

Parę tuzinów łatwych zadań z języków formalnych i złożoności obliczeniowej

Jerzy Marcinkowski

przygotowane w lutym 2011
jako aneks do podobnego zbioru zadań z roku 2009

Niech $p \geq 5$ będzie pewną liczbą pierwszą, a L_p językiem tych słów nad $\{0,1\}$ które czytane jako liczba w zapisie binarnym dają, jako resztę z dzielenia przez p , jedną z liczb $\{1, 2, \dots, (p-1)/2\}$, przy czym liczby czytamy „od prawej”, czyli od najmniej znaczącego bitu (to znaczy pierwszy znak słowa jest ostatnią cyfrą liczby).

Zadanie A1. Czy istnieje niedeterministyczny automat skończony o mniej niż $p+3$ stanach rozpoznający język L_p ?

Zadanie A2. Czy istnieje deterministyczny automat skończony o mniej niż $2p$ stanach rozpoznający język L_p ?

Zadanie A3. Język $L \in \{0,1\}^*$ jest regularny. Czy wynika z tego, że język

$$\sqrt{L} = \{w \in \{0,1\}^* : \exists x, y \in \{0,1\}^* \quad wx = y \wedge |y| = |w|^2\}$$

jest regularny?

Zadanie A4. Czy zbiór takich słów nad alfabetem $\{0,1\}$, które mają parzystą długość, i w których pierwszej połowie jest przynajmniej tyle samo jedynek, co w drugiej połowie, jest bezkontekstowy?

Zadanie A5. Na wykładzie udowodniliśmy, że rozszerzenie definicji automatu skończonego o możliwość poruszania się po słowie wejściowym w obie strony nie zmieni klasy rozpoznawanych języków. Czy podobnie jest w przypadku automatów ze stosem? Mówiąc dokładniej, rozważamy automaty, których relacja przejścia zawiera się w

$$(Q \times T \times U) \times (Q \times U^* \times \{L, R\})$$

gdzie Q jest skończonym zbiorem stanów („w jakim stanie jestem”), T jest zbiorem symboli taśmowych („co widzę na taśmie”), U zbiorem symboli stosowych (z analogicznymi jak dla zwykłych automatów ze stosem założeniami dotyczącymi symbolu dna stosu), zaś L i R należy rozumieć jako instrukcje „idź w lewo” i „idź w prawo”. Automaty takie są uruchamiane dla słów, których koniec i początek zaznaczone są dodatkowym symbolem taśmowym, nie występującym wewnątrz słowa. Czy każdy język jaki można rozpoznać przy pomocy takiego automatu jest bezkontekstowy? *Wskazówka: wystarczy rozważać takie deterministyczne automaty akceptujące po osiągnięciu jakiegoś końcowego stanu akceptującego.*

Niech $\mathcal{A}$ będzie skończonym alfabetem i niech $L \subseteq \mathcal{A}^*$. Przez $\text{Lustro}(L)$ będziemy w kolejnych zadaniach rozumieli język $\{wv^R \in \mathcal{A}^* : wv \in L\}$

Zadanie A6. Pokaż, że jeśli L jest regularny, to $\text{Lustro}(L)$ również jest regularny.

Zadanie A7. Pokaż, że deterministyczny automat skończony rozpoznający język $\text{Lustro}(L)$ może potrzebować liczby stanów wykładniczo większej niż deterministyczny automat skończony rozpoznający język L .

Zadanie A8. Czy teza Zadania 6. pozostanie prawdziwa, jeśli oba wystąpienia słowa „regularny” zmienimy w nim na „bezkontekstowy”?

W kolejnych dwóch zadaniach niech M_n będzie językiem tych słów nad alfabetem $\{1, 2, \dots, n\}$ (gdzie n jest pewną liczbą parzystą) w których występują wszystkie litery alfabetu oprócz być może jednej. Przez $\overline{M_n}$ rozumiemy dopełnienie języka M_n do zbioru $\{1, 2, \dots, n\}^*$.

