ZLOŻONOŚĆ OBLICZENIOWA - WYK. 9

1. Tw. Baker-Gill-Solovay (1975): Istnieją języki A i B takie, że:

$$\mathbf{P}^A = \mathbf{NP}^A \qquad \text{i} \qquad \mathbf{P}^B \neq \mathbf{NP}^B.$$

Dowód: Za A wystarczy wziąć QBF. Porachujmy:

$$\mathbf{NP}^{\mathrm{QBF}} \subseteq \mathbf{NPSPACE} = \mathbf{PSPACE} = \mathbf{P}^{\mathrm{QBF}}.$$

Po kolei kroki:

- zamiast pytać wyroczni QBF o słowo, maszyna może sama obliczyć odpowiedź: pytanie jest wielomianowo długie, więc mieści się w wielomianowej pamięci, a QBF jest rozwiązywalny w wielomianowej pamięci.
- tw. Savitcha,
- ostatnia inkluzja wynika bezpośrednio z **PSPACE**-zupełności (w sensie Karpa) problemu QBF.
- ćwiczenie: sprawdzić co się rozsypuje w tym dowodzie, jeśli zamiast QBF weźmiemy SAT.

Teraz skonstruujemy pewną wyrocznię B i rozważymy język

$$L = \{1^n \mid \text{pewne słowo } w \text{ długości } n \text{ jest w } B\}.$$

Po pierwsze, $L \in \mathbf{NP}$, bo maszyna może zgadnąć właściwe słowo w.

Deterministyczna maszyna rozpoznająca L ma problem: może tylko pytać wyroczni o kolejne słowa i nie starczy jej czasu by sprawdzić wszystkie. Trzeba tylko zaprojektować B tak, by faktycznie nie dało się zrobić nic mądrzejszego.

Wybierzmy enumerację $M_1, M_2, M_3, \ldots$ wszystkich maszyn Turinga z wyrocznią, działających w czasie wielomianowym (wyrocznia nie jest częścią definicji maszyny), podobnie jak w dowodzie tw. Ladnera. Będziemy konstruować język B po trochu, oszukując kolejne maszyny.

Skonstruujemy języki $B_1 \subseteq B_2 \subseteq B_3 \cdots B$ i długości $n_1 \leq n_2 \leq n_3 \leq \cdots$ takie, że:

- $B = \bigcup_{i \in \mathbb{N}} B_i$,
- $M_i^{B_i}$ źle rozpoznaje słowo 1^{n_i},
- M_i^B zgadza się z $M_i^{B_i}$ na słowie 1^{n_i}.

Zaczynamy od $B_0 = \emptyset$. Potem dla $i = 1, 2, \ldots$:

- wybieramy n_i takie duże, żeby żadna maszyna M_j (dla $j < i$) na słowie 1^{n_j} nie była w stanie wyprodukować zapytania długości n_i (to gwarantuje, że oszukane wcześniej maszyny nadal są oszukane), oraz żeby M_i na słowie 1^{n_i} działała na pewno w mniej niż 2^{n_i} krokach.

- uruchom M_i z wyrocznią B_{i-1} na słowie 1^{n_i}. Zauważmy, że ta wyrocznia na każde pytanie długości n_i odpowiada negatywnie.
- jeśli M_i zaakceptowała, to weź $B_i = B_{i-1}$. Wtedy $1^{n_i} \notin L$ i oszukaliśmy maszynę M_i.
- jeśli M_i odrzuciła, to znajdź słowo w długości n_i, o które maszyna M_i nie zapytała (takie słowo musi być, bo M_i zrobiła mniej niż 2^{n_i} kroków) i ustal $B_i = B_{i-1} \cup \{w\}$. Wtedy $1^{n_i} \in L$ i oszukaliśmy maszynę M_i.

Język B jest obliczalny (w czasie wykładniczym), ale dla tego twierdzenia to nieważne.

2. maszyny ze źródłem bitów losowych (probabilistyczne)
   - maszyna deterministyczna
   - dodatkowa taśma do odczytu, na której taśma porusza się tylko w prawo, za każdym razem o jeden krok
3. Redefinicja **NP**: język K jest w **NP** jeżeli istnieje wielomian $p(n)$ i maszyna M ze źródłem bitów, która zawsze działa w dokładnie $p(n)$ krokach, taka że:
   - $w \in L \implies \exists s \in \{0,1\}^{p(|w|)}.\ (w,s) \in L_M$
   - $w \notin L \implies \nexists s.\ (w,s) \in L_M$

   Intuicja: słowo jest w L wtw. jakiś świadek to potwierdza.
4. Klasa **RP**, czyli *randomized polynomial time*: ta sama definicja, ale warunki zmieniają się na:
   - $w \in L \implies Pr_s[(w,s) \in L_M] \geq \frac{1}{2}$
   - $w \notin L \implies \nexists s.\ (w,s) \in L_M$

   Intuicja: słowo jest w L jeśli co najmniej połowa świadków to potwierdza. Innymi słowy: losując świadka, jeśli słowo nie jest w L to na pewno odrzucimy, a jeśli jest w L to zaakceptujemy z prawdopodobieństwem co najmniej $\frac{1}{2}$.

