Egzamin z równań w słowach, 13 lutego 2004

czas trwania egzaminu: 150 min., wymagania:
na ocenę dostateczną – ok. 30%, na ocenę bardzo dobrą – ok. 75% punktów

Zad. 1. (7 pkt.) Rozwiąż równanie w słowach $(ax)^n = (yb)^m$, gdzie a i b są literami stałymi, x i y – zmiennymi, a n i m ustalonymi dodatnimi liczbami naturalnymi.

Rozwiązanie. Po pierwsze przypomnijmy sobie lemat 35.6. Stwierdza on, że jeżeli słowa X^n i Y^m mają wspólny prefiks długości $\mid XY \mid$, to słowa X i Y komutują i są potęgami pewnego słowa. (Szkic uzasadnienia: jeżeli założymy, że słowo X jest dłuższe niż Y, to XY jest tym wspólnym prefiksem i daje się przedstawić w postaci Y^tP dla pewnego prefiksu P słowa Y. Przyjmijmy, że $Y = PS$. Wtedy $XPS = XY = Y^tP = (PS)^tP$. Stąd wynika, że $PS = SP$ itd.) W dowodzie lematu lub rozwiązując zadanie bezpośrednio można skorzystać też z własności sformułowanej w zadaniu 6 z listy 3.

Przypuśćmy, że X i Y spełniają równanie $(ax)^n = (yb)^m$. Jeżeli $n = 1$ lub $m = 1$, to oczywiście aX i Yb są potęgami pewnego słowa. Jeżeli $n, m \geqslant 2$, to to samo wynika z przytoczonego lematu. Przyjmijmy więc, że $aX = T^i$ oraz $Yb = T^j$ dla pewnego słowa T i liczb $i, j > 0$. Teraz łatwo zauważyć, że słowo T zaczyna się literą a i kończy się literą b. Mamy więc $T = aZb$ dla pewnego słowa Z. Wyliczenie X i Y już nie przedstawia trudności: $X = Zb(aZb)^{i-1}$ oraz $Y = (aZb)^{j-1}aZ$. Każde rozwiązanie analizowanego równania jest więc takiej postaci. Nietrudno zauważyć, że jeżeli X i Y spełniają podane wzory, to aby były rozwiązaniem równania $(ax)^n = (yb)^m$, muszą mieć odpowiednie długości. Jeżeli podstawimy wzory na X i Y do równania, to bez trudu otrzymamy, że $i \cdot n = j \cdot m$.

Ostatecznie, słowa X i Y spełniają równanie $(ax)^n = (yb)^m$ wtedy i tylko wtedy, gdy istnieje słowo Z i liczby $i, j > 0$ takie, że $X = Zb(aZb)^{i-1}$, $Y = (aZb)^{j-1}aZ$ oraz $i \cdot n = j \cdot m$.

Zad. 2. (8 pkt.) Pokaż, że problem niesprzeczność równań w słowach jest rozstrzygalny wtedy i tylko wtedy, gdy istnieje obliczalna funkcja $f : R \to R$ taka, że jeżeli jakiekolwiek równanie w słowach $u = v$ ma rozwiązanie, to ma także rozwiązanie α takie, że $\mid \alpha(u) \mid \leqslant f(\mid uv \mid)$.

Rozwiązanie. Jest oczywiste, że istnieje algorytm, który dla danego równaniach w słowach R i danego podstawienia α odpowiada na pytanie, czy α jest rozwiązaniem R.

Wiadomo też, że są algorytmy, które dla danego niesprzecznego równania w słowach zwracają jedno z jego rozwiązań. Taki jest algorytm Plotkina. (rozwiązanie niedokończone!)

Zad. 3. (6 pkt.) Podaj znane Ci rezultaty o oszacowaniach z góry i z dołu złożoności obliczeniowej następujących problemów:

1. problem niesprzeczności równań w słowach (dane: równanie w słowach; czy dane równanie ma rozwiązanie),

2. problem niesprzeczności kwadratowych równań w słowach (dane: kwadratowe równanie w słowach; czy dane równanie ma rozwiązanie),

3. problem niesprzeczności równań w słowach z więzami regularnymi (dane: równanie i lista języków regularnych; czy dane równanie ma rozwiązanie o składowych należących do odpowiednich języków regularnych),

Rozwiązanie. jest zawarte w notatkach z wykładu.

Zad. 4. (7 pkt.) Pokaż, że dla pewnej stałej c, jeżeli równanie w słowach $u = v$ z jedną zmienną ma rozwiązanie, to ma rozwiązanie α takie, że $|\,\alpha(u)\,| \leqslant c \cdot |\,uv\,|^2$.

