



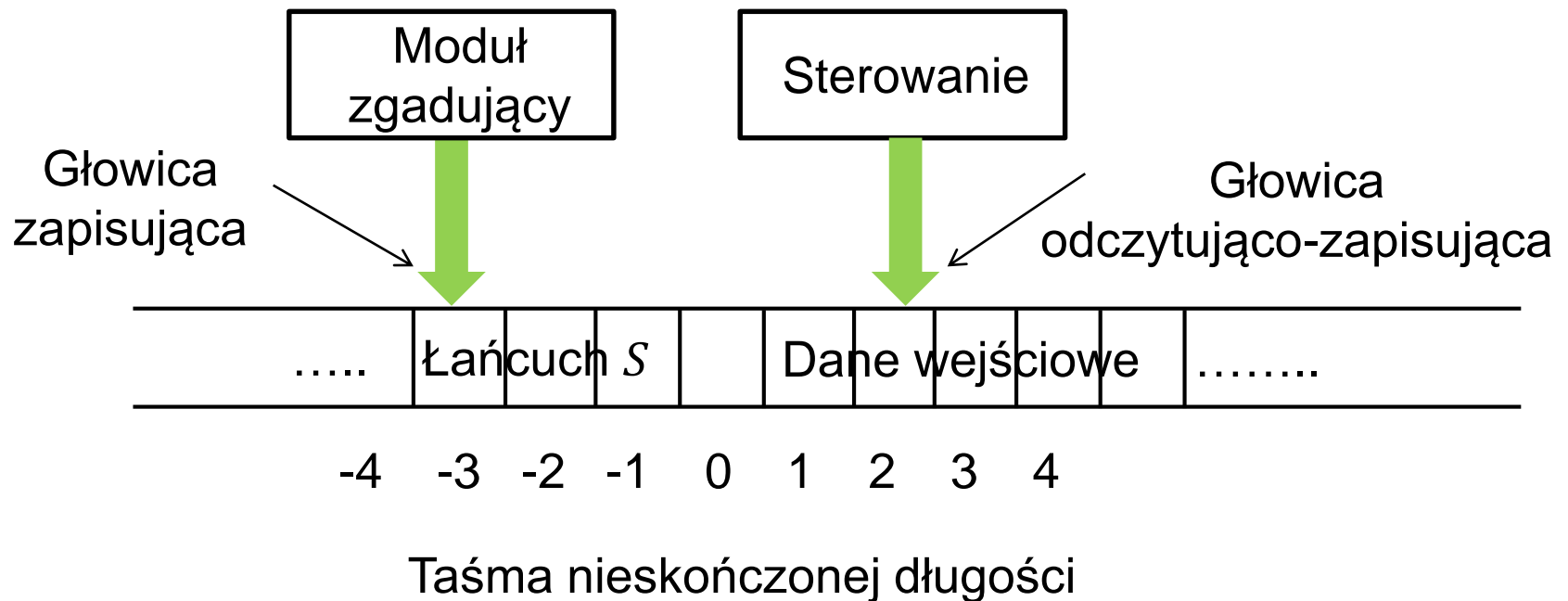
Politechnika Wroclawska

Struktury danych
i złożoność obliczeniowa
Wykłady 8., 9.

Prof. dr hab. inż. Jan Magott



Niedeterministyczna jednośmowa maszyna Turinga (NDMT)





Niedeterministyczna maszyna Turinga (NDMT)

Moduł zgadujący tylko zapisuje na taśmie odgadnięte rozwiązanie (np. podzbiór zbioru elementów w problemie podziału, kolejność odwiedzania wszystkich miast w problemie komiwojażera, podzbiór elementów do upakowania w plecaku problemu plecakowego).

Wykonanie programu składa się z wielu wykonań pary działań:

1. Zgadywania rozwiązania - generowania łańcucha S symboli,
2. Sprawdzania jaka jest odpowiedź na pytanie problemu decyzyjnego dla wygenerowanego rozwiązania.