Zadanie A9. a. Jaką minimalną liczbę stanów musi mieć deterministyczny automat skończony rozpoznający $\overline{M_n}$?

b. Jaką minimalną liczbę stanów musi mieć niedeterministyczny automat skończony rozpoznający $\overline{M_n}$?

Zadanie A10. Udowodnij, że każdy niedeterministyczny automat skończony rozpoznający język M_n musi mieć więcej niż $2^{\frac{n}{2}-1}$ stanów.

Wskazówka. Dla liczby naturalnej k , takiej, że $1 \leq k \leq n/2$ nazwijmy parę liczb $\{2k-1, 2k\}$ rodziną. Powiemy że słowo $x \in \{1, 2, \dots, n\}^*$ nie rozdziela rodzin, jeśli zawsze wtedy, gdy jedna z liter z jakiejś rodziny występuje w słowie x , w słowie tym występuje również druga z tych liczb. Powiemy że słowo x jest rosnące, gdy każda jego kolejna litera jest liczbą większą niż poprzednia. Ile jest słów nie należących do M_n , które są rosnące i nie rozdzielają rodzin?

Zadanie A11. Niech $L_k = \{0^n : k \text{ nie dzieli } n\}$. Dla języka regularnego L niech $d(L)$ oznacza minimalną liczbę stanów automatu deterministycznego rozpoznającego L , zaś $m(L)$ niech oznacza minimalną liczbę stanów automatu niedeterministycznego rozpoznającego L . Podaj nieskończenie wiele liczb naturalnych k dla których zachodzi $d(L_k) = n(L_k)$, i nieskończenie wiele k naturalnych dla których ta równość nie zachodzi.

Zadanie A12. Pokaż, że $L \subseteq \{0\}^*$ jest bezkontekstowy wtedy i tylko wtedy gdy jest regularny.

Zadanie A13. Przez funkcję liniową będziemy w tym zadaniu rozumieć funkcję postaci $h(x) = ax + b$, gdzie a, b są liczbami naturalnymi.

Instancją problemu skaczącej pchły będzie w tym zadaniu skończony ciąg funkcji liniowych $\langle f_1, g_1, \dots, f_k, g_k \rangle$. Powiemy, że instancja $\mathbf{c} = \langle f_1, g_1, \dots, f_k, g_k \rangle$ problemu skaczącej pchły ma rozwiązanie jeśli istnieje niepusty ciąg $s_1, s_2, \dots, s_l$ liczb ze zbioru $\{1, 2, \dots, k\}$ taki, że punkt $\langle f_{s_l} f_{s_{l-1}} \dots f_{s_1}(0), g_{s_l} g_{s_{l-1}} \dots g_{s_1}(0) \rangle$ leży na prostej $y = x$ (zatem wyobrażamy sobie, że pchła zaczyna skakać w punkcie $\langle 0, 0 \rangle$, w każdym kolejnym skoku umie przemieścić się z punktu $\langle x, y \rangle$ do dowolnego spośród $\langle f_i(x), g_i(y) \rangle$; pytamy o to, czy potrafi kiedykolwiek znów znaleźć się w punkcie o obu współrzędnych równych).

Pokaż, że problem skaczącej pchły, to znaczy problem czy dla danej instancji istnieje rozwiązanie, jest nierozstrzygalny.

Zadanie A14. Niech A, B będą podzbiorami zbioru liczb naturalnych. Załóżmy, że f jest redukcją świadczącą o tym, że $A \leq_{rek} B$. Załóżmy, że f jest „na” (tzn. jej obrazem jest cały zbiór liczb naturalnych). Pokaż, że w takim razie zachodzi również $B \leq_{rek} A$.

Zadanie A15. Niech $\phi(x, y)$ będzie pewną formułą arytmetyki liczb naturalnych z dodawaniem i mnożeniem, z dwiema zmiennymi wolnymi.