Rozwiązanie. Będziemy rozważać równanie $u = v$ z jedną zmienną, które ma postać $Ax \ldots = x \ldots$ Wiadomo, że rozwiązaniami równania $u = v$ są (niektóre) słowa $X = A^t B$ dla $B < A$ (i tylko takie słowa). Zgodnie z omawianym na ćwiczeniach algorytmem, aby ustalić niesprzeczność równania jedną zmienną, wystarcza dla pewnego g przebadać tylko, czy równanie spełnia chociaż jedno takie słowo, dla którego $t \leqslant g$. Jeżeli oszacujemy g, to również oszacujemy długość pewnego (a więc także minimalnego) rozwiązania.

Poprawność wspomianego algorytmu wynika z następującego twierdzenia:

Twierdzenie 0.1 *Przypuśćmy, że mamy słowa $P_0, \ldots, P_n$ i $Q_0, \ldots, Q_m$, oraz liczby $k = \min\{n, m\}$ i $G = \max\{|\,|\,P_0 \ldots P_i\,| - |\,Q_0 \ldots Q_i\,|\,| : i \leqslant k\}$. Jeżeli A jest niepustym słowem i $l \geqslant 1 + \frac{G}{|\,A\,|}$ taką, że*

$$P_0 A^l P_1 A^l \ldots A^l P_n = Q_0 A^l Q_1 A^l \ldots A^l Q_m,$$

to $n = m$ i zachodzą równości

$$P_0 A^t P_1 A^t \ldots A^t P_n = Q_0 A^t Q_1 A^t \ldots A^t Q_m$$

dla wszystkich $t \in N$.

Dowolne równanie z jedną zmienną, a więc także $u = v$, jest postaci

$$AxP_1 x \ldots xP_n = xQ_1 x \ldots xQ_m$$

(będziemy rozważać sytuacje, w której $P_0 = A \neq \varepsilon$ i $Q_0 = \varepsilon$). Jeżeli to równanie ma rozwiązanie α, to $\alpha(x) = X = A^l B$ dla pewnego $B < A$ i $l \in N$. Podstawiając X do równania otrzymujemy

$$AA^l(BP_1)A^l \ldots A^l(BP_n) = A^l(BQ_1)A^l \ldots A^l(BQ_m).$$

Chcemy teraz skorzystać z podanego twierdzenia. Zauważmy, że liczba G w rozważanym przypadku nie zależy od B i może być definiowana tym samym wzorem, co w sformułowaniu twierdzenia. Wobec tego,

$$\begin{aligned}
&G = \max\{|\,|\,P_0 \ldots P_i\,| - |\,Q_0 \ldots Q_i\,|\,| : i \leqslant k\} \leqslant \\
&\leqslant \max\{|\,P_0 \ldots P_i\,|, |\,Q_0 \ldots Q_i\,| : i \leqslant k\} \leqslant \\
&\leqslant |\,P_0 \ldots P_n\,| + |\,Q_0 \ldots Q_m\,| \leqslant |\,uv\,|.
\end{aligned}$$

Z twierdzenia wynika też, że jeżeli rozważane równanie ma rozwiązanie, to ma także rozwiązanie $X = A^l B$ takie, że $l < 1 + \frac{G}{|\,A\,|}$. Oszacujemy długość takiego X. Z podanej nierówności wynika, że $|\,X\,| = |\,A^l B\,| < |\,A\,| + |\,B\,| + G$. Stąd otrzymujemy, że

$$\begin{aligned}
&|\,\alpha(u)\,| = |\,AXP_1 X \ldots XP_n\,| = |\,AP_1 \ldots P_n\,| + n|\,X\,| \leqslant \\
&\leqslant |\,AxP_1 x \ldots xP_n\,| \cdot |\,X\,| \leqslant |\,u\,|(|\,A\,| + |\,B\,| + G) \leqslant 3|\,uv\,|^2.
\end{aligned}$$

Zad. 5. (10 pkt.) Podaj algorytm rozwiązujący problem niesprzeczności kwadratowego układu równań, wykaż jego poprawność i oceń złożoność (rozumianą na przykład jako liczbę operacji elementarnych w zależności liczby znaków potrzebnych do zapisania układu).

Rozwiązanie. Dobrze jest załozyć, że rozważamy rozwiązania o niepustych skladowych. Wystarczy trochę uogólnić algorytm Plotkina. Przypuśćmy, że $U = \{u_1 = v_1, \dots, u_n = v_n\}$ jest kwadratowym o długości $d = \sum_{i=1}^{n} |\, u_i v_i \,|$. Zdefiniujemy układ U_δ. W tym celu wybieramy z układu U jedno z równań. Przyjmijmy, że wybraliśmy równanie $x \dots = y \dots$, w którym obie strony zaczynają się zmiennymi. Następnie wybieramy spośród φ_{xy}, φ_{yx} i ε_{xy} podstawienie δ, tworzymy zbiór $\{\delta(u_1) = \delta(v_1), \dots, \delta(u_n) = \delta(v_n)\}$ i z każdego równania z tego zbioru usuwamy wspólny dla obu stron prefiks. Otrzymany w ten sposób układ jest definiowanym U_δ. (Jeżeli wybraliśmy równanie postaci $x \dots = a \dots$, gdzie a jest stałą, to δ wybieramy tylko spośród φ_{ax} i ε_{ax}, a nie przekształcamy równań postaci $a \dots = b \dots$ dla różnych stałych a, b.)

Układ U_δ ma następujące własności:

1. zmienne i stałe występujące w U_δ występują w U (oczywiste),
2. długość układu U_δ nie przekracza długości układu U (ponieważ U jest kwadratowy, więc wykonując podstawienie δ w najgorszy razie w dwóch miejscach dopiszemy po jednej literze, a następnie usuniemy w wybranym (w konstrukcji U_δ) równaniu przynajmniej dwie litery),
3. układ U_δ jest kwadratowy (uzasadniamy to podobnie jak poprzednią własność),
4. jeżeli α jest rozwiązaniem U_δ, to $\alpha\delta$ jest rozwiązaniem U (oczywiste),
5. jeżeli układ U ma rozwiązanie α długości $d_\alpha = \sum_{i=1}^{n} |\, \alpha(u_i) \,|$, to dla pewnego δ układ U_δ ma rozwiązanie β takie, że $d_\beta < d_\alpha$.

Dokładniejszego uzasadnienia wymaga tylko ostatnia własność. Zrobimy to w jednym przypadku, w pozostałych postępujemy podobnie. Przypuśćmy, że z układu U wybraliśmy równanie postaci $xw_1 = yw_2$ oraz rozwiązanie α spełnia nierówność $\alpha(x) > \alpha(y)$. Wtedy bierzemy $\delta = \varphi_{yx}$ i definiujemy $\beta(z) = \alpha(z)$ dla $z \neq x$ oraz $\beta(x)$ jako słowo spełniające równość $\alpha(x) = \alpha(y)\beta(x)$. Tak zdefiniowany homomorfizm β spełnia równość $\alpha = \beta\delta$. Równość ta implikuje, że β jest rozwiązaniem U_δ a ponadto, że $d_\beta \leqslant d_\alpha$, gdyż analogiczne nierówności zachodzą dla odpowiadających sobie składników tych sum. Aby dowieść ostrą nierówność $d_\beta < d_\alpha$ wystarczy zauważyć, że ostra nierówność zachodzi dla jednego składnika. Jest składnik odpowiadający równaniu $xw_1 = yw_2 \in U$. W układzie U_δ mamy skrócone równanie otrzymame z $\delta(xw_1) = \delta(yw_2)$, czyli z $yx\delta(w_1) = y\delta(w_2)$. To skrócone równanie jest równe $x\delta(w_1) = \delta(w_2)$. Zauważmy, że

$$|\, \beta\delta(w_2) \,| = |\, \alpha(w_2) \,| < |\, \alpha(yw_2) \,|.$$

Dla danego kwadratowego układu równań U tworzymy graf $\mathcal{G}$, którego wierzchołki są etykietowane kwadratowymi układami o długości nie przekraczającej długości U, zbudowanymi z tych samych zmiennych i stałych. Z wierzchołka odpowiadającego U w grafie $\mathcal{G}$ wychodzą krawędzie dokładnie do wierzchołków postaci U_δ. Bez trudu dowodzi się, że układ U ma rozwiązanie wtedy i tylko wtedy, gdy w grafie $\mathcal{G}$ jest ścieżka od wierzchołka odpowiadającego U, do wierzchołka odpowiadającego układowi $\{\varepsilon = \varepsilon\}$. Istnienie takiej ścieżki gwarantuje niesprzeczność układu U na podstawie własności 4. Własność 5 pozwala skonstruwać ścieżkę prowadzącą od wierzchołka wyznaczonego przez układ niesprzeczny do wierzchołka z układem $\{\varepsilon = \varepsilon\}$. (brakuje jeszcze szacowania złozonosci)

Zad. 6. (7 pkt.) Udowodnij, że każde rozwiązanie $\alpha : X^* \to A^*$ równania R w słowach bez stałych, ze zmiennymi ze zbioru X, o różnych stronach jest złożeniem $\theta\beta$ rozwiązania $\beta : X^* \to B^*$ tego równania i homomorfizmu (podstawienia) $\theta : B^* \to A^*$ dla pewnego zbioru B o mniejszej liczbie elementów niż X.

Rozwiązanie. Po pierwsze zauważmy, że jeżeli X jest alfabetem, a X^* jest **zbiorem słów nad alfabetem** X, to X^* (z wiadomym działaniem) jest z definicji półgrupą wolną. Jeżeli bierzemy równanie $u = v$ o różnych stronach, to $u \neq v$ i nie ma szans, aby zachodziła równość $u = v$. Mimo to, równanie zapisujemy zwykle w postaci $u = v$ stwarzając wiele możliwości robienia błędów logicznych. Tradycyjnie równania tak właśnie zapisujemy. Jeżeli komplikacje logiczne stają za bardzo uciążliwe, to równanie zapisuje się w postaci (u, v). Zgodnie z formalną definicją równanie jest po prostu parą słów.

Przypuśćmy, że mamy równanie $u = v$ i rozwiązanie α tego równania. Strony tego równania są różne, a więc $u \neq v$. Za to zachodzi równość $\alpha(u) = \alpha(v)$.

Zbiór $\alpha(X^*)$ (czyli obraz zbioru X^* wyznaczony przez funkcję α) jest zawarty w A^* (w zbiorze słów nad alfabetem A). Zbiór ten jest półgrupą (jest zamknięty ze względu na konkatenację i należy do niego ε), jest też podpółgrupą A^*. Jako półgrupa jest generowany przez $\alpha(X)$ (obraz zbioru X). Wobec tego, $\alpha(X^*) = (\alpha(X))^*$ (ten drugi zbiór to **domknięcie Kleene'go** zbioru $\alpha(X)$). Jeżeli $u = x_1 \dots x_n$ oraz $v = y_1 \dots y_m$ i $x_1, \dots, x_n, y_1, \dots, y_m \in X$, to

$$\alpha(x_1) \dots \alpha(x_n) = \alpha(u) = \alpha(v) = \alpha(y_1) \dots \alpha(y_m).$$

Ponieważ słowa u i v są różne, więc pewien element półgrupy $\alpha(X^*)$ ma dwa różne przedstawienia w postaci konkatenacji elementów $\alpha(X)$. Oznacza to, że $\alpha(X)$ nie jest zbiorem wolnych generatorów (a więc albo półgrupa $\alpha(X^*)$ nie jest wolna, albo jest wolna, ale $\alpha(X)$ nie jest jej najmniejszym zbiorem generatorów). Zgodnie z twierdzeniem o defekcie, półgrupę $\alpha(X^*)$ można zastąpić większą półgrupą wolną $M \subseteq A^*$ z mniejszym zbiorem (wolnych) generatorów Y. Powtórzmy jeszcze to samo bardziej formalnie: istnieje zbiór $Y \subseteq A^*$ taki, że $|Y| < |\alpha(X)|$ oraz domknięcie Kleene'go Y^* (podpółgrupa generowana przez Y) zawiera M (a więc $M \subseteq Y^*$); ponadto dla Y prawdziwe jest twierdzenie o jednoznaczności rozkładu w postaci konkatenacji elementów Y. Teraz możemy przystąpić do rozwiązywania zadania.

Weźmy dowolny alfabet B równoliczny z Y i funkcję $\theta : B \to Y$ ustalającą równoliczność tych zbiorów. Można ją rozszerzyć do homomorfizmu $\theta : B^* \to Y^*$ (B^* to zbiór słów nad alfabetem B, Y^* to domknięcie Kleene'go zbioru Y). Oczywiście, każdy element Y jest wartością θ. Stąd można wywnioskować, że każdy element Y^* też jest wartością θ. Ponadto, z twierdzenia o jednoznaczności rozkładu w postaci konkatenacji elementów Y wynika, że θ jest różnowartościowy. Wszystko razem świadczy to o tym, że θ jest izomorfizmem B^* i Y^*.

Teraz wystarczy zauważyć, że rozwiązanie α spełnia oczywisty wzór

$$\alpha = \theta(\theta^{-1}\alpha).$$

Jeżeli przyjmiemy, że $\beta = \theta^{-1}\alpha$, to dla B, θ i β będą prawdziwe wszystkie warunki wymienione w treści zadania. Uff! przypuszczam, że tego rozwiązania nie da się przeczytać. Mam nadzieję, że przestudiowanie pozwoli niektórym osobom zrozumieć błędy zrobione na egzaminie.

Zad. 7. (6 pkt.) Podaj graf związany z równaniem $xyz = zyx$. Odczytaj z tego grafu rozwiązanie α podanego równania (jedno z wielu) takie, że $\alpha(x) \neq \alpha(z)$ i $|\,\alpha(xyz)\,| \geqslant 10$.

Rozwiązanie. Wystarczy narysować graf, wybrać dostatecznie długą ścieżkę i wyliczyć rozwiązanie.