   Uwaga: Są maszyny M, które nie akceptują żadnego języka w sensie **RP**. To, czy dana maszyna jest dobra w sensie **RP**, jest nierozstrzygalne, nawet jeśli znamy wielomian $p(n)$. Mówimy, że klasa **RP** jest "semantyczna", a nie "syntaktyczna", bo nie mamy łatwo rozpoznawalnego modelu obliczeń, który jej odpowiada. To się wiąże z tym, że w **RP** nie ma problemów zupełnych.

   Klasa **coRP** to dopełnienia języków z **RP**, czyli te języki, w których akceptując mamy rację, a odrzucając mamy rację z prawdopodobieństwem co najmniej $\frac{1}{2}$.

5. Fakt: $\mathbf{P} \subseteq \mathbf{RP} \subseteq \mathbf{NP}$. Obie inkluzje są trywialne.

6. Fakt (*amplifikacja*): w definicji **RP** liczbę $\frac{1}{2}$ można zmienić na dowolną liczbę $q \in (0,1)$, a klasa języków pozostanie ta sama.

   Dowód: Niech $\mathbf{RP}_p$ będzie klasą z prawdopodobieństwem błędu p, tj. taką, gdzie zamiast $\frac{1}{2}$ jest $1-p$.

   Jest jasne, że jeśli $p \leq q$ to $\mathbf{RP}_p \subseteq \mathbf{RP}_q$.

   Pokażemy, że $\mathbf{RP}_p \subseteq \mathbf{RP}_{p^2}$. Dla maszyny M z błędem p konstruujemy maszynę M', która na tym samym słowie w niezależnie losuje dwa świadki i akceptuje jeśli któryś z nich zaakceptuje. Maszyna działa dwa razy dłużej (a więc wielomianowo długo), a prawdopodobieństwo błędu spada do p^2: mylimy się tylko jeśli błędnie odrzucimy, czyli jeśli maszyna M błędnie odrzuciła dwa razy.

7. Przykład problemu w **coRP** to tożsamość wielomianu: czy dany (w jakiejś niejawnej postaci) wielomian jest tożsamościowo równy zeru? Formalnie, dany jest *obwód arytmetyczny*, gdzie każda bramka wejściowa odpowiada zmiennej, a bramki są typu $+$, $\times$ i unarna bramka $-$. Pytamy czy powstały wielomian jest stale równy zero. Problem można rozważać nad ciałem liczb wymiernych, albo jakimś skończonym ciałem. Tak czy inaczej nie wiadomo czy ten problem jest w **P**.

8. Lemat Schwartza-Zippela: niech p będzie niezerowym wielomianem k zmiennych, całkowitego stopnia d, nad ciałem F (skończonym lub nie), niech S będzie skończonym podzbiorem tego ciała i niech $r_1, \ldots, r_k$ będą losowo wybranymi elementami S. Wtedy

$$Pr[p(r_1, \ldots, r_k) = 0] \leq \frac{d}{|S|}.$$

   Dowód: przez indukcję po k. Dla $k = 1$ to po prostu twierdzenie Bezout: wielomian stopnia d może mieć tylko d pierwiastków.

   Dla kroku indukcyjnego, przedstawmy p jako wielomian jednej zmiennej x_1:

$$p(x_1, \ldots, x_k) = \sum_{i=0}^{d} x^i p_i(x_2, \ldots, x_k)$$

   Skoro p jest niezerowy, to istnieje i takie że p_i jest niezerowe. Weźmy największe takie i. Wtedy $x^i p_i$ ma stopień najwyżej d, więc stopień p_i to co najwyżej $d - i$. Z założenia indukcyjnego:

$$Pr[p_i(r_2, \ldots, r_k) = 0] \leq \frac{d-i}{|S|}.$$

   Jeśli $p_i(r_2, \ldots, r_k) \neq 0$ to $p(x_1, r_2, \ldots, r_k)$ jest stopnia i (bo wzięliśmy największe i), więc

$$Pr[p(r_1, \ldots, r_n) = 0 \mid p_i(r_2, \ldots, r_k) \neq 0] \leq \frac{i}{|S|}$$

Dla każdych zdarzeń A, B zachodzi

$$Pr[A] \leq Pr[B] + Pr[A|B^c],$$

więc

$$Pr[p(r_1, \ldots, r_k) = 0] \leq Pr[p_i(r_2, \ldots, r_k) = 0] + Pr[p(r_1, \ldots, r_n) = 0 \mid p_i(r_2, \ldots, r_k) \neq 0] = \frac{d}{|S|}$$

9. Lemat Schwartza-Zippela pokazuje, że problem *Czy dany wielomian jest niezerowy?*, przynajmniej nad dużym skończonym ciałem, jest w **RP**: świadek to losowe wartościowanie zmiennych.

   Dokładniej: obwód arytmetyczny o n bramkach definiuje wielomian o stopniu całkowitym co najwyżej 2^n. Jeśli jest k zmiennych, to wystarczy wylosować k liczb z przedziału $0..10 \cdot 2^n$ (trzeba na to $\mathcal{O}(k \cdot n)$ bitów), obliczyć na nich wartość wielomianu (czyli obliczyć obwód) i zaakceptować jeśli wyszło 0. Pomylimy się z prawdopodobieństwem najwyżej $\frac{1}{10}$.

   Nad ciałem liczb wymiernych pojawia się dodatkowy problem: jak obliczyć wartość wielomianu w czasie wielomianowym? Nawet jeśli jest zerowy, to wartości pośrednie mogą być wykładniczo długie. *Rozwiązanie:* trzeba wylosować liczbę m z przedziału $2..2^{2n}$ i wszystko liczyć modulo m. Jeśli wartość wielomianu $Y = p(r_1, \ldots, r_k)$ miała wyjść 0, to modulo m też wyjdzie 0. Z kolei jeśli wyszła niezerowa, to pokażemy że z prawdopodobieństwem $\frac{1}{10n}$ nie dzieli się przez m.

   Istotnie, z twierdzenia o liczbach pierwszych, liczba m jest pierwsza z prawdopodobieństwem co najmniej $\frac{1}{5n}$. (Twierdzenie mówi, że

$$\lim_{n \to \infty} \frac{\pi(n)}{n/\ln n} = 1$$

   gdzie $\pi(n)$ to liczba liczb pierwszych mniejszych od n). Z kolei liczba Y może mieć wartość najwyżej $(10 \cdot 2^n)^{2^n}$. Liczba jej dzielników pierwszych to najwyżej logarytm z tego, czyli mniej niż $5n2^n$. Tak więc losowo wybrane m jest wśród tych dzielników z prawdopodobieństwem najwyżej

$$\frac{5n2^n}{2^{2n}} < \frac{1}{10n}.$$

   Z tego wynika, że prawdopodobieństwo że k jest pierwsza *i nie jest* dzielnikiem Y jest co najmniej $\frac{1}{10n}$.

   No dobra, ale wciąż mylimy się z prawdopodobieństwem $1 - \frac{1}{10n}$. Co dalej? Wystarczy powtórzyć cały algorytm niezależnie $\mathcal{O}(n)$ razy. Pomylimy się tylko jeśli za każdym razem się pomyliliśmy. To działa, bo wiadomo że

$$\lim_{n \to \infty} (1 - \frac{1}{n})^n = \frac{1}{e}$$

   więc dla dużych n prawdopodobieństwo błędu spadnie poniżej $\frac{9}{10}$, a dalej już zwykłą amplifikacją.

10. klasa **PP**, czyli *probabilistic polynomial*: tak samo jak **RP**, ale:

    - $w \in L \Longrightarrow Pr_s[(w,s) \in L_M] \geq \frac{1}{2}$
    - $w \notin L \Longrightarrow Pr_s[(w,s) \in L_M] < \frac{1}{2}$

    Intuicja: akceptacja przez głosowanie świadków. Zalety tej klasy:

    - symetryczne pomyłki
    - klasa "syntaktyczna", i faktycznie ma problem zupełny MAJSAT: czy dana formuła boolowska jest spełniana przez co najmniej połowę wartościowań?

    Wada tej klasy: jest bardzo duża i trudno ją uznać za klasę problemów efektywnie rozpoznawalnych. Mianowicie $\mathbf{NP} \subseteq \mathbf{PP}$. *Dowód:* ćwiczenie.

11. klasa **BPP**, czyli *bounded probabilistic polynomial*: błędy symetryczne, ale małe:

    - $w \in L \Longrightarrow Pr_s[(w,s) \in L_M] \geq \frac{3}{4}$
    - $w \notin L \Longrightarrow Pr_s[(w,s) \in L_M] \leq \frac{1}{4}$

    Innymi słowy, prawdopodobieństwo pomyłki to co najwyżej $\frac{1}{4}$ w każdą stronę.

12. proste fakty jako ćwiczenia: $\mathbf{RP} \subseteq \mathbf{BPP} \subseteq \mathbf{PP}$. Otwarty problem: jak **BPP** ma się do **NP**? Hipoteza: $\mathbf{BPP} = \mathbf{P}$.

13. Amplifikacja dla **BPP**: zamiast $\frac{1}{4}$ można wziąć dowolną liczbę $p \in (0, \frac{1}{2})$.

    Dowód: Niech początkowe prawdopodobieństwo błędu to p. Zróbmy $2m+1$ niezależnych eksperymentów i niech zdecyduje większość. Fakt: prawdopodobieństwo błędu spada do $(4p(1-p))^m$. Dla $p = \frac{1}{4}$, wychodzi $(\frac{3}{4})^m$.

    Dowód faktu: niech prawdopodobieństwo sukcesu jakiegoś eksperymentu będzie $q > \frac{1}{2}$. Wykonajmy eksperyment niezależnie $2m+1$ razy. Prawdopodobieństwo, że uzyskamy sukces najwyżej m razy, szacuje się z góry przez:

    $$\sum_{i=0}^{m} \binom{2m+1}{i} q^i (1-q)^{2m-i+1} \leq \sum_{i=0}^{m} \binom{2m+1}{i} q^m (1-q)^m = 2^{2m}(q(1-q))^m$$