Złożoność obliczeniowa (bioinformatyka) - egzamin
3 lutego 2020

1. Które z następujących języków są regularne? Które są bezkontekstowe, a które są nierozstrzygalne? Do jakich klas złożoności należą te języki?

   (a) $L_1 = \{a^m b^n a^k b^\ell \mid m, n, k, \ell \in \mathbb{N}\}$;

   (b) $L_2 = \{a^m b^n a^k b^\ell \mid m, n, k, \ell \in \mathbb{N} \wedge m \leq \ell \wedge n \leq k\}$;

   (c) $L_3 = \{a^m b^n a^k b^\ell \mid m, n, k, \ell \in \mathbb{N} \wedge m \leq k \wedge n \leq \ell\}$.

2. Rozpatrzmy następujące problemy decyzyjne:

   (a) Dana formuła zdaniowa φ. Czy φ ma co najmniej dwa wartościowania spełniające?

   (b) Dany graf G i liczba k. Czy istnieje taki k-elementowy zbiór S wierzchołków grafu G, że żadne dwa wierzchołki ze zbioru S nie są połączone krawędzią?

   (c) Dany graf G i liczby d, k. Czy istnieje taki k-elementowy zbiór S wierzchołków grafu G, że każdy z pozostałych wierzchołków jest osiągalny z jakiegoś wierzchołka należącego do S drogą składającą się z co najwyżej d krawędzi?

   Czy któryś z tych problemów jest rekurencyjnie przeliczalny, NP-zupełny, NL-zupełny, rozstrzygalny?

3. Czy jeśli $\textsc{Ntime}(n^2) \subseteq \mathrm{P}$, to $\mathrm{P} = \mathrm{NP}$?

Przykładowe rozwiązania

1: Wszystkie trzy języki są oczywiście rozstrzygalne i mieszczą się w klasie LOGSPACE. Do ich rozpoznawania potrzebne są dwa liczniki służące do zliczania liter a i b. Wartości liczników nie przekraczają długości słowa wejściowego, wystarczy więc logarytmiczna pamięć.

1a: Język L_1 jest regularny, bo to po prostu język $a^*b^*a^*b^*$.

1b: Język L_2 nie jest regularny, bo ma nieskończenie wiele różnych ilorazów, w tym np. języki:

$$L_2 \backslash a^d = \{a^m b^n a^k b^\ell \mid m, n, k, \ell \in \mathbb{N} \wedge m + d \leq \ell \wedge n \leq k\},$$

dla dowolnego $d \in \mathbb{N}$. Jest to język bezkontekstowy generowany przez gramatykę

$$\xi_0 ::= a\xi_0 b \mid \xi_1, \quad \xi_1 ::= \xi_1 b \mid \xi_2, \quad \xi_2 ::= b\xi_2 a \mid \xi_3, \quad \xi_3 ::= \xi_3 a \mid \varepsilon$$

1c: Ten język nie jest bezkontekstowy. W przeciwnym razie niech N będzie stałą z lematu o pompowaniu i niech $w = a^N b^N a^N b^N$. Ponieważ $w \in L_3$ i $|w| \geq N$, więc $w = xyzuv$ dla pewnych słów x, y, z, u, v, gdzie $|yzu| \leq N$, $yu \neq \varepsilon$ oraz każde słowo $xy^k zu^k v$ należy do L_3. Przypuśćmy, że w słowie yu jest choć jedna litera b (w przeciwnym razie jest jakieś a i rozumujemy podobnie). Ponieważ fragment yzu jest krótki, więc wszystkie litery b w słowie yu pochodzą z tego samego segmentu złożonego z N liter b. Jeśli jest to pierwszy segment, to słowo xy^2zu^2v ma za dużo liter b w pierwszej części aby należało do L_3. A jeśli jest to drugi segment, to słowo xy^0zu^0v, czyli słowo xzv ma w drugiej części za mało liter b. W obu przypadkach mamy sprzeczność z lematem o pompowaniu.

2: Wszystkie trzy problemy są NP-zupełne, a zatem oczywiście rozstrzygalne a tym bardziej rekurencyjnie przeliczalne. Jeśli któryś z nich jest NL-zupełny, to NL = P = NP.

2a: Do tego problemu sprowadza się problem SAT. Jeśli ψ jest dowolną formułą i zmienna zdaniowa r nie występuje w ψ, to ψ jest spełnialna wtedy i tylko wtedy, gdy formuła $r \vee \psi$ ma co najmniej dwa wartościowania spełniające.

2b: Do tego problemu sprowadza się problem kliki. Zbiór S wierzchołków grafu G jest kliką wtedy i tylko wtedy, gdy w grafie dualnym $\overline{G}$ żadne dwa jego wierzchołki nie są połączone krawędzią.

2c: Szczególnym przypadkiem tego problemu (dla $d = 1$) jest problem pokrycia wierzchołkowego.

3: Tak. Przypuśćmy, że $\text{NTIME}(n^2) \subseteq \text{P}$; udowodnimy, że P=NP. Weźmy dowolny język $L \in \text{NP}$, powiedzmy $L \in \text{NTIME}(n^{2k})$. Wtedy język $L' = \{w\$^{|w|^k - |w|} \mid w \in L\}$ należy do $\text{NTIME}(n^2)$, bo słowa postaci $w\$^{|w|^k - |w|}$ długości $n = |w|^k$, można rozpoznawać w czasie $|w|^{2k} = n^2$. Skoro $\text{NTIME}(n^2) \subseteq \text{P}$, to także $L' \in \text{P}$, tj. $L' \in \text{DTIME}(n^r)$, dla pewnego r. Wtedy jednak $L \in \text{DTIME}(n^{rk}) \subseteq \text{P}$, bo słowo $w \in L$ o długości n można rozpoznawać tak samo jak słowo $w\$...\$$ długości n^k, czyli w czasie $(n^k)^r$.