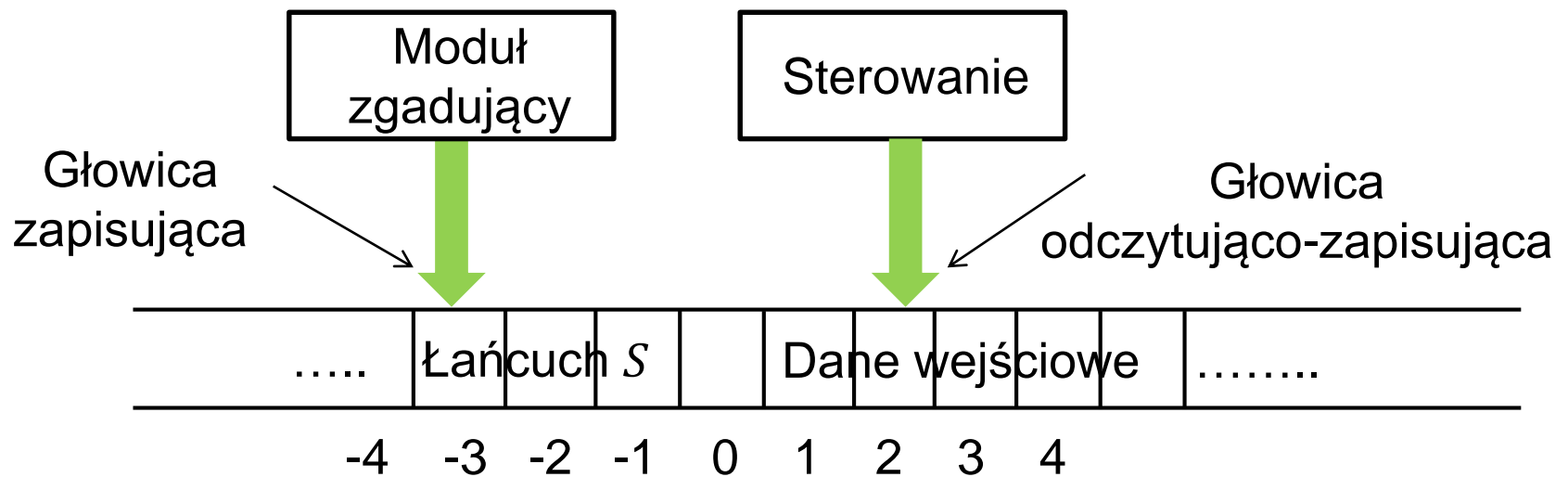
Niedeterministyczna jednotaśmowa maszyna Turinga (NDMT) składa się z DMT i modułu zgadującego (generującego).

Wykonanie programu NDMT dla łańcucha danych $x(I)$ instancji I przebiega następująco:

1. Moduł zgadujący zapisuje na taśmie łańcuch S symboli ze skończonego zbioru symboli taśmy,
2. NDMT sprawdza, tak jak wykonywany jest program przez DMT, czy wygenerowany łańcuch S spełnia warunki pytania instancji I .

Dla jednej instancji I może istnieć wiele łańcuchów S reprezentujących rozwiązania. ?

NDMT dla Problemu podziału zbioru



Łańcuch S - liczba binarna, której i -ta pozycja wskazuje czy i -ty element zbioru X należy do wygenerowanego rozwiązania X_1

Dane wejściowe

$$x_1 \sqcup x_2 \sqcup \dots \sqcup x_k \sqcup B$$



NDMT dla Problemu podziału zbioru

$2^k - 2$ - liczba generowanych łańcuchów reprezentujących zbiór X_1

Czy koszt sprawdzenia relacji

$$\sum_{x_i \in X_1} x_i = B$$

jest ograniczony od góry przez wielomian od k ?

Jaka byłaby złożoność generowania łańcuchów i sprawdzania, gdyby program był wyrażony w języku wysokiego poziomu.



NDMT rozwiązuje problem decyzyjny π , jeśli dla każdej instancji $I \in D_\pi$ 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 $I \in D_\pi$, 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 $O(p(N(I)))$ dla pewnego wielomianu p .



Twierdzenie

Jeśli jednotaśmowa NDMT rozwiązuje decyzyjny problem π w czasie wielomianowym, to istnieje wielomian p taki, że jednotaśmowa DMT rozwiązuje ten problem w czasie $O(2^{p(N(I))})$, gdzie $I \in D_\pi$ a $N(I)$ jest rozmiarem danych wejściowych instancji I .



Interpretacja Twierdzenia o relacji między NDMT a DMT

Działanie DMT

DMT powinna generować kolejno rozwiązania np. poprzez tworzenie kolejnych liczb binarnych k cyfrowych i sprawdzać jaka jest odpowiedź na pytanie.



Adekwatność NDMT jako modelu obliczeń

- Sekwencyjny dostęp do danych wejściowych i wyników pośrednich (ze względu na organizację taśmową).
- Nie istnieje rzeczywisty odpowiednik NDMT rozwiązujący problemy decyzyjne w prezentowany sposób.
- NDMT obrazuje zdolność weryfikacji pozytywnej odpowiedzi dla rozwiązania (wygenerowanego łańcucha S) instancji $I \in D_\pi$.



