Politechnika Wrocławska

Struktury danych i złożoność obliczeniowa
Wykład 5

Prof. dr hab. inż. Jan Magott

DMT rozwiązuje problem decyzyjny π przy kodowaniu e w co najwyżej wielomianowym czasie, jeśli dla wszystkich łańcuchów wejściowych $x(I)$ takich, że $I \epsilon D_{\pi}$ zatrzymuje się po czasie działania $t \leqslant p(|x(I)|)$ dla każdego $x(I)$ i pewnego wielomianu p, oraz kończy obliczenia w stanie q_y, dla wszystkich $x(I)$ takich, że $I \epsilon Y_{\pi}$ i tylko dla nich.

NDMT rozwiązuje problem decyzyjny π , jeśli dla każdej instancji $I \epsilon D_{\pi}$ są spełnione warunki:

- Jeśli odpowiedź dla I brzmi „tak”, to zostanie wygenerowany łańcuch S , który wraz z $x(I)$ spowoduje, że po wykonaniu programu przez NDMT maszyna ta osiągnie stan końcowy q_{tak} ,
- Jeśli odpowiedź dla I brzmi „nie”, to dla każdego wygenerowanego łańcucha S albo NDMT osiągnie stan końcowy q_{nie}, albo etap sprawdzania nie zostanie zakończony.

NDMT rozwiązuje problem decyzyjny π w (co najwyżej) **wielomianowym czasie**, jeśli dla każdej instancji $I \epsilon D_{\pi}$, dla której odpowiedź brzmi „tak”, zostanie wygenerowany taki łańcuch S , że czas wykonania etapów zgadywania i sprawdzania zakończonego odpowiedzią „tak” przez NDMT (dla I oraz S) jest $O\left(p\left(N(I)\right)\right)$ dla pewnego wielomianu p .

Klasę $\boldsymbol{P}$ tworzą wszystkie problemy decyzyjne, które w co najwyżej wielomianowym czasie może rozwiązać DMT.

Klasa $\boldsymbol{NP}$ zawiera wszystkie problemy decyzyjne, które w co najwyżej wielomianowym czasie może rozwiązać NDMT.

$$P \subseteq NP$$

Ze względu na wiele lat nieudanych prób udowodnienia relacji $\mathrm{P} = NP$, jest prawie pewne, że:

$$P \subset NP$$

(jest prawie pewne, że P jest właściwą podklasą klasy NP).

Jednak czy $P \subset NP$ jest problemem otwartym. ?

Intuicyjne pojmowanie klas P i NP

Klasa P zawiera te wszystkie problemy decyzyjne, dla których znaleziono wielomianowe algorytmy ich rozwiązania.

Klasa NP zawiera te wszystkie problemy decyzyjne, dla których znaleziono ponadwielomianowe algorytmy ich rozwiązania (w wielomianowym czasie można odgadnąć rozwiązanie i sprawdzić czy to rozwiązanie daje odpowiedź "tak").

- Pytanie, czy problemy NP-zupełne można rozwiązywać w czasie wielomianowym, jest największą zagadką informatyki teoretycznej.
- Problem $P = NP$ czy $P \neq NP$ jest problemem otwartym umieszczonym na liście Problemów milenijnych.

- **Problemy milenijne** (ang. Millennium Prize Problems) - zestaw siedmiu zagadnień matematycznych ogłoszonych przez Instytut matematyczny Claya 24 maja 2000 roku.

- Za rozwiązanie każdego z nich wyznaczono milion dolarów nagrody.

- Do dziś rozwiązano tylko jeden: hipoteza Poincarego została potwierdzona w 2006 przez rosyjskiego matematyka Grigorija Perelmana.

Transformacją wielomianową problemu $\boldsymbol{\pi_2}$ do problemu $\boldsymbol{\pi_1}$ ($\pi_2 \propto \pi_1$) jest funkcja $f: D_{\pi_2} \rightarrow D_{\pi_1}$ spełniająca warunki:

1. Dla każdej instancji $I_2 \epsilon D_{\pi_2}$ odpowiedź brzmi „tak", wtedy i tylko wtedy, gdy dla instancji $f(I_2)$ odpowiedź również jest „tak",

2. Czas obliczenia funkcji f przez DMT dla każdej instancji $I_2 \epsilon D_{\pi_2}$ jest ograniczony od góry przez wielomian od $N(I_2)$.

