Michał Górniak

Zadanie 29. Język $L \subseteq \{0,1\}^*$ jest regularny. Czy wynika z tego, że język

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

jest regularny?

Rozwiązanie. Tak, język $\sqrt{L}$ jest regularny. Przedstawię tego dowód. W dowodzie tym pokażę DFA, który rozpoznaje język $\sqrt{L}$. Jako że język L jest regularny, możemy wziąć DFA $\mathcal{A} = \langle Q, \Sigma, \delta, q_0, F\rangle$, który go rozpoznaje. Niech $n = |Q|$ oraz $Q = \{q_0, q_1, \ldots, q_{n-1}\}$.

Automat, który pokażę, będzie tak naprawdę krotką $n+1$ automatów. Zastosuję sztuczkę podobną do tej, która była stosowana w zadaniu 22, czyli wczytanie słowa będzie tak naprawdę wczytaniem słowa do wszystkich $n+1$ automatów, każdy z nich da nam pewną informację, a na ich podstawie będziemy w stanie decydować, czy stan jest akceptujący czy nie.

Pierwszym z tych $n+1$ automatów będzie po prostu $\mathcal{A}$. Dzięki niemu, będziemy wiedzieli w jakim stanie znajdujemy się po wczytaniu wybranego słowa.

Oznaczmy pozostałe automaty przez $\mathcal{B}_0, \mathcal{B}_1, \ldots, \mathcal{B}_{n-1}$. Każdy automat jest powiązany z odpowiednim stanem automatu $\mathcal{A}$, tzn. automat $\mathcal{B}_i$ ze stanem q_i. Dokładniej, chciałbym aby automat B_i po wczytaniu dowolnego słowa długości l, był w takim stanie, który jest podzbiorem stanów automatu $\mathcal{A}$ takim, że stany w nim będące są wszystkimi i tylko takimi stanami, do których można dojść, wczytując pewne słowo długości $l^2 - l$, zaczynając w automacie $\mathcal{A}$ w stanie q_i. Zauważmy, że to w szczególności oznacza, że zachowanie automatów $\mathcal{B}_i$ zależy wyłącznie od długości wczytywanego słowa.

Załóżmy na chwilę, że umiemy skonstruować takie automaty. Pokażę dlaczego gdyby się to udało, to zadanie byłoby rozwiązane. Ustalmy zatem słowo $w \in \{0,1\}^*$ oraz niech $k = |w|$. Wczytujemy nasze słowo do każdego z tych automatów. Dzięki pierwszemu automatowi wiemy, że po wczytaniu tego słowa jesteśmy w pewnym stanie q_i w automacie $\mathcal{A}$. Możemy zatem teraz spojrzeć na automat $\mathcal{B}_i$ oraz na podzbiór stanów automatu $\mathcal{A}$, który otrzymaliśmy. Jeśli w tym podzbiorze był jakiś stan $f \in F$, to znaczy, że w automacie $\mathcal{A}$ po wczytaniu słowa w znaleźliśmy się q_i, a później istniało słowo, które było długości $k^2 - k$ oraz gdy je przeczytaliśmy startując w stanie q_i, to doszliśmy do stanu f, który jest akceptujący. Zatem $w \in \sqrt{L}$. Oczywiście jeśli w podzbiorze otrzymanym od automatu $\mathcal{B}_i$ nie ma stanu akceptującego, to znaczy że nieważne jakie słowo będziemy wczytywać, nie uda nam się dojść w $k^2 - k$ krokach do stanu akceptującego, czyli $w \notin \sqrt{L}$.

Pozostało jedynie opowiedzieć, jak skonstruować automaty $\mathcal{B}_i$. W tym celu wprowadzę podobne, ale prostsze automaty $\mathcal{C}_i$, które pomogą mi stworzyć automaty $\mathcal{B}_i$. Niech zatem $\mathcal{C}_i$ (podobnie jak $\mathcal{B}_i$) będzie powiązany ze stanem q_i automatu $\mathcal{A}$, ale po wczytaniu słowa długości l, chcemy otrzymywać podzbiór stanów $\mathcal{A}$ takich, że można do nich dojść wczytując do automatu $\mathcal{A}$ pewne słowo długości l (a nie $l^2 - l$ jak w przypadku $\mathcal{B}_i$), zaczynając w stanie q_i. Zauważmy, że w przypadku automatu $\mathcal{C}_i$ stan po wczytaniu słowa długości $l+1$ jesteśmy w stanie jednoznacznie wyznaczyć na podstawie stanu, w którym jesteśmy po wczytaniu słowa długości l. Do nowego podzbioru stanów damy dokładnie takie stany, do których jesteśmy w stanie przejść pojedynczą krawędzią z $\mathcal{A}$ z pewnego stanu z podzbioru stanów $\mathcal{A}$ dla słowa długości l. Stanów w automacie $\mathcal{C}_i$ jest nie więcej niż 2^n, zatem po pewnym czasie wpadniemy w cykl, czyli stany tego automatu możemy opisać jako $w_0, w_1, \ldots, w_x,\ c_0, c_1, \ldots, c_{m-1}$, gdzie $w_0 = \{q_i\}$ jest stanem początkowym, a funkcja przejścia wygląda następująco:

$$\begin{cases} \delta_{\mathcal{C}}(w_j, a) = w_{j+1} & \text{dla } 0 \leqslant j < x, a \in \Sigma, \\ \delta_{\mathcal{C}}(w_x, a) = c_0 & \text{dla } a \in \Sigma, \\ \delta_{\mathcal{C}}(c_j, a) = c_{j+1 \bmod m} & \text{dla } 0 \leqslant j < m, a \in \Sigma. \end{cases}$$

Twierdzę teraz, że przy pomocy automatu $\mathcal{C}_i$, skonstruuję automat $\mathcal{B}_i$. Zauważmy, że w przypadku automatu $\mathcal{B}_i$, po wczytaniu słowa długości l, nie wystarczy pamiętać podzbioru stanów automatu $\mathcal{A}$, do którego możemy dojść w $l^2 - l$ krokach. Dzieje się tak dlatego, że w przypadku automatu $\mathcal{C}_i$ podczas przejścia chcieliśmy zrobić tylko dodatkowy jeden krok, a w przypadku automatu $\mathcal{B}_i$ chcielibyśmy

przesunąć się o $(l+1)^2-(l+1)-(l^2-l)=2l$ kroków. Musimy zatem wzbogacić stany automatu $\mathcal{C}_i$ o pewne dodatkowe informacje. Gdybyśmy przez chwilę nie martwili się liczbą stanów w automacie (tzn. że mogłaby być nieskończona), to moglibyśmy stworzyć stany $w_{0^2-0}, w_{1^1-1}, w_{2^2-2}, \ldots, w_{y^2-y}$ oraz (c_i, l) dla $0 \leqslant i < m$ oraz wszystkich $l \in \mathbb{N}$. Do stanów w_{j^2-j} nie potrzebujemy dodawać żadnej informacji, bo wiemy, że wczytane słowo miało długość dokładnie j. Jest tutaj drobny niuans, musimy rozróżniać stan $w_{0^2-0} = w_0 = w_{1^2-1}$, ale nie jest to żaden problem. Przyjrzyjmy się dokładniej funkcji przejścia dla stanów (c_j, l):

$$\delta_{\mathcal{B}}((c_j, l), a) = (c_{j+2l \bmod m}, l+1) \text{ dla } 0 \leqslant j < m, l \in \mathbb{N}, a \in \Sigma.$$

W naszym rozwiązaniu chcemy, żeby automaty $\mathcal{B}_i$ dawały odpowiednie podzbiory automatu $\mathcal{A}$, zatem zauważmy, że jeśli oznaczymy przez r_l resztę z dzielenia l przez m, to

$$j + 2l \bmod m = j + 2r_l \bmod m,$$

czyli nie musimy mieć dla każdego c_j wszystkich stanów (c_j, l), $l \in \mathbb{N}$, wystarczy stworzyć stany (c_j, l) dla $0 \leqslant j < m$ oraz $0 \leqslant l < m$ i zadać funkcję przejścia wzorem:

$$\delta'_{\mathcal{B}}((c_j, l), a) = (c_{j+2l \bmod m}, l+1 \bmod m) \text{ dla } 0 \leqslant l, j < m, a \in \Sigma.$$

Opisany automat ma skończoną liczbę stanów i po wczytaniu słowa długości l znajduje się w stanie, który jest parą, a jej pierwszym elementem jest podzbiór stanów A takich, że można do nich dojść pewnym słowem długości l^2-l zaczynając w stanie q_i. Skonstruowaliśmy zatem poszukiwane automaty $\mathcal{B}_i$, co kończy dowód.