Klasę P tworzą wszystkie problemy decyzyjne, które w co najwyżej wielomianowym czasie może rozwiązać DMT.

Klasa 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 $P = NP$, 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. ?



- 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.



Problemy NP-zupełne



Transformacją wielomianową problemu π_2 do problemu π_1 ($\pi_2 \propto \pi_1$) jest funkcja $f: D_{\pi_2} \rightarrow D_{\pi_1}$ spełniająca warunki:

1. Dla każdej instancji $I_2 \in D_{\pi_2}$ 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 $I_2 \in D_{\pi_2}$ 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 $\pi_1 \in NP$, to $\pi_2 \in NP$.

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

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. $\pi_1 \in NP$,
2. Dla każdego innego problemu decyzyjnego $\pi_2 \in NP$ 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.



Do udowodnienia NP-zupełności problemu decyzyjnego π wystarczy:

1. Dowieść, że $\pi \in NP$,
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.



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 (wartość) $w(a_i) > 0$ 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 $A' \subset A$ w co najwyżej wielomianowym czasie i sprawdzić w co najwyżej wielomianowym czasie, czy odpowiedź dla tego problemu brzmi „tak”. Należy sprawdzić nierówności

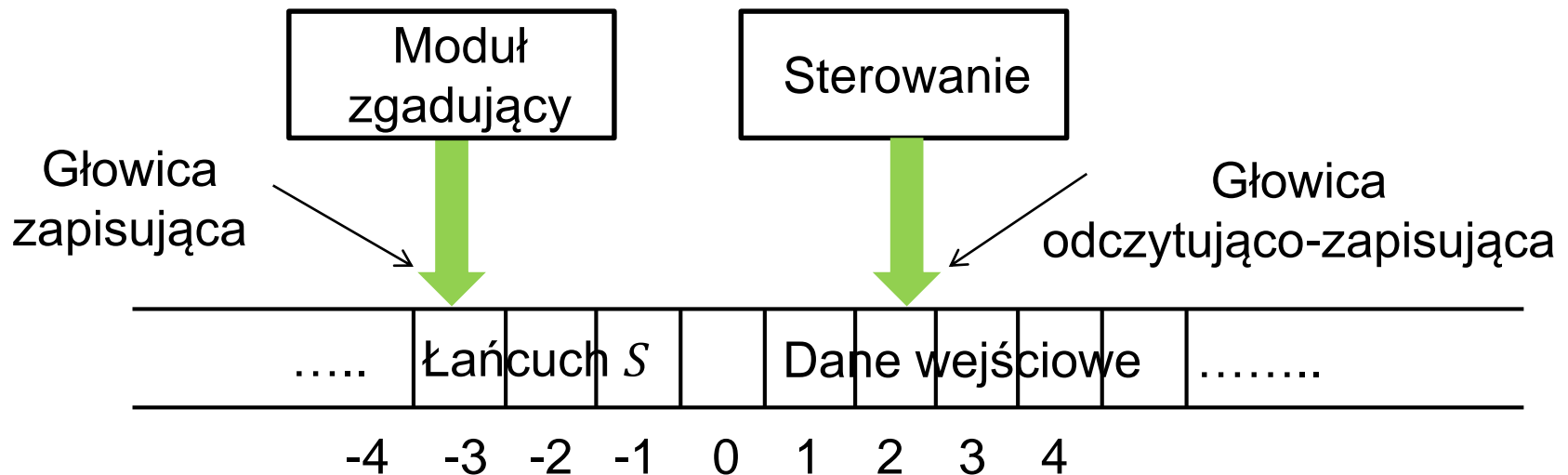
$$\sum_{a_i \in A'} s(a_i) \leq b$$

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



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

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



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

Dokładnie

$$\lceil \log_2(n + 1) \rceil$$

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

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

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

$$N(I) \approx$$

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



Model *RAM* (ang. random access machine)

- Jeden procesor,
- Operacje elementarne (zapisania, dodawania, odejmowania, porównania dwu liczb, itp.) wymagają jednego kroku czasowego,
- Taśma wejściowa z głowicą odczytującą,
- Taśma wyjściowa z głowicą zapisującą,
- Pamięć danych jest zbiorem rejestrów z wyróżnionym rejestrem - akumulatorem, w którym wykonywane są obliczenia,
- Dostęp do pamięci wymaga jednego kroku,
- Możliwość adresowania pośredniego,
- Program, który nie jest przechowywany w pamięci jest sekwencją rozkazów (nie podlega automodyfikacji),
- Licznik rozkazów.