Własności transformacji wielomianowej

Lemat 1 Transformacja wielomianowa jest przechodnia, tzn. jeśli $\pi_2 \propto \pi_1$ i $\pi_3 \propto \pi_2$, to $\pi_3 \propto \pi_1$.

Lemat 2 Jeżeli $\pi_2 \propto \pi_1$ i $\pi_1 \epsilon NP$, to $\pi_2 \epsilon NP$.

Lemat 3 Jeżeli $\pi_2 \propto \pi_1$ i $\neg \pi_2 \epsilon NP$, to $\neg \pi_1 \epsilon NP$.

Wniosek Jeżeli $\pi_2 \propto \pi_1$, to problem π_1 jest co najmniej tak trudny jak π_2.

Problem decyzyjny π_1 jest nazywany ***NP-zupełnym***, jeśli:

1. $\pi_1 \epsilon NP$,
2. Dla każdego innego problemu decyzyjnego $\pi_2 \epsilon NP$ jest $\pi_2 \propto \pi_1$.

Zatem, jeśli istniałby algorytm wielomianowy do rozwiązywania jakiegokolwiek problemu NP-zupełnego, to każdy problem z klasy NP (w tym również problemy NP-zupełne) mógłby być rozwiązany za pomocą algorytmu wielomianowego.

Z bezskuteczności poszukiwań algorytmu wielomianowego dla któregokolwiek problemu NP-zupełnego wynika, że prawie na pewno wszystkie problemy NP-zupełne można rozwiązać tylko przy użyciu algorytmów ponadwielomianowych.

Wnioski

1. Klasa problemów NP-zupełnych zawiera **problemy równoważne wielomianowo**, tzn. jeśli π_1 jest NP-zupełny i π_2 jest NP-zupełny, to $\pi_2 \propto \pi_1$ i $\pi_1 \propto \pi_2$.

2. Klasa problemów NP-zupełnych zawarta jest w klasie NP.

3. Jeśli dla pewnego problemu NP-zupełnego istnieje wielomianowy algorytm rozwiązania, to wszystkie problemy NP-zupełne są rozwiązywalne w czasie wielomianowym.

4. Klasa problemów NP-zupełnych zawiera najtrudniejsze problemy z klasy NP.

Podsumowanie

P – klasa problemów rozwiązywalnych w czasie wielomianowym.

NP – klasa problemów nie rozwiązywalnych w czasie wielomianowym.

Problemy otwarte to takie, dla których nie znaleziono algorytmu wielomianowego rozwiązania ani nie wykazano NP-zupełności.

Problem decyzyjny π_1 jest nazywany ***NP-zupełnym***, jeśli:

1. $\pi_1 \epsilon NP$,
2. Dla każdego innego problemu decyzyjnego $\pi_2 \epsilon NP$ jest $\pi_2 \propto \pi_1$.

Zatem, jeśli istniałby algorytm wielomianowy do rozwiązywania jakiegokolwiek problemu NP-zupełnego, to każdy problem z klasy NP (w tym również problemy NP-zupełne) mógłby być rozwiązany za pomocą algorytmu wielomianowego czyli $NP = P$.

Z bezskuteczności poszukiwań algorytmu wielomianowego dla któregokolwiek problemu NP-zupełnego (chyba $NP \neq P$) wynika, że **prawie na pewno wszystkie problemy NP-zupełne można rozwiązać tylko przy użyciu algorytmów ponadwielomianowych.**

Jak wykazać

2. Dla każdego innego problemu decyzyjnego $\pi_2 \epsilon NP$ jest $\pi_2 \propto \pi_1$.

?

Niech π_2 będzie dowolnym problemem NP-zupełnym.

Zatem każdy problem $\pi_3 \in NP$ można przetransformować wielomianowo w π_2 tzn. $\pi_3 \propto \pi_2$ (z definicji).

Założmy, że wykażemy $\pi_2 \propto \pi_1$

Stąd z przechodniości relacji $\propto$ mamy

$(\pi_3 \propto \pi_2 \wedge \pi_2 \propto \pi_1) \Rightarrow \pi_3 \propto \pi_1$
wykażemy, że każdy problem $\pi_3 \in NP$ można przetransformować wielomianowo w π_1.

Wystarczy wykazać, że $\boldsymbol{\pi_2 \propto \pi_1}$.

