

Matematyczny świat języków. Wprowadzenie i języki regularne materiały do ćwiczeń

Projekt „Matematyka dla ciekawych świata”
spisał: Michał Korch

18 maja 2017

1 Na rozgrzewkę intelektualną

Poniższe dwa zadania pochodzą z Olimpiady Lingwistyki Matematycznej.

1. Poniżej zapisano kilka zależności o dodatnich liczbach całkowitych w języku ndom, który jest językiem z nowogwinejskiej rodziny językowej używanym przez ok. 1000 mieszkańców indonezyjskiej wyspy Yos Sudarso. W zdaniach nie występują liczby większe niż 20.

- $\text{mer abo meregh} + \text{thef} > \text{mer an thef}$
- $\text{mer abo thef} + \text{mer abo thonith} = \text{tondor}$
- $\text{mer an thef abo ithin} + \text{meregh} = \text{tondor abo thef}$
- $\text{thonith} \cdot \text{thonith} = \text{mer an thef abo thonith}$
- $\text{mer} + \text{mer abo sas} = \text{mer an thef abo sas}$

- (a) Na tej podstawie podajcie co oznaczają następujące liczebniki:

- i. thonith
- ii. meregh
- iii. mer an thef abo sas
- iv. tondor abo thef
- v. mer abo ithin

- (b) Zapiszcie w języku ndom liczby 19 oraz 22.

Wskazówka: Zaczynaj od analizy ostatniej zależności. Co znaczą an, thef i abo?

Wskazówka: Następnie zanalizuj przedostatnią zależność. Można z niej wymyślić znaczenia słów thonith i mer.

Wskazówka: Teraz z drugiej zależności z łatwością odczytasz znaczenie liczebnik tondor!

Wskazówka: Pora zajrzeć do trzeciej zależności. Ile wynosi suma ithin i meregh? Jakimi liczebnikami mogą więc być? Zauważ, że meregh ma podobny człon do mer.

Wskazówka: Pozostaje do rozszefrowania sas. Jakiej cyfry Ci brakuje?

2. Bislama to powstały na bazie angielszczyzny kreolski język używany na Vanuatu, gdzie jest obok angielskiego i francuskiego językiem urzędowym. Poniżej znajduje się zapisana w bislamie zagadka logiczna dotycząca Samuela, Rika, Pita i Zazi – mieszkańców czterech stojących w linii prostej domów ponumerowanych kolejno liczbami od 1 do 4. Każdy dom ma inny kolor: niebieski, zielony, czarny lub czerwony. Ich mieszkańcy wykonują cztery różne zawody: policjanta, ministra, żołnierza i lekarza. Każde z nich ma jedno zwierzę: kota, psa, szczura lub węża, i pije tylko jeden rodzaj napoju: herbatę, kawę, sok lub wodę.

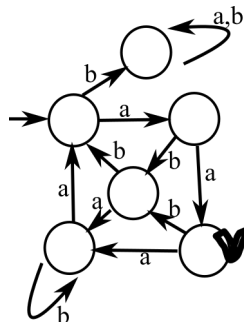
- (a) Zazi hem i stap slip long nambatu haos.
- (b) Nambatri haos we i blufala mo nambafo haos we i redfala.
- (c) Rik hem i stap klosap long redfala haos.
- (d) Man blong rat i stap klosap long haos blong man blong snek.
- (e) Samuel hem i stap slip long grinfala haos mo hem i gat wan puskat.
- (f) Wan man i stap dring wota mo hem i stap klosap long nambatri haos.
- (g) Polisman hem i stap klosap long haos blong man blong snek.
- (h) Wan man hem i stap slip long grinfala haos mo hem i stap dring ti.
- (i) Pit blong snek i stap dring wota.
- (j) Wan man hem i stap dring kafe mo hem i stap klosap long docta.
- (k) Zazi hem i soldia.
- (l) Rik hem i stap dring jus.

- (a) Odpowiedz na pytanie z zagadki: Wanem nem blong man we i ministra? (Jak ma na imię osoba, która jest ministrem?)
- (b) Przetłumacz na polski zdania (d) i (f).
- (c) W którym domu mieszka wąż? W którym domu pije się kawę? W którym domu mieszka kot?

Wskazówka: Np.: drugie zdanie brzmi: Trzeci dom jest niebieski, a czwarty jest czerwony.

2 Automaty skończone

1. Niech będzie dany następujący automat nad językiem $\{a, b\}$.



(a) Dla jakich liczb k , słowo $\underbrace{a \dots a}_k$ jest akceptowane przez ten automat?

(b) Które z poniższych słów są akceptowane przez ten automat?

$abbaa$

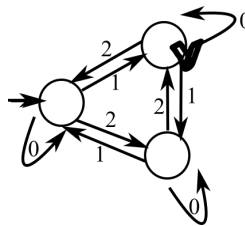
$baab$

$aaaaababb$

$aaabaaa$

$aabbaa$

2. Niech będzie dany następujący automat nad językiem $\{0, 1, 2\}$.



(a) Które z następujących słów są akceptowane przez powyższy automat?

01012

111

1222

120210

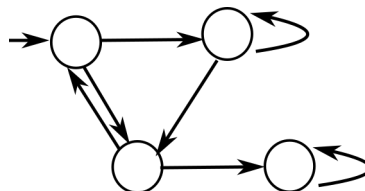
112220

22222

(b) Jaki język słów jest rozpoznawany przez ten automat?

Wskazówka: Zastanów się nad sumą cyfr (liter) w akceptowanych słowach.

3. W poniższym automacie nad alfabetem $\{a, b\}$ podpisz literami poszczególne przejścia oraz oznacz stany końcowe, aby automat akceptował dokładnie wszystkie słowa, poza tymi, w których występują dwie litery b pod rząd.



4. Znajdź automat nad alfabetem $\{a, b\}$, który rozpoznaje język złożony ze słów, w których a występuje parzyście wiele razy, a pierwsza i ostatnia litera to b .

3 Języki, które nie są regularne

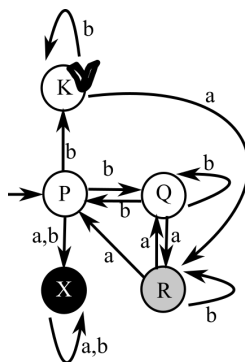
Udowodnij, że następujące języki nie są regularne:

- język wszystkich słów nad alfabetem $\{a, b\}$, które są postaci $\underbrace{a \dots a}_k b \underbrace{a \dots a}_k$.
- język wszystkich słów nad alfabetem $\{a, b, c\}$ będących palindromami (czyli identycznych niezależnie od tego, czy czyta się je od przodu, czy od tyłu).

Wskazówka: Aby udowodnić, że język nie jest regularny, przeprowadź dowód nie wprost. Załóż, że język jest rozpoznawalny przez pewien skończony automat, który ma N stanów i weź do rozpatrzenia słowo zdecydowanie dłuższe niż liczba stanów.

4 Automaty niedeterministyczne

- Niech będzie dany następujący automat niedeterministyczny nad językiem $\{a, b\}$.



- (a) Które z poniższych słów są akceptowane przez ten automat?

abbaab

baaaaab

bbbaaaaabaabb

aba

baabbaabbb

- (b) Opisz język rozpoznawany przez ten automat.
Wskazówka: Zauważ, że czarny stan to stan śmieciowy (nie da się z niego dojść do stanu akceptującego). Kolory biały i szary odpowiadają parzystości liczby wystąpień litery a .
- (c) Zbuduj automat deterministyczny na podstawie tego automatu stosując poznaną na wykładzie konstrukcję.
Wskazówka: Stany tworzonego automatu deterministycznego to zbiory stanów automatu niedeterministycznego, do których można dojść w wyniku danego ciągu liter.

- Udowodnij, że jeśli L, K są językami regularnymi, to również $L \cup K$ jest językiem regularnym.
Wskazówka: Zatem istnieją automaty A i B , które rozpoznają odpowiednio języki L i K . Jak skonstruować z nich automat (niedeterministyczny), który rozpoznaje język $L \cup K$?

5 Wyrażenia regularne

1. Wypisz wszystkie słowa z następujących języków nad alfabetem $\{a, b, c\}$:

$$(a + a.b).(c.(a + c) + \varepsilon),$$

$$a.(a + \varepsilon).(a + b + c).a,$$

$$a.(a + \varepsilon).(a.a + \varepsilon).(a + \varepsilon),$$

2. Opisz języki nad alfabetem a, b zdefiniowane następującymi wyrażeniami regularnymi:

$$a^*.b.a^*,$$

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

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

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

3. Znajdź wyrażenie regularne opisujące następujące języki:

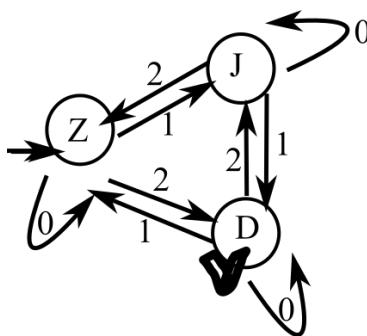
(a) język nad alfabetem $\{a, b\}$ wszystkich słów, w których występuje nieparzysta liczba liter a .

(b) język nad alfabetem $\{a, b, c\}$ wszystkich słów, w których pomiędzy dwoma literami a występuje co najmniej jedna litera b lub c .

(c) język nad alfabetem $\{a, b\}$, wszystkich słów, w którym żadne dwie kolejne litery nie są takie same.

4. Narysuj automat (może być niedeterministyczny), który rozpoznaje język $(a^*.a.(b+c))^* + b^*$.

5. Oblicz wyrażenie regularne odpowiadające poniższemu automatowi (a zatem opisujące język wszystkich słów nad alfabetem $\{0, 1, 2\}$, których suma cyfr dzieli się przez 3 z resztą 2).



Wskazówka:

$$Z = \varepsilon + Z.0 + J.2 + D.1$$

$$J = J.2 + D.2 + Z.1$$

$$D = D.0 + J.1 + Z.2$$