Model RAM c.d.

- Każda komórka taśm: wejściowej i wyjściowej oraz każdy rejestr może zawierać dowolną liczbę całkowitą.
- Po odczycie z komórki taśmy wejściowej (zapisie w komórce taśmy wyjściowej), głowica taśmy jest przesuwana o jedną pozycję w prawo. Treści wpisanej na taśmie wyjściowej nie można zmienić.
- Rozkazy (zbiór rozkazów nie jest precyzyjnie zdefiniowany, ale nie może zawierać instrukcji niespotykanych w rzeczywistych komputerach):
 - Arytmetyczne (+, -, x, /),
 - Wejścia-wyjścia,
 - Rozgałęzienia przepływu sterowania.



Twierdzenie

Modele procesu obliczeń:

- *Jednotaśmowa maszyna Turinga,*
- *Wielotaśmowa maszyna Turinga,*
- *Maszyna RAM*

są równoważne w tym sensie, że jeśli dany problem jest rozwiązywany przez jeden model w czasie ograniczonym od góry przez wielomian zależny od rozmiarów problemu, to przy założeniu logarytmicznego kryterium kosztów jest on również rozwiązywany przez każdy inny model w czasie ograniczonym od góry przez wielomian zależny od jego rozmiarów.

(Można czerpać z intuicji programowania w assemblerze.)



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

Generacja rozwiązania

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

Złożoność liniowa



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

Weryfikacja rozwiązania

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 $\lceil \log_2 b_2 \rceil$, natomiast RAM przy kryterium logarytmicznym i jednorodnym.

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



Dowód, że Problem podziału $\pi_2 \propto \pi_1$

Problem podziału π_2 jest NP-zupełny

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

$$n = k,$$

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

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

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

$$b = y = B.$$



Transformacją wielomianową problemu π_2 do problemu π_1 ($\pi_2 \propto \pi_1$) jest funkcja $f: D_{\pi_2} \rightarrow D_{\pi_1}$ spełniająca warunki:

1. Dla każdej instancji $I_2 \in D_{\pi_2}$ 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 $I_2 \in D_{\pi_2}$ jest ograniczony od góry przez wielomian od $N(I_2)$.



Dowód, że Problem podziału $\pi_2 \propto \pi_1$

1. Dowód, że dla każdej instancji $I_2 \in D_{\pi_2}$ 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 $i \in \overline{\{1, n\}}$ oraz $b = y = B$, a więc dla zbioru $C' = \{c_i : c_i = g^{-1}(a_i) \wedge a_i \in A'\}$ prawdziwe jest $\sum_{c_i \in C'} s(c_i) = B$.

Zatem dla instancji $I_2 \in D_{\pi_2}$ odpowiedź brzmi „tak”.



Dowód, że Problem podziału $\pi_2 \propto \pi_1$

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

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

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

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

Zatem dla instancji $I_1 \in \pi_1$ odpowiedź brzmi „tak”.



Dowód, że Problem podziału $\pi_2 \propto \pi_1$

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

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



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 – zupełne - klasa problemów prawie na pewno nie rozwiązywalnych w czasie wielomianowym.

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



Do udowodnienia NP-zupełności problemu decyzyjnego π wystarczy:

1. Dowieść, że $\pi \in NP$,
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.



Problemami NP-zupełnymi są:

- Problem podziału,
- Problem komiwojażera w wersji decyzyjnej,
- Problem cyklu Hamiltona.



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 $f: D_{\pi_2} \rightarrow D_{\pi_1}$,
4. Pokazać, że f jest obliczana w czasie wielomianowym,
5. Pokazać, że $\pi_2 \Rightarrow \pi_1$,
6. Pokazać, że $\pi_1 \Rightarrow \pi_2$ lub $\neg\pi_2 \Rightarrow \neg\pi_1$.

Zwykle najtrudniejsze są punkty 2. i 6.



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?