Do udowodnienia NP-zupełności problemu decyzyjnego π wystarczy:

1. Dowieść, że $\pi \epsilon NP$,
2. Przetransformować wielomianowo dowolny znany problem NP-zupełny do problemu π.

W celu zbadania złożoności obliczeniowej danego problemu, staramy się znaleźć dla niego optymalny deterministyczny algorytm wielomianowy lub wykazać trudność tego problemu. Aby wykazać trudność, wystarczy udowodnić NP-zupełność.

Do klasy problemów NP-zupełnych należą najtrudniejsze problemy klasy NP.

Fundamentalne dla teorii złożoności obliczeniowej jest Twierdzenie Cook'a, które wskazuje pewien problem jako NP-zupełny.

Twierdzenie Cook'a

Problem spełnialności wyrażeń logicznych jest NP-zupełny.

Problemami NP-zupełnymi są:

- Problem podziału,
- Problem komiwojażera,
- Problem cyklu Hamiltona.

Twierdzenie

Problem plecakowy jest NP-zupełny.

Cel:

Udowodnić NP-zupełność problemu plecakowego poprzez wielomianową transformację problemu podziału, który jest NP-zupełny, do plecakowego.

Problem podziału zbioru

Dane:

- $C = \{c_1, \dots, c_i, \dots, c_k\}$ - zbiór k elementów,
- Rozmiar $s(c_i) > 0$ elementu c_i, gdzie $s(c_i) \in N_+$, $N_+ = \{1,2,\dots\}$,
- $B \in N_+$,
- $\sum_{i=1}^{k} s(c_i) = 2B$.

Pytanie:

Czy istnieje podzbiór $C' \subset C$ taki, że

$$\sum_{c_i \in C'} s(c_i) = B \;?$$

Problem plecakowy - wersja decyzyjna

Dane:

Skończony zbiór elementów $A = \{a_1, a_2, \dots, a_n\}$.

Rozmiar $s(a_i) > 0$ i waga $w(a_i) > 0$ elementu a_i .

Pojemność plecaka $b > 0$ i stała $y > 0$.

Zadanie:

Czy istnieje podzbiór $A' \subset A$ taki, że:

$$\sum_{a_i \in A'} s(a_i) \leq b$$

$$\sum_{a_i \in A'} w(a_i) \geq y \quad ?$$

Dowód, że Problem plecakowy $\pi_1 \in NP$

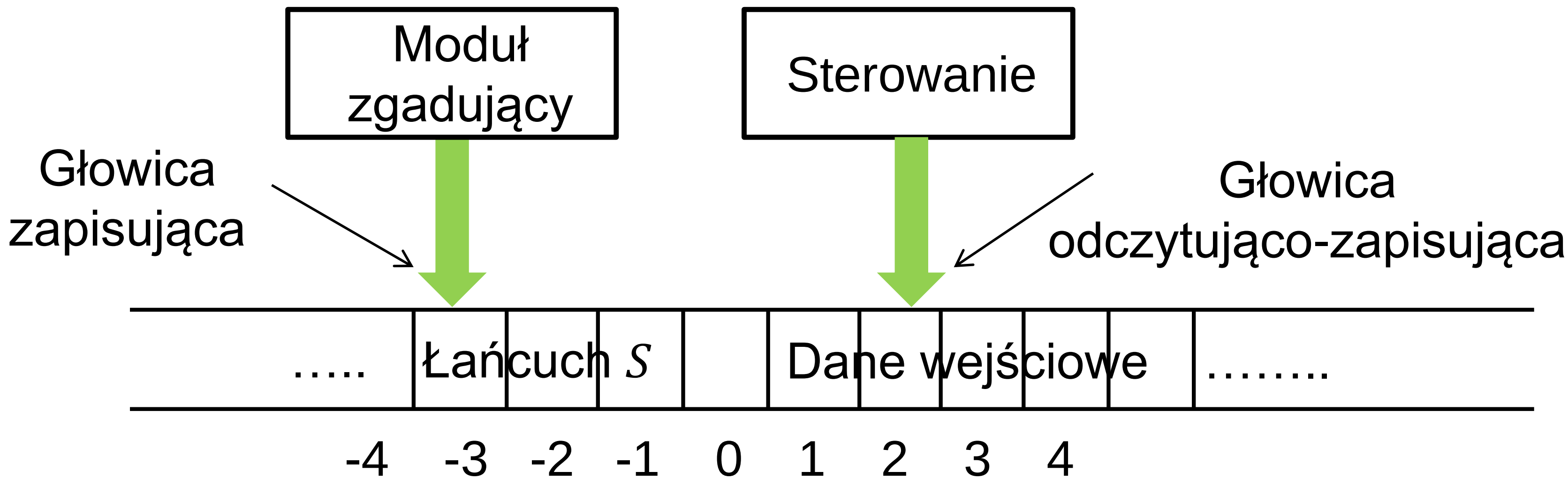
Aby rozwiązać instancję (konkretny problem) $I \in \pi_1$, NDMT musi wygenerować podzbiór $A' \subset A$ i sprawdzić w co najwyżej wielomianowym czasie, czy odpowiedź dla tego problemu brzmi „tak". Należy sprawdzić nierówności

$$\sum_{a_i \in A'} s(a_i) \leq b$$

$$\sum_{a_i \in A'} w(a_i) \geq y$$

Dowód, że Problem plecakowy $\pi_1 \in NP$

NDMT dla Problemu plecakowego



Łańcuch S - liczba binarna, której i-ta pozycja wskazuje czy i-ty element zbioru A należy do wygenerowanego rozwiązania A'

Dane wejściowe

$$n \sqcup s(a_1) \sqcup \ldots \sqcup s(a_n) \sqcup$$
$$w(a_1) \sqcup \ldots \sqcup w(a_n) \sqcup b \sqcup y$$

Dowód, że Problem plecakowy $\boldsymbol{\pi_1} \in \boldsymbol{NP}$

$n \sqcup s(a_1) \sqcup \ldots \sqcup s(a_n) \sqcup w(a_1) \sqcup \ldots \sqcup w(a_n) \sqcup b \sqcup y$

$$N(I) = \begin{array}{c} \lceil log_2 n \rceil + 1 + \sum_{i=1}^{n} (\lceil log_2 s(a_i) \rceil + 1) \\ + \sum_{i=1}^{n} (\lceil log_2 w(a_i) \rceil + 1) + \lceil (log_2 b \rceil + 1 + \lceil (log_2 y \rceil \end{array}$$

$\lceil x \rceil$ - najmniejsza liczba całkowita nie mniejsza niż x

$$N(I) < (2n+3)(\lceil log_2 \max\{n, \left\{s(a_i): \{\overline{1,n}\}\right\}, \{w(a_i): \{\overline{1,n}\}\}, b, y\} \rceil + 1)$$

1. Zgadnięcie rozwiązania to wygenerowanie ciągu n bitów.
2. Do sprawdzania nierówności: $\sum_{a_i \in A'} s(a_i) \leq b$, $\sum_{a_i \in A'} w(a_i) \geq y$ wystarcza $2|A'| - 2 \leq 2n$ operacji dodawania i 2 operacje porównania.

Dowód, że Problem plecakowy $\boldsymbol{\pi_1} \in \boldsymbol{NP}$

Do sprawdzania nierówności: $\sum_{a_i \in A'} s(a_i) \leq b$, $\sum_{a_i \in A'} w(a_i) \geq y$
wystarcza $2|A'| - 2 \leq 2n$ operacji dodawania i 2 operacje porównania.

Operacje porównania i dodawania dwóch liczb b_1 i b_2 DTM może wykonać w czasie wielomianowym zależnym od $\lceil (log_2 b_1 \rceil$ i $\lceil (log_2 b_2 \rceil$.

Zatem złożoność weryfikacji odgadniętego rozwiązania jest ograniczona od góry przez wielomian $p(N(I))$ czyli $\boldsymbol{\pi_1} \in \boldsymbol{NP}$.

Dowód, że Problem podziału $\boldsymbol{\pi_2 \propto \pi_1}$

Problem podziału π_2 jest NP-zupełny

Dla instancji $I_2 \epsilon D_{\pi_2}$ konstruujemy instancję $I_1 \epsilon D_{\pi_1}$ taką, że:

$$n = k,$$

wzajemnie jednoznaczne przyporządkowanie $g(c_i) = a_i$,

$$s(a_i) = s(c_i) \text{ dla } i \epsilon \{\overline{1,n}\},$$

$$w(a_i) = s(c_i) \text{ dla } i \epsilon \{\overline{1,n}\},$$

$$b = y = B.$$

Transformacją wielomianową problemu $\boldsymbol{\pi_2}$ do problemu $\boldsymbol{\pi_1}$ ($\pi_2 \propto \pi_1$) jest funkcja $f: D_{\pi_2} \to D_{\pi_1}$ spełniająca warunki:

1. Dla każdej instancji $I_2 \epsilon D_{\pi_2}$ odpowiedź brzmi „tak”, wtedy i tylko wtedy, gdy dla instancji $f(I_2)$ odpowiedź również jest „tak”,

2. Czas obliczenia funkcji f przez DMT dla każdej instancji $I_2 \epsilon D_{\pi_2}$ jest ograniczony od góry przez wielomian od $N(I_2)$.

Dowód, że Problem podziału $\boldsymbol{\pi_2 \propto \pi_1}$

1. Dowód, że dla każdej instancji $I_2 \epsilon D_{\pi_2}$ odpowiedź brzmi „tak", wtedy i tylko wtedy, gdy dla instancji $I_1 \in \pi_1$ odpowiedź również jest „tak".

Niech odpowiedź dla $I_1 \in \pi_1$ brzmi „tak". Zatem istnieje $A' \subset A$ taki, że: $\sum_{a_i \in A'} s(a_i) \leq b$, $\sum_{a_i \in A'} w(a_i) \geq y$. Ponieważ $s(a_i) = w(a_i)$=$s(c_i)$ dla $i \epsilon \{\overline{1,n}\}$ oraz $b = y = B$, a więc dla zbioru $C' = \{c_i : c_i =$

Dowód, że Problem podziału $\boldsymbol{\pi_2 \propto \pi_1}$

1. Dowód, że dla każdej instancji $I_2 \epsilon D_{\pi_2}$ odpowiedź brzmi „tak”, wtedy i tylko wtedy, gdy dla instancji $I_1 \in \pi_1$ odpowiedź również jest „tak”.

Niech odpowiedź dla $I_2 \epsilon D_{\pi_2}$brzmi „tak”. Zatem istnieje $C' \subset C$ taki, że: $\sum_{c_i \in C'} \quad s(c_i) = B$. Ponieważ

$s(a_i) = w(a_i)$=$s(c_i)$ dla $i \epsilon \{\overline{1,n}\}$ oraz $b = y = B$, a więc dla zbioru $A' = \{a_i : a_i = s(c_i) \wedge c_i \epsilon C' \ \}$ prawdziwe jest

$$\sum_{a_i \in A'} s(\,a_i) = \sum_{a_i \in A'} w(a_i) = b = y$$

Zatem dla instancji $I_1 = f(I_2)$, $I_1 \in \pi_1$ odpowiedź brzmi „tak”.

Dowód, że Problem podziału $\boldsymbol{\pi_2 \propto \pi_1}$

2. Czas obliczenia funkcji f przez DMT dla każdej instancji $I_2 \epsilon D_{\pi_2}$ jest ograniczony od góry przez wielomian od $N(I_2)$.

Czas konstrukcji danych $I_1 \epsilon D_{\pi_1}$ jest ograniczony od góry przez wielomian od rozmiaru $I_2 \epsilon D_{\pi_2}$, ponieważ DMT musi przepisać $2n + 3$ liczb.

W celu wykazania NP-zupełności problemu π_1 należy:

1. Pokazać, że $\pi_1 \in NP$,
2. Wybrać odpowiedni NP-zupełny problem π_2,
3. Skonstruować transformację $f: D_{\pi_2} \to D_{\pi_1}$,
4. Pokazać, że f jest obliczana w czasie wielomianowym,
5. Pokazać, że:

   Jeśli odpowiedź dla $I_1 \epsilon D_{\pi_1}$ brzmi „tak", to

   dla instancji $I_2 \epsilon D_{\pi_2}$ odpowiedź brzmi „tak",

   (ew. dowód nie wprost).

6. Pokazać, że:

   Jeśli odpowiedź dla $I_2 \epsilon D_{\pi_2}$ brzmi „tak", to

   dla instancji $I_1 \epsilon D_{\pi_1}$ odpowiedź brzmi „tak"

Zwykle najtrudniejsze są punkty 2. i 6.

Problem optymalizacyjny jest NP-trudny, jeśli odpowiadający mu problem decyzyjny jest NP-zupełny.

NP-zupełny ang. NP-complete

NP-trudny ang. NP-hard

Aby wykazać NP-trudność problemu optymalizacyjnego należy udowodnić NP-zupełność jego wersji decyzyjnej.

Zasadnicze techniki dowodzenia NP-zupełności problemów decyzyjnych:

- Ograniczanie,
- Lokalna zamiana,
- Projektowanie części składowych.

Którą z technik zastosowano dowodząc NP-zupełności Problemu plecakowego poprzez przetransformowanie do niego Problemu podziału?