Zadanie 29

Artur Błaszkiewicz

Wrocław, April 6, 2020

Treść zadania: 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?

Definicja: Automat $A = \langle \{0,1\}, Q, q_0, F, \delta \rangle$ to pewien automat rozpoznający język L.
Niech $s = |Q| - 1$
$Q = \{q_0, q_1, q_2, ..., q_s\}$

Definicja: Zbior $X_{i,j} = \{q_k : \exists w \in \{0,1\}^j \ \hat{\delta}(q_i, w) = q_k\}$

Lemat 1: $\forall_{i \in \{0,1,...,s\}} \ \exists_{N,k \in \mathbb{N}} \ \forall_{n>N} \ X_{i,n} = X_{i,n+k}$

Dowód lematu 1: $X_{i,j} \in \mathcal{P}(Q)$ i $|\mathcal{P}(Q)| < \infty$, więc istnieją $l, l' \in \mathbb{N}$, takie, że

$$l < l' \wedge \ X_{i,l} = X_{i,l'}$$

Załóżmy nie wprost, że istnieje $m \in \mathbb{N}$, że $X_{i,l+m} \neq X_{i,l'+m}$, wtedy istnieje takie q_x, które należy tylko do jednego ze zbiorów $X_{i,l+m}, X_{i,l'+m}$, bez utraty ogólności $q_x \in X_{i,l+m}$. Oznacza to, że istnieje takie $q_y \in X_{i,l}$ i $w \in \{0,1\}^m$, że $\hat{\delta}(q_y, w) = q_x$, ale ponieważ $X_{i,l} = X_{i,l'}$, to $q_y \in X_{i,l'}$, czyli $q_x \in X_{i,l'+m}$. **sprzeczność.**

Definicja: Niech l_i będzie najmniejszą taką liczbą naturalną, że $X_{i,l_i} = X_{i,l_i+m}$ dla jakiegoś $m > 0$ naturalnego.
Niech $d_i > 0$ będzie najmniejszą taką liczbą naturalną, że $X_{i,l_i} = X_{i,l_i+d_i}$.
Niech $r_i = l_i \ mod \ d_i$.
Niech $+_q$ oznacza dodawanie modulo q.

Definicja: Automat $A' = \langle \{0,1\}, Q', q'_{0,0,\langle 0,0\rangle,\langle 0,0\rangle,...,\langle 0,0\rangle}, F', \delta' \rangle$
$Q' = \{q'_{n,m,\langle a_0,b_0\rangle,\langle a_1,b_1\rangle,...,\langle a_s,b_s\rangle} : n \in \{0,1,...,s\}, m \in \{0,1,...,|\mathcal{P}(Q)|\},\ a_i, b_i \in \{0,1,...,d_i - 1\}\}$

$F' = \{q'_{n,m,\langle a_0,b_0\rangle,\langle a_1,b_1\rangle,...,\langle a_s,b_s\rangle} : m \in \{0,1,...,|\mathcal{P}(Q)|-1\}, \exists_{w\in\{0,1\}^{m^2-m}} \exists_{q_f\in F} \hat{\delta}(q_n, w) = q_f\}$
$\cup$
$\{q'_{n,m,\langle a_0,b_0\rangle,\langle a_1,b_1\rangle,...,\langle a_s,b_s\rangle} : m = |\mathcal{P}(Q)| \wedge \exists_{q_f\in F}\ q_f \in X_{n,l_n+a_n+(d_n-r_n)}\}$

$\delta'(q'_{n,m,\langle a_0,b_0\rangle,\langle a_1,b_1\rangle,...,\langle a_s,b_s\rangle}, x) =$
$q'_{n',m',\langle a_0+_{d_0}b_0, b_0+_{d_0}2\rangle,\langle a_1+_{d_1}b_1, b_1+_{d_1}2\rangle,...,\langle a_s+_{d_s}b_s, b_s+_{d_s}2\rangle}$

gdzie n' to numer stanu $\delta(q_n, x)$, (to znaczy $\delta(q_n, x) = q_{n'}$)
$m' = m + 1$, gdy $m < |\mathcal{P}(Q)|$, $m' = m$, gdy $m = |\mathcal{P}(Q)|$.

Idea automatu A': A' pamięta rozwiązania dla krótkich słów, natomiast dla długich pamięta stan w jakim byłby A po przeczytaniu tego słowa. Dla każdego stanu z A automat A' pamięta jakiej długości słowa mogą doprowadzić A do stanu akceptującego, jeśli zaczynałby z tego konkretnego stanu. Żeby spamiętać te długości, których jest być może nieskończenie wiele, A' wykorzystuje fakt, że dla większych długości wystarczy znać tę długość modulo. Wiedząc jakiej długości słowa potrafią doprowadzić A do stanu akceptującego A' liczy ile liter trzeba dopisać do aktualnego słowa, aby podnieść jego długość do kwadratu. Jest to dokładnie $|w|^2 - |w|$. Ponieważ pamięć jest ograniczona to pamięta tą wartość modulo. Wykorzystuje do tego prosty fakt, że $((x+1)^2 - (x+1)) - (x^2 - x) = 2x$.

Lemat 2: Dla każdego słowa w, jeżeli po przeczytaniu w, A jest w stanie q_i to wtedy i tylko wtedy A' po przeczytaniu w jest w stanie $q'_{i,min(|w|,|\mathcal{P}(Q)|),\langle a_0,b_0\rangle,\langle a_1,b_1\rangle,...,\langle a_s,b_s\rangle}$, gdzie a_j, b_j to pewne liczby.

Dowód lematu 2: Przez prostą indukcję.

Lemat 3: A' poprawnie rozpoznaje słowa z $\sqrt{L}$ krótsze niż $|\mathcal{P}(Q)|$.

Dowód lematu 3: Po przeczytaniu słowa w A' jest w pewnym stanie $q'_{n,|w|,\langle a_0,b_0\rangle,\langle a_1,b_1\rangle,...,\langle a_s,b_s\rangle}$. Oznacza to, że po przeczytaniu w A jest w stanie q_n. w należy do $\sqrt{L}$ wtedy i tylko wtedy, gdy:

$$\exists t \in \{0,1\}^{|w|^2-|w|} \; \exists q_f \in F \; \hat{\delta}(q_n, t) = q_f$$

Co dla $|w| < |\mathcal{P}(Q)|$ pokrywa się z definicją F'

Lemat 4: Jeśli po przeczytaniu w A' jest w stanie $q'_{i,j,\langle a_0,b_0\rangle,\langle a_1,b_1\rangle,...,\langle a_s,b_s\rangle}$, to $|w|^2 - |w| \; mod \; d_k = l_k + a_k + (d_k - r_k) \; mod \; d_k$ dla każdego $k \in \{0, 1, ..., s\}$

Dowód lematu 4: Przez indukcję:

Baza indukcyjna: Po przeczytaniu ϵ A' jest w stanie $q'_{0,0,\langle 0,0\rangle,\langle 0,0\rangle,...,\langle 0,0\rangle}$, $0 = l_k + a_k + d_k - r_k \; mod \; d_k = 0$

Krok indukcyjny: Jeżeli po przeczytaniu dowolnego słowa w długości x lemat zachodzi, to każde słowo długości x ma takie same a_k w stanie automatu A' po przeczytaniu tego słowa. Po przeczytaniu kolejnej literki wartość a_k zamieni się na $a'_k = a_k + 2b_k \; mod \; d_k = a_k + 2(x \; mod \; d_k) \; mod \; d_k$. Dla dowolnego słowa w' długości $x + 1$ mamy:

$$(x+1)^2 - (x+1) \; mod \; d_k = l_k + a_k + 2x + d_k - r_k \; mod \; d_k$$

$$x^2 - x + 2x \; mod \; d_k = l_k + a_k + d_k - r_k + 2x \; mod \; d_k$$

$$2x \; mod \; d_k = 2x \; mod \; d_k$$

Wniosek z lematu 4: Dla słów w spełniających $|w|^2 - |w| >= l_i$ z lematu 1 mamy:

$$X_{i,|w|^2-|w|} = X_{i,l_i+a_i+d_i-r_i}$$

Lemat 5: A' poprawnie rozpoznaje słowa z $\sqrt{L}$ nie krótsze niż $|\mathcal{P}(Q)|$

Dowód lematu 5: Po przeczytaniu słowa w ($|w| \geq |\mathcal{P}(Q)|$) A' jest w pewnym stanie $q'_{n,|\mathcal{P}(Q)|,\langle a_0,b_0\rangle,\langle a_1,b_1\rangle,...,\langle a_s,b_s\rangle}$. Oznacza to, że po przeczytaniu w A jest w stanie q_n. w należy do $\sqrt{L}$ wtedy i tylko wtedy, gdy:

$$\exists t \in \{0,1\}^{|w|^2-|w|} \; \exists q_f \in F \; \hat{\delta}(q_n, t) = q_f$$

Stan $q'_{n,|\mathcal{P}(Q)|,\langle a_0,b_0\rangle,\langle a_1,b_1\rangle,...,\langle a_s,b_s\rangle}$ jest akceptujący wtedy i tylko wtedy, gdy:

$$\exists_{q_f \in F} \; q_f \in X_{n,l_n+a_n+(d_n-r_n)}$$

, co z wnioskiem z lematu 4 daje:

$$\exists_{q_f \in F}\ q_f \in \{q_k : \exists t \in \{0,1\}^{|w|^2-|w|} \hat{\delta}(q_n, t) = q_k\}$$

Oznacza to, że w należy do $\sqrt{L}$ wtedy i tylko wtedy, gdy stan $q'_{n,|\mathcal{P}(Q)|,\langle a_0,b_0\rangle,\langle a_1,b_1\rangle,\ldots,\langle a_s,b_s\rangle}$ jest akceptujący.