Problem cyklu Hamiltona

Dane:

Graf nieskierowany $G = \langle V, E \rangle$, $n = |V|$

Pytanie:

Czy G zawiera cykl Hamiltona, tzn. czy istnieje takie uporządkowanie wierzchołków grafu $\langle v_{i[1]}, v_{i[2]}, \dots, v_{i[n]} \rangle$, że:

$$\{v_{i[j]}, v_{i[j+1]}\} \in E \text{ dla } j \in \overline{1, n-1}$$

oraz

$$\{v_{i[n]}, v_{i[1]}\} \in E ?$$



Problem ścieżki Hamiltona

Dane:

Graf nieskierowany $G = \langle V, E \rangle$, $n = |V|$

Pytanie:

Czy G zawiera ścieżkę Hamiltona, tzn. czy istnieje takie uporządkowanie wierzchołków grafu $\langle v_{i[1]}, v_{i[2]}, \dots, v_{i[n]} \rangle$, że $\{v_{i[j]}, v_{i[j+1]}\} \in E$ dla $j \in \overline{\{1, n-1\}}$?



Twierdzenie

Problemy cyklu Hamiltona i ścieżki Hamiltona są NP-zupełne.



Problem najdłuższej ścieżki

Graf nieskierowany $G = \langle V, E \rangle$, liczba naturalna $K < |V|$.

Pytanie:

Czy G zawiera ścieżkę prostą (ścieżkę przechodzącą przez każdy z jej wierzchołków dokładnie jeden raz) zawierającą K lub więcej krawędzi?



Zasadnicze techniki dowodzenia NP-zupełności problemów decyzyjnych:

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



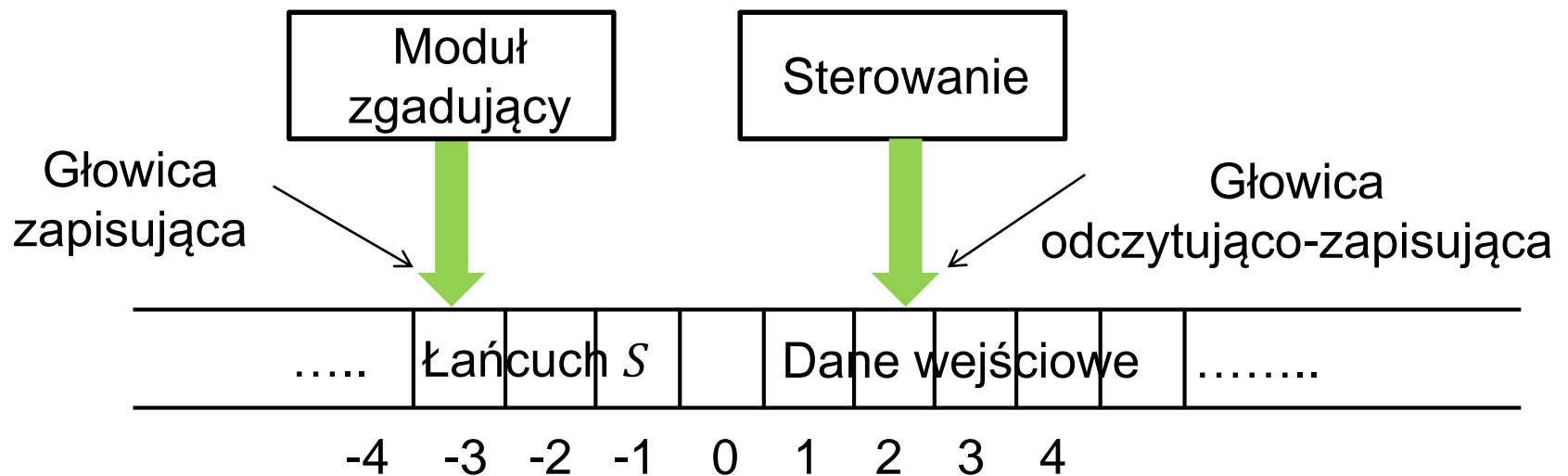
Twierdzenie

Problem najdłuższej ścieżki jest NP-zupełny.

Dowód wykonany techniką ograniczania.

Jako znany problem NP-zupełny przyjmujemy Problem ścieżki Hamiltona.

NDMT dla *Problemu najdłuższej ścieżki*



Łańcuch S - ciąg $K+1$ indeksów wierzchołków, którego i -ta pozycja wskazuje na i -ty wierzchołek wygenerowanej sekwencji.

Dane wejściowe

$$|N| \sqcup v_0, v_j \sqcup v_k, v_l \sqcup \dots \sqcup v_p, v_r \sqcup$$



Twierdzenie

Problem najdłuższej ścieżki π_1 jest NP-zupełny.

Dowód, że $\pi_1 \in NP$

Podczas rozwiązywania problemu π_1 , moduł zgadujący NDMT generuje sekwencję $(K + 1)$ wierzchołków, tzn. K krawędzi.

W jaki sposób? Jaka jest złożoność obliczeniowa?

Deterministyczna część NