,-1}$ oznacza, że w aktualnie przeczytanym prefiksie symboli a jest o 2 więcej od symboli b oraz, że symboli a jest o 1 mniej niż symboli c. Jeśli któraś z różnic wyniesie 3 lub -3 to automat przechodzi do stanu odrzucającego q_{reject}. Fragment automatu:



Stąd, język L_1 jest regularny.

[NIE] **(B)** $L_2 = \{w \epsilon \{a,b,c\}^* \mid |w|_a = |w|_b$ oraz dla każdego prefiksu x słowa w zachodzi $||x|_c - |x|_a| \leq 2\}$

W tym przypadku nie da się narysować automatu akceptującego powyższy język.
Niech $M = L_2 \cap$ **(ac)*b*** $= \{(ac)^n b^n \mid n \geq 0\}$. Z lematu o rozrastaniu dla języków formalnych, niech p - stała z lematu. Pompujemy słowo $w = (ac)^p b^p = xyz$, część y musi być zawarta w $(ac)^p$, gdyż $|xy| \leq p$. Część y nie może zawierać pojedynczej litery a ani pojedynczej litery c, bo wtedy pompowanie naruszy warunek $|x_c - x_a| \leq 2$. Jeżeli pompujemy część y, która zawiera a i c żeby spełnić powyższy warunek, to zostanie naruszony drugi z warunków ($|w|_a = |w|_b$). Więc język M nie jest regularny, stąd język L_2 też nie jest regularny, bo języki bezkontekstowe są zamknięte ze względu na przecięcie.

4.7. Czy następujące języki są regularne?

[NIE] **(A)** $L_1 = \{xx^R w \mid x,w \in \{a,b\}^+\}$

Weźmy język $M = L_1 \cap (ab)^*(ba)^*a = \{(ab)^k(ba)^m a \mid m \geq k\}$
Stosujemy lemat o pompowaniu dla języków regularnych:
n – stała z lematu
$w = xyz = (ab)^n(ba)^n a$
$|xy| \leq n$
$\forall k \geq 0\ xy^k z \in M$
xy musi leżeć w całości w części $(ab)^n$, żeby spełniać warunek $|xy| \leq n$. Mamy więc sytuację:
$i + j \leq n$
$x = (ab)^i$
$y = (ab)^j$
$z = (ab)^{n-i-j}(ba)^n a$
Zgodnie z lematem do języka musiałoby należeć też słowo:
$xy^2z = (ab)^i(ab)^{2j}(ab)^{n-i-j}(ba)^n a = (ab)^{n+j}(ba)^n a$
Z powyższego wynika, że M nie jest regularny, więc L_1 też nie jest regularny.

[TAK] **(B)** $L_2 = \{xwx^R \mid x,w \in \{a,b\}^+\}$

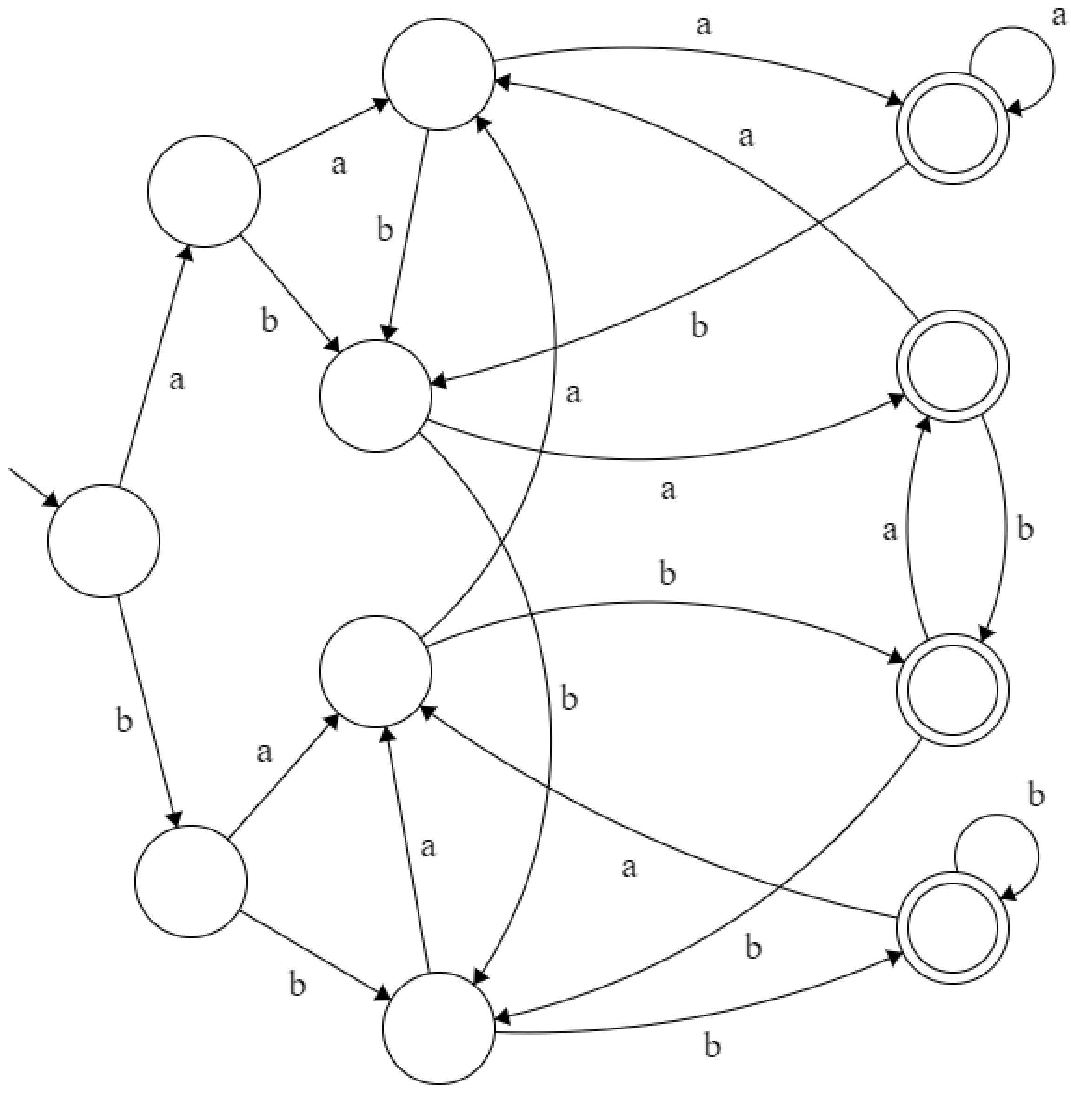
Wyrażenie regularne: $\mathbf{a(a|b)^+a \mid b(a|b)^+b}$

[NIE] **(C)** $L_3 = \{w \in \{a,b\}^* \mid \exists x,y \in \{a,b\}^*, w = xaay, |x| = |y|\}$

Stosujemy lemat o pompowaniu dla języków regularnych:
n – stała z lematu
$w = xyz = b^n aab^n$

$|xy| \leq n$
$\forall k \geq 0 \;\; xy^k z \in L_3$
xy musi leżeć w całości w części pierwszego b^n, żeby spełniać warunek $|xy| \leq n$. Mamy więc sytuację:
$i + j \leq n$
$x = b^i$
$y = b^j$
$z = b^{n-i-j}aab^n$
Zgodnie z lematem do języka musiałoby należeć też słowo:
$xy^2z = b^i b^j b^j b^{n-i-j}aab^n = b^{n+j}aab^n$

[TAK] **(D)** $L_4 = \{w \in \{a,b\}^* \mid \exists x, y \in \{a,b\}^*, w = xy, |y| = 3, y = y^R\}$
Wyrażenie regularne: $(\mathbf{a}|\mathbf{b})^*(\mathbf{a}(\mathbf{a}|\mathbf{b})\mathbf{a} \mid \mathbf{b}(\mathbf{a}|\mathbf{b})\mathbf{b})$



4.8. Czy następujące języki są regularne?

[TAK] **(A)** $L_1 = \{a^n b^n w \mid w \in \{a, b\}^*, n \geq 0\}$

Jeżeli przyjęte zostanie $n = 0$, to cały język będzie równy $\{a, b\}^*$, a więc jest to język regularny.

[NIE] **(B)** $L_1 = \{a^n b^n w \mid w \in \{a, b\}^*, n > 0\}$

Nie jest to język regularny, a bezkontekstowy. Świadczy o tym zbiór prefiksów $\{a^n b^n \mid n > 0\}$, który nie jest możliwy do wygenerowania za pomocą gramatyki regularnej. Aby to udowodnić, zastosować można lemat o pompowaniu języków regularnych dla słowa $w = a^k b^k$, gdzie k jest stałą z lematu (było na wykładzie).

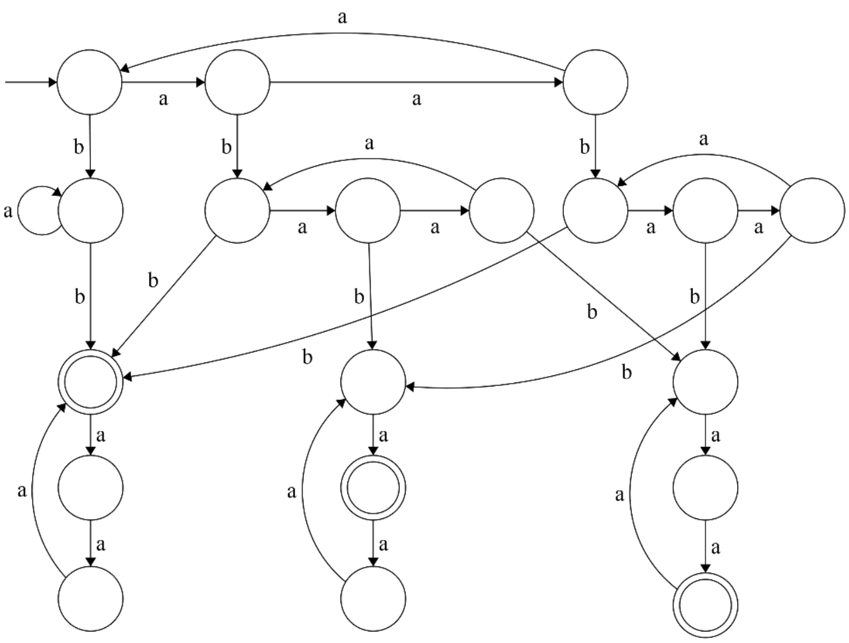
Rozważamy słowo $w = a^k b^k = xyz$, którego długość wynosi $2k > k$. Wówczas słowo y może zawierać od jednego do k symboli a. Inne przypadki nie mogą być brane pod uwagę ze względu na warunek z lematu: $|xy| \leq k$.

W pierwszym przypadku rozważyć można łańcuch xy^2z. Zawiera on od $k+1$ do $2k$ symboli a. W takiej sytuacji łańcuch ten nie należy do języka, ponieważ liczba symboli b pozostaje niezmieniona. A więc łańcuch xy^2z nie należy do języka w żadnym możliwym przypadku.

[TAK] **(C)** $L_3 = \{w \in \Sigma^* \mid \exists x \in \Sigma^*, xx$ jest podłańcuchem $w\}$

Jeżeli przyjęte zostanie $x = \varepsilon$ to wtedy język będzie miał następującą formę: $L_3 = \{w \in \Sigma^*\}$. Jest to więc język regularny.

[TAK] **(D)** $L_4 = \{a^m b a^n b a^p \mid p \equiv mn$ (mod $3), m,n,p \geq 0\}$

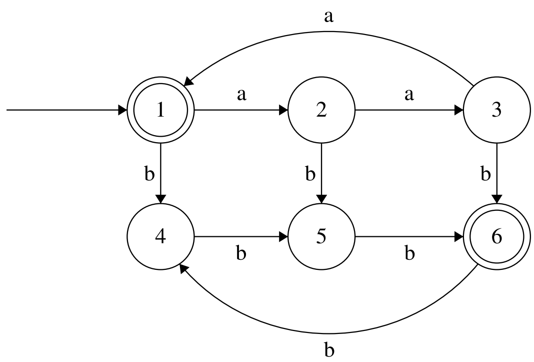


Ze względu na zdolność automatów skończonych do liczenia modulo jesteśmy w stanie stworzyć DFA, który zaakceptuje słowa tego języka. Automat ten znajduje się powyżej.

4.9. Czy następujące języki są regularne?

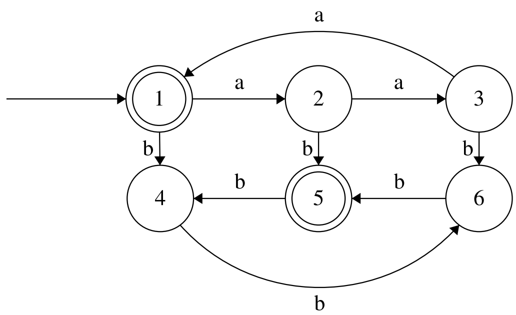
[TAK] **(A)** $L_1 = \{a^i b^j \mid i+j$ jest podzielne przez $3\}$

Deterministyczny automat skończony:



[TAK] **(B)** $L_2 = \{a^i b^j \mid i,j \geq 0,$ i oraz j mają taką samą resztę z dzielenia przez $3\}$

Deterministyczny automat skończony:



[TAK] **(C)** $L_3 = \{xcy \mid x, y \in \{a,b\}^*, |x|_a + |x|_b$ jest podzielne przez $3\}$, gdzie $|x|_a$ oznacza liczbę wystąpień symbolu a w słowie x.

Deterministyczny automat skończony: