# Języki Formalne i Złożoność Obliczeniowa
# Lista zadań nr 2

Kwiatek

2 kwietnia 2020

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?

- Najpierw podzielimy zupełnie przestrzeń słów nad alfabetem $\Sigma = \{0,1\}^*$ na regularne języki $P_q$ (tak, by wszystkie języki $P$ sumowały się do $\Sigma$, oraz by każdy $P$ był językiem regularnym).
- Następnie postaramy się przemianować zbiór podany w zadaniu na język, do którego potrafimy prosto skonstruować automat.
- Skonstruujemy automat rozpoznający $\sqrt{L} \cap P_q$.
- Ostatecznie wywnioskujemy, $\sqrt{L}$ jest regularny.

## 1 Podział przestrzeni

Weźmy automat $A = \langle Q, \Sigma, q_0, F, \delta \rangle$ rozpoznający L. Niech $P_q = \{w \in \Sigma : \delta(q_0, w) = q\}$. Czyli, $P_q$ jest zbiorem słów, że po wczytaniu dowolnego z nich do automatu $A$ dostajemy stan $q$. Zauważmy, że $\bigcup_{q \in Q} P_q = \Sigma$. Ponieważ dla dowolnego słowa wczytanego do automatu, skończy się ono w pewnym stanie q. A oczywiście $P_q$ jest podzbiorem $\Sigma^*$. Zauważmy też, że $P_q$ jest regularny, ponieważ możemy stworzyć automat podobny do $A$, ale z jednym stanem akceptującym $F = \{q\}$. Zdefiniujmy przy okazji język $S_q$, który bierze sufiksy słów z L, zaczynając od stanu q. Możemy zatem stworzyć automat podobny do $A$, ale stan początkowy $q_0 = q$. Regularność języka wynika z konstrukcji automatu.

## 2 Język $\sqrt{L}$

Weźmy przekrój języka $\sqrt{L}$ oraz $P_q$. Uzasadnimy następujący ciąg przejść:
$\sqrt{L} \cap P_q = \{w \in P_q : \exists x \in \{0,1\}^* \; \exists y \in L \; wx = y \wedge |y| = |w|^2\}$
Pierwsze przejście jest przepisaniem definicji w nowy sposób - mówimy, że szukamy $y$ równego słowu $w$ skonkatenowanym z elementem z $\Sigma^*$.
$\sqrt{L} \cap P_q = \{w \in P_q : \exists y \in L \; y \in w\Sigma^* \wedge |y| = |w|^2\}$
Dalej, zapewniamy, że $y$ ma odpowiednią długość, wybierając tylko niektóre elementy z $\Sigma^*$.
$\sqrt{L} \cap P_q = \{w \in P_q : \exists y \in L \; y \in w\Sigma^{|w|^2 - |w|}\}$
Zamiast szukać $y$ należących do języka, szukamy słów s, które pod skonskatenowaniu z $w$ dadzą nam słowo z L.
$\sqrt{L} \cap P_q = \{w \in P_q : \exists s \in \Sigma^* \; |s| = |w|^2 - |w| \wedge \delta(q, s) \in F\}$
Ostatecznie aplikujemy definicję języka sufiksów. Jako że startujemy w $q$ wiemy, że $s$ należy do $S_q$.
$\sqrt{L} \cap P_q = \{w \in P_q : \exists s \in S_q \; |s| = |w|^2 - |w|\}$

## 3 Konstrukcja automatu rozpoznającego $\sqrt{L} \cap P_q$

Mamy już automat $A_q$ rozpoznający $P_q$ - został opisany w części pierwszej.
Szukamy automatu, którym po wczytaniu słowa $w$ sprawdzi, czy istnieje słowo długości $|w|^2 - |w|$ w języku $S_q$.
Łatwo możemy stworzyć automat sprawdzający, czy słowo długości $|w|$ należy do $S_q$. Mianowicie weźmy automat rozpoznający $S_q$ i wszystkie krawędzie przemianujmy na cały alfabet $\Sigma$. Wtedy po wczytaniu dowolnych n liter będziemy w zbiorze stanów. Jeżeli do tego zbioru należy stan akceptujący, to istnieje słowo o długości n akceptowane przez $S_q$.
Z drugiej strony wiemy, że jeżeli regularny język nad jednoliterowym alfabetem (w którym jedyną różnicą pomiędzy słowami jest ich długość) będzie skończony lub rozpoznające go DFA -reprezentowane jako graf - będzie kończyć się cyklem. Oznaczmy jako $p$ długość tego cyklu. Zauważmy, że istnieje $n_0$, że dla dowolnych słów długości $n > n_0$, $0^n$ jest akceptowane przez ten DFA wtedy i tylko wtedy, gdy $0^{n+p}$ też jest akceptowane - po wykonaniu pełnego cyklu jesteśmy w tym samym stanie.
Łatwo zauważyć, że DFA rozpoznające język regularny nad jednoliterowym alfabetem musi spełniać powyższą własność. Pozwolę sobie to uzasadnić przedstawiając szkic dowodu nie wprost. Weźmy nieskończony język na alfabetem jednoelementowym. Załóżmy nie wprost, że nie ma w nim cykli. Ale wtedy weźmy słowo długości $n$ (gdzie $n$ jest dowolnie duże). Wtedy istnieje ścieżka długości $n$ w tym grafie, która jest w każDym stanie najwyżej raz. Z dowolności $n$ istnieje nieskończenie długa ścieżka, co prowadzi do sprzeczności ze skończonością naszego automatu.
Weźmy zatem język: $\{|w| : |w|^2 - |w| \in L\}$ dla pewnego języka regularnego L. Pokażemy, że język ten jest regularny.

Zauważmy fakt dotyczący funkcji $n^2 - n$:
$n^2 - n = (n+p)^2 - (n+p) \ mod \ p$
$n^2 - n = n^2 + 2np + p^2 - n - p \ mod \ p$
$n^2 - n = n^2 - n \ mod \ p$

Weźmy zatem $n_0, p$ oraz język regularny $G = \{0^{|w|} : w \in S_q\}$ dla których spełnione jest:
$n > n_0 \Rightarrow 0^n \in G \Leftrightarrow 0^{n+p} \in G$
Co za tym idzie dla dowolnego $n_1$ i $n_2$
$n_1 = n_2 \ mod \ p \Rightarrow 0^{n_1} \in G \Leftrightarrow 0^{n_2} \in G$
Ale z powyższego faktu implikuje to
$0^{n^2-n} \in G \Leftrightarrow 0^{(n+p)^2-(n+p)} \in G$
Zatem $n > n_0 \Rightarrow n \in \{|w| : |w|^2 - |w| \in G\} \Leftrightarrow n + p \in \{|w| : |w|^2 - |w| \in G\}$
Powyższy język jest więc regularny. Zatem z definicji G mamy język regularny $\overline{S_q} = \{|w| : \exists s \in S_q |s| = |w|^2 - |w|\}$ Weźmy zatem przekrój języków $P_q$ i $\overline{S_q}$.
Przekrój języków regularnych także jest regularny.
Zatem zaakceptujemy tylko takie słowa, które są w $P_q$ oraz mają odpowiednią długość czyli są w $\overline{S_q}$.

## 4 Wniosek

Weźmy zatem sumę wszystkich języków przedstawionych powyżej:
$\bigcup_{q \in Q} \sqrt{L} \cap P_q = \sqrt{L}$
Co wynika z dowolności doboru pośredniego stanu p. A skoro jest to język, który jest sumą języków regularnych, to sam jest regularny.