Zadanie 1. Rozważmy język złożony ze słów zerojedynkowych o długości co najmniej 3, w których druga i trzecia litera od końca są równe. Narysować diagram minimalnego automatu deterministycznego akceptującego ten język. Uzasadnić minimalność tego automatu.

Rozwiązanie. Rozważmy najpierw deterministyczny automat dla tego języka, który nie jest minimalny. Automat ten trzyma w stanie ostatnie trzy litery przeczytanego słowa. Jeśli słowo jest krótsze niż trzy litery, to trzyma w pamięci całe słowo. Tak więc stany tego automatu to wszystkie słowa zerojedynkowe długości co najwyżej trzy. Stan początkowy to ϵ. Jeśli automat jest w stanie q, który jest słowem długości co najwyżej dwa, i czyta literę $a \in \{0,1\}$, to przechodzi do stanu qa. Jeśli jest w stanie xyz, który jest słowem długości trzy, i czyta literę $a \in \{0,1\}$, to przechodzi do stanu yza. Stanami akceptującymi są te, gdzie równe są litery trzecia i druga od końca, a więc stany $\{000, 001, 110, 111\}$.

Zminimalizujmy ten automat. Dla stanu q tego automatu i słowa $w \in \{0,1\}^*$, oznaczmy przez qw stan do którego automat dojedzie z q po przeczytaniu w. Jeśli w ma przynajmniej trzy litery, to qw nie zależy od w i jest po prostu ostatnimi trzema literami w. Dla słów długości co najwyżej dwa, poniższa tabelka pokazuje dla których par q i w, stan qw jest akceptujący. Stany akceptują te same języki wtedy i tylko wtedy gdy mają ten sam wiersz w tabelce.

stan q	$w = \epsilon$	$w = 0$	$w = 1$	$w = 00$	$w = 01$	$w = 10$	$w = 11$
ϵ	nie	nie	nie	nie	nie	nie	nie
0	nie	nie	nie	tak	tak	nie	nie
1	nie	nie	nie	nie	nie	tak	tak
00	nie	tak	nie	tak	tak	nie	nie
01	nie	nie	nie	nie	nie	tak	tak
10	nie	nie	nie	tak	tak	nie	nie
11	nie	nie	tak	nie	nie	tak	tak
000	tak	tak	nie	tak	tak	nie	nie
001	tak	nie	nie	nie	nie	tak	tak
010	nie	nie	nie	tak	tak	nie	nie
011	nie	nie	tak	nie	nie	tak	tak
100	nie	tak	nie	tak	tak	nie	nie
101	nie	nie	nie	nie	nie	tak	tak
110	tak	nie	nie	tak	tak	nie	nie
111	tak	nie	tak	nie	nie	tak	tak

Stany odpowiadające słowom jedno- i dwuliterowym mają wiersze, które pojawiają się przy stanach trzyliterowych, więc mogą być usunięte. Po zminimalizowaniu, automat ma dziewięć stanów: stan ϵ i wszystkie słowa trzyliterowe. Funkcja przejścia jest określona tak:

$$\delta(\epsilon, 0) = 010 \qquad \delta(\epsilon, 1) = 101 \qquad \delta(xyz, a) = yza.$$

Stan początkowy to ϵ, a akceptujące są $\{000, 001, 110, 111\}$. Niezmiennik automatu jest następujący:

- po przeczytaniu słowa pustego stan to ϵ

- po przeczytaniu słowa długości jeden lub dwa stan to 010 lub 101 w zależności od ostatniej litery przeczytanego słowa
- po przeczytaniu słowa trzyliterowego stan automatu to ostatnie trzy litery

Taki niezmienni, oraz wybór stanów akceptujących, pokazuje, że automat rozpoznaje język z zadania.

Tabelka, którą wcześniej sporządziliśmy, pokazuje, że dla każdych dwóch stanów istnieje słowo, które jeden stan doprowadza do akceptacji, a drugi stan doprowadza do nie-akceptacji, a więc automat ma minimalną liczbę stanów.

Zadanie 2. Niech L będzie zbiorem wszystkich niepustych słów zerojedynkowych, w których liczba wystąpień litery 1 jest równa liczbie wystąpień litery 0. Czy język L^2 jest regularny?
Odpowiedź uzasadnić.

Rozwiązanie.

Korzystając z lematu o pompowaniu, pokażemy, że język L^2 nie jest regularny. Załóżmy, dla sprzeczności, że język ten jest regularny. Niech M stała z lematu o pompowaniu. Rozważmy słowo

$$w = 0^M 1^M 01,$$

które należy do L^2, bo $0^M 1^M$ oraz 01 należą do L. Teza lematu o pompowaniu mówi, że istnieje podział

$$w = w_1 w_2 w_3 \qquad \text{gdzie } |w_1 w_2| \leqslant M \text{ oraz } w_2 \text{ jest niepuste,}$$

taki, że dla każdego $i \in \mathbb{N}$, słowo $w_1 w_2^i w_3$ też należy do L. Skoro słowo $w_1 w_2$ jest nie dłuższe niż M, to słowo w_2 składa się z samych zer. Jest też niepuste. Tak więc, dla $i = 2$, w słowie $w_1 w_2^2 w_3$, które rzekomo ma należeć do L^2, jest więcej zer niż jedynek. W każdym słowie z L^2 jest tyle samo zer co jedynek, a więc otrzymana sprzeczność z tezą lematu o pompowaniu dowodzi, że język L^2 nie spełnia założenia lematu, a więc nie jest regularny.

Zadanie 3. Czy następujący język jest bezkontekstowy, jeśli tak to napisać gramatykę bezkontekstową generującą ten język i objaśnić jej działanie.

$$L' \;=\; \{\; 0^i\, 1^j \;:\; i \neq j \;\&\; 2i \neq j \;\}$$

Rozwiązanie. Język można podzielić na sumę trzech języków $L_1 \cup L_2 \cup L_3$ opisanych poniżej. Języki bezkontekstowe są zamknięte na skończone sumy, a więc wystarczy pokazać, że języki L_1, L_2 i L_3 są bezkontekstowe.

L_1 Słowa postaci $0^i 1^j$, gdzie $j < i$. Język ten opisuje się następującą gramatyką.

$$S \to 0S1 \quad | \quad S \to Y \quad | \quad Y \to 0Y \quad | \quad Y \to 0,$$

gdzie startowym symbolem jest S. Symbol Y opisuje niepuste słowa składające się z samych zer. Symbol S opisuje słowa postaci $0^i w 1^i$, gdzie słowo w jest niepuste i składa się z samych zer, a więc język L_1.

L_2 Słowa postaci $0^i 1^j$, gdzie $i < j < 2i$. Język ten opisuje się następującą gramatyką.

$$S \to 0S1 \quad | \quad S \to 0S11 \quad | \quad S \to 011.$$

W każdej produkcji jest generowana jedna lub dwie jedynki na każde zero, a więc jedynek będzie przynajmniej tyle co zer, ale co najwyżej dwa razy tyle. Ostatnia produkcja, która musi być użyta w każdym wyprowadzeniu, gwarantuje, że jedynek będzie więcej niż zer.

L_3 Słowa postaci $0^i 1^j$, gdzie $j > 2i$. Język ten opisuje się następującą gramatyką.

$$S \to 0S11 \quad | \quad S \to Y \quad | \quad Y \to 1Y \quad | \quad Y \to 1,$$

gdzie startowym symbolem jest S. Symbol Y opisuje niepuste słowa składające się z samych jedynek. Symbol S opisuje słowa postaci $0^i w 1^{2i}$, gdzie słowo w jest niepuste i składa się z samych jedynek, a więc język L_3.

Zadanie 4. Czy dla każdego alfabetu skończonego, bezkontekstowy jest język słów mających nietrywialny okres? Odpowiedź uzasadnij. (Słowo ma nietrywialny okres gdy ma właściwy niepusty prefiks będący sufiksem tego słowa, np. *rytter* ma nietrywialny okres, a *wojciech* nie ma).

Rozwiązanie. Język ten nie jest bezkontekstowy już dla alfabetu $\{0, 1\}$. Niech K oznacza przecięcie języka z zadania z regularnym językiem $0^+1^+0^+1^+$. By pokazać, że nie jest bezkontekstowy język z zadania, wystaczy pokazać, że K nie jest bezkontekstowy, bo języki bezkontekstowe są zamknięte na przecięcia z językami regularnymi.

Pokażemy, że język K nie spełnia tezy lematu o pompowaniu dla języków bezkontekstowych. Niech więc M stała z lematu. Rozważmy słowo

$$w = 0^M 1^M 0^M 1^M.$$

Słowo to należy do K, bo jego pierwsza połowa jest równa drugiej połowie. Na mocy lematu o pompowaniu, istnieje podział

$$w = w_1 w_2 w_3 w_4 w_5 \qquad \text{gdzie } |w_2 w_3 w_4| \leqslant M \text{ oraz } w_2 w_4 \text{ jest niepuste},$$

taki, że dla każdego $i \in \mathbb{N}$, słowo $w_1 w_2^i w_3 w_4^i w_5$ też należy do K. Jeśli słowo w_2 zawiera naraz zera i jedynki, to przy dwukrotnym pompowaniu $i = 2$ dostajemy słowo spoza $0^+1^+0^+1^+$, a więc spoza K. A więc słowo w_2 musi być w całości zawarte w jednym z czeterech bloków (pierwszy blok to pierwszy blok zer, drugi blok to pierwszy blok jedynek itd.). Podobnie dla w_4. Mało tego, cały pompowany kawałek $w_2 w_3 w_4$ musi mieścić się w całości w dwóch sąsiednich blokach, na mocy założenia $|w_2 w_3 w_4| \leqslant M$. Rozważmy pięć przypadków, w zależności od tego, w którym bloku jest w_2.

- w_2 jest w pierwszym bloku. Po dwukrotnym pompowaniu $i = 2$, słowo jest postaci

  $$0^{M+x} 1^{M+y} 0^M 1^M \qquad \text{gdzie } x \geqslant 1 \text{ i } y \geqslant 0.$$

  Żaden właściwy prefiks tego słowa nie jest równy żadnemu sufiksowi. Jeśli prefiks składa się z samych zer, to oczywiście nie jest sufiksem. W przeciwnym przypadku, prefiks zaczyna się od 0^{M+x}, a to słowo można przypasować tylko do samego początku.

- w_2 jest w drugim bloku. Po zerokrotnym pompowaniu $i = 0$, słowo jest postaci

  $$0^M 1^{M-x} 0^{M-y} 1^M \qquad \text{gdzie } x \geqslant 1 \text{ i } y \geqslant 0.$$

  Żaden właściwy prefiks tego słowa nie jest równy żadnemu sufiksowi. Jeśli prefiks składa się z samych zer, to oczywiście nie jest sufiksem. W przeciwnym przypadku, sufiks kończy się na 1^M, a to słowo można przypasować tylko do samego końca.

- w_2 jest w trzecim bloku. Po zerokrotnym pompowaniu $i = 0$, słowo jest postaci

  $$0^M 1^M 0^{M-x} 1^{M-y} \qquad \text{gdzie } x \geqslant 1 \text{ i } y \geqslant 0.$$

  Żaden właściwy prefiks tego słowa nie jest równy żadnemu sufiksowi. Jeśli prefiks składa się z samych zer, to oczywiście nie jest sufiksem. W przeciwnym przypadku, prefiks zaczyna się od 0^M, a to słowo można przypasować tylko do samego początku.

- w_2 jest w czwartym bloku. Po dwukrotnym pompowaniu $i = 2$, słowo jest postaci

  $$0^M 1^M 0^M 1^{M+x} \qquad \text{gdzie } x \geqslant 1$$

  Żaden właściwy prefiks tego słowa nie jest równy żadnemu sufiksowi. Jeśli sufiks składa się z samych jedynek, to oczywiście nie jest sufiksem. W przeciwnym przypadku, sufiks kończy się na 1^{M+x}, a to słowo można przypasować tylko do samego końca.

- w_2 jest puste. Wówczas w_4 jest niepuste. Robimy wtedy takie samo rozumowanie jak powyżej, tylko z przypadkami dotyczącymi bloku, który zawiera w_4.