Napisz zdanie ψ arytmetyki liczb naturalnych z dodawaniem i mnożeniem, które będzie prawdziwe wtedy i tylko wtedy, gdy istnieją liczba $l \geq 1$ i skończony ciąg liczb naturalnych $a_1, a_2, \dots, a_l$, taki, że $a_1 = 1$, $a_l = 2$, oraz taki, że dla każdego $1 \leq i \leq l - 1$ zachodzi $\phi(a_i, a_i + 1)$.

Wskazówka: Możesz na przykład użyć chińskiego twierdzenia o resztach (choć są również inne rozwiązania). Postuż się makrami, podobnymi do tych, których używaliśmy na wykładzie — na przykład Pierwsza(z), Kolejne-Pierwsze(z, t).

Zadanie A16. Podzbiór zbioru liczb naturalnych nazywamy co-r.e. jeśli jego dopełnienie jest rekurencyjnie przeliczalne. Odcinkami początkowymi zbioru liczb naturalnych nazywamy zbiory postaci $\{1, 2, \dots, n\}$ dla $n \in \mathcal{N}$. Oznaczmy przez B zbiór numerów tych programów, których dziedziny są odcinkami początkowymi zbioru liczb naturalnych.

a. Czy zbiór B jest co-r.e.?

b. Udowodnij, że istnieje zbiór trójek liczb naturalnych A , który jest co-r.e. i którego rzut na pierwszą oś jest zbiorem B . *Uwaga. Wymaga to oczywiście milczącego rozszerzenia definicji zbiorów co-r.e. na zbiory trójek liczb.*

W kolejnych dwóch zadaniach, po napisanym na skończonej taśmie słowie poruszają się żuczki. Dwa albo trzy. Każdy z żuczków jest rodzajem dwukierunkowego deterministycznego automatu skończonego, to znaczy ma skończony zbiór stanów i funkcję przejścia (inną dla każdego żuczka), która w zależności od tego w jakim jest stanie, jaki symbol widzi w aktualnej komórce taśmy, i które z pozostałych żuczków znajdują się wraz z nim w aktualnej komórce taśmy każe mu odpowiednio zmienić stan i poruszyć się w lewo lub w prawo (dla porządku zakładamy, że końce słowa oznaczone są unikalnymi symbolami, dzięki czemu żuczek nigdy nie opuści taśmy i że na początku wszystkie żuczki stoją na początku słowa, w pewnym ustalonym stanie początkowym). Kolejny ruch żuczka, czyli wykonanie funkcji przejścia, następuje zawsze wtedy, gdy usłyszysz on tyknięcie zegara.

Przez *problem niepustości* rozumiemy pytanie, czy dla danych funkcji przejścia żuczków istnieje słowo, które żuczki zaakceptują, to znaczy takie, na którym któryś z nich osiągnie, po skończonej liczbie kroków, ustalony stan akceptujący.

Zadanie A17. Dwa synchroniczne żuczki. Pokaż, że problem niepustości jest nierozstrzygalny jeśli rozważamy pary (funkcji dla) żuczków i zakładamy, że są one synchroniczne, to znaczy oba słyszą tykanie tego samego zegara.

Zadanie A18. Trzy asynchroniczne żuczki. Pokaż, że z tezy Zadania 17 wynika, że nierozstrzygalny jest również problem niepustości dla trójek (funkcji dla) żuczków jeśli zakładamy, że są one asynchroniczne, to znaczy w każdej komórce taśmy słyszą osobny zegarek, który tyka jak chce (np. czasem wolniej czasem szybciej). *Uwaga. Różne zachowania zegarków mogą tu skutkować różnymi obliczeniami, czyli różnymi zachowaniami żuczków. Myślimy o tym jak o obliczeniu niedeterministycznym: słowo zostaje zaakceptowane, gdy istnieje zachowanie zegarków, które prowadzi do obliczenia akceptującego.*

Komentarz. Nie znam rozwiązania zadania o dwóch asynchronicznych żuczkach.

W kolejnych zadaniach rozważamy płaszczyznę, po której biega k psów. Startują one jednocześnie z początku układu współrzędnych, a następnie każdy z nich, w każdej jednostce czasu, przesuwa się o jednostkę odległości na północ, południe, wschód lub zachód (zatem po każdym takim kroku psy znajdują się w punktach kratowych płaszczyzny). O kierunku kolejnego ruchu psa decyduje jego funkcja przejścia (psy są różne, i mogą mieć różne funkcje przejścia). Argumentami funkcji przejścia są: aktualny stan psa (każdy pies ma skończoną

liczbę stanów) oraz zapach punktu kratowego aktualnie zajmowanego przez psa. Przez zapach pola rozumiemy tu informację o tym, które psy odwiedziły już to pole.

Wśród stanów każdego psa wyróżniamy jeden stan szczekający.

Mówiąc bardziej formalnie, funkcja przejścia i -tego psa δ_i ma typ $\delta_i : Q_i \times Z \rightarrow Q_i \times \{N, S, E, W\}$, gdzie Q_i to skończony zbiór stanów i -tego psa, zaś $Z = \mathcal{P}(\{1, \dots, k\})$ jest zbiorem zapachów (czyli rodziną wszystkich podzbiorów zbioru wszystkich biegających psów). Funkcja ruchu i -tego psa $\hat{\delta}_i$ definiowana jest przy pomocy δ_i w naturalny dla teorii automatów skończonych sposób. Wreszcie zapach $z(p, t)$ punktu kratowego p w chwili t jest równy $z(p, t-1) \cup S(p, t-1)$, gdzie $S(p, t)$ to zbiór numerów psów znajdujących się w punkcie p w chwili t . W chwili zerowej zapachem wszystkich punktów jest zapach pusty.

Problem szczeku dla układu k psów jest następujący: dane funkcje przejścia k psów, oraz informacja o tym które stany są szczekające. Czy któryś z biegających zgodnie z tymi funkcjami psów osiągnie kiedyś stan szczekający?

Zadanie A19. Czy problem szczeku dla układu 3 psów jest rozstrzygalny? *Wskazówka: Czy punkt kratowy w pierwszej ćwiartce układu współrzędnych ma coś wspólnego z konfiguracją automatu z dwoma licznikami?*

Zadanie A20. Czy problem szczeku dla układu z jednym psem jest rozstrzygalny? *Wskazówka: Zadanie z trzema psami jest znacznie łatwiejsze.*

Zadanie A21. Jaka jest złożoność problemu 3-kolorowania grafów, jeśli ograniczymy się do grafów o stopniu wierzchołków równym co najwyżej 4?

Zadanie A22. Rozważmy następujący Problem Trójek Klasowych (PTK).

Rodzice uczniów każdej klasy w szkole wybierają spośród siebie (do różnych niezmiernie ważnych zadań) trójkę – zwaną trójką klasową. Ponieważ można być rodzicem więcej niż jednego dziecka w szkole, można być członkiem więcej niż jednej takiej trójki.

Spośród członków tych trójek dyrektor szkoły chce powołać szkolny komitet rodzicielski, do którego należy **dokładnie jeden** członek każdej trójki klasowej. PTK jest problemem rozstrzygnięcia, czy, dla danej listy trójek klasowych, powołanie takiego komitetu jest możliwe.

Udowodnij, konstruując odpowiednią redukcję, że $3\text{COL} <_P \text{PTK}$.

Zadanie A23. Jaka będzie złożoność problemu z poprzedniego zadania, jeśli każde wystąpienie słowa *trójka* zamienimy w nim słowem *dwójka*?

Zadanie A24. Przez *liczbę chromatyczną grafu nieskierowanego* $G = \langle V, E \rangle$ rozumiemy najmniejszą liczbę naturalną n dla której istnieje funkcja $l : V \rightarrow \{1, 2, \dots, n\}$ taka, że zachodzi formuła $\forall x, y \in V \ E(x, y) \Rightarrow l(x) \neq l(y)$. Liczbę chromatyczną grafu G oznaczamy przez $\chi(G)$

Udowodnij, że jeżeli istnieje wielomianowy algorytm który, dla danego grafu G , zwróci zawsze jedną z liczb $\{\chi(G), \chi(G) + 1\}$, to $\text{PTIME} = \text{NP}$

Przez *pokrycie cyklowe* nieskierowanego grafu $G = \langle V, E \rangle$ rozumiemy zbiór rozłącznych (wierzchołkowo) cykli o wierzchołkach w V i krawędziach w E , taki, że każdy wierzchołek ze zbioru V należy do któregoś z tych cykli. *Moc pokrycia wierzchołkowego* to liczba cykli z których to pokrycie się składa. Dla grafu G przez $\sigma(G)$ oznaczmy minimalną moc pokrycia cyklowego G (ponieważ wierzchołek jest sam w sobie cyklem, więc liczba $\sigma(G)$ jest zawsze określona i nie większa od liczby wierzchołków G).

Zadanie A25. Niech $c > 3$. Jaka jest złożoność problemu stwierdzenia, dla danego grafu G , czy $\sigma(G) \leq n/c$, gdzie n to liczba wierzchołków G ?

Zadanie A26. Załóżmy, że istnieje wielomianowy algorytm, który dla każdego grafu G takiego że $\sigma(G) = 1$ zwraca pokrycie cyklowe G składające się z nie więcej niż dwóch cykli. Pokaż, że w takim razie $P=NP$.

Zadanie A27. a. (od -1 do 7 punktów) Pokaż, że jeśli P jest różne od NP to funkcja $\sigma(G)$ nie może być, przez żaden wielomianowy algorytm, aproksymowana z dokładnością do stałej multiplikatywnej. Mówiąc dokładniej, pokaż że nie istnieje wtedy wielomianowy algorytm M i liczba $c > 0$ taka, że dla każdego grafu G algorytm M uruchomiony dla G zwróci jego pokrycie cyklowe o mocy nie większej niż $c\sigma(G)$.

Zadanie A28. Czy istnieje język $L \in PTIME$, który nie daje się rozpoznawać, na maszynie Turinga, w czasie kwadratowym? (Przez *czas kwadratowy* rozumiemy czas ograniczony przez $c(n^2+1)$ gdzie n jest wielkością wejścia a c pewną stałą niezależną od wielkości wejścia).

Niech $G = \langle V, E \rangle$ będzie grafem nieskierowanym. Przypominam, że kliką w G nazywamy zbiór $U \subset V$ taki, że $\forall x, y \in U \{x, y\} \in E$, zaś zbiorem niezależnym w G nazywamy zbiór $U \subset V$ taki, że $\forall x, y \in U \{x, y\} \notin E$.

W poniższych dwóch zadaniach wolno korzystać ze wszystkiego, czego w czasie wykładu i ćwiczeń dowiedzieliście się o złożoności różnych wersji problemu spełnialności formuł boolewskich. Nie wolno natomiast korzystać z zadań o złożoności problemu klik, ani innych problemów grafowych.

Zadanie A29. a. Jaka jest złożoność problemu rozstrzygnięcia, dla danego grafu $G = \langle V, E \rangle$, czy istnieje kliką $U \subset V$ taka, że $V \setminus U$ jest zbiorem niezależnym?

b. Jak zmieni się złożoność problemu z punktu **a.**, jeżeli dodatkowo zażądamy, aby moc U była dokładnie połową mocy V ?

Zadanie A30. Jaka jest złożoność problemu rozstrzygnięcia, dla danego grafu $G = \langle V, E \rangle$, czy istnieją kliką $U \subset V$ i zbiór niezależny $Y \subset V$, oba o mocy nie mniejszej niż jedna czwarta mocy V .

Zadanie A31. *Problem wesołego dwukolorowania* zdefiniujemy w tym zadaniu następująco. Instancją problemu jest graf nieskierowany. Pytamy, czy da się pokolorować wierzchołki tego grafu dwoma kolorami w taki sposób, aby każdy wierzchołek miał wśród swoich sąsiadów przynajmniej po jednym wierzchołku każdego koloru. Pokaż, że problem wesołego dwukolorowania jest NP-zupełny.

Wskazówka: Wolno skorzystać z NP-zupełności problemu not-all-equal-3-SAT. Instancjami tego problemu są formuły w postaci 3-CNF, a pytamy, czy istnieje takie wartościowanie zmiennych, że w każdej klauzuli jest przynajmniej jeden literał zwartościowany jako prawdziwy i przynajmniej jeden literał zwartościowany jako fałszywy.

Zadanie A32. *Grę w kompromis* definiujemy, na potrzeby tego zadania, następująco. Planując do gry stanowi skierowany graf dwudzielny $\langle V, E \rangle$ (gdzie $V = L \cup P$ jest podziałem V wynikającym z dwudzielności grafu; zakładamy ponadto, że minimalny stopień wyjściowy

wierzchołka jest ≥ 2), z wyróżnionym wierzchołkiem $c_0 \in L$ zwanym początkowym, i ze zbiorem $W \subseteq L$, zwanym zbiorem pozycji wygrywających. W kompromis grają dwaj gracze, $\mathcal{L}$ i $\mathcal{P}$, wykonujący ruchy na przemian.

Protokół ruchu gracza $\mathcal{L}$ jest następujący. Zastaje on planszę z kamieniem umieszczonym w jakimś wierzchołku $s \in L$. Następnie wybiera dwa różne wierzchołki $t_1, t_2 \in P$ takie, że zachodzi $E(s, t_1)$ i $E(s, t_2)$ i mówi: *wybierz sobie bracie, gdzie chcesz bym się ruszył*. Jego przeciwnik wybiera jeden spośród wierzchołków t_1, t_2 , a gracz $\mathcal{L}$ przesuwa tam kamień. Protokół ruchu gracza $\mathcal{P}$ jest analogiczny, z tym że zamienione są role zbiorów P i L .

Na początku gry kamień leży w c_0 , zatem pierwszy ruch wykonuje gracz $\mathcal{L}$. Gra kończy się zwycięstwem gracza $\mathcal{P}$ gdy uda mu się postawić kamień w wierzchołku należącym do W . Gra kończy się zwycięstwem gracza $\mathcal{L}$ jeśli gracz $\mathcal{P}$ nie wygra w ciągu $2|V|$ ruchów.

Jaka jest złożoność problemu rozstrzygnięcia, dla danej planszy do gry w kompromis, który z graczy ma w niej strategię wygrywającą?

Wskazówka: The owls are not what they seem.

Na planecie Nijak używa się logiki o trzech wartościach: T (prawda), F (fałsz) i M (mhm). Spójniki logiczne $\vee, \wedge$ i $\neg$ są dla wartości T i F określone tak jak na Ziemi, $\neg(M) = M$, zaś $\vee$ i $\wedge$ są symetryczne i $M \vee M = M \wedge M = M$, $T \vee M = T \wedge M = T$ i $F \vee M = F \wedge M = F$. Definicje postaci CNF, 2CNF itd. są na Nijak takie same jak na Ziemi. Podobnie, formuła jest na Nijak *spełnialna* jeśli istnieje wartościowanie zmiennych przy którym ma ona wartość logiczną T.

Zadanie A33. Jaka jest na Nijak złożoność problemu spełnialności formuł w postaci 2CNF?

Zadanie A34. Jaka jest na Nijak złożoność problemu spełnialności formuł w postaci 3CNF? W zadaniu tym zakładamy, że literałami są zmienne, negacje zmiennych, oraz stałe T, F i M.

Zadanie A35. Jaka jest na Nijak złożoność problemu spełnialności formuł w postaci 3CNF? W zadaniu tym zakładamy, że literałami są zmienne i negacje zmiennych, ale nie są nimi stałe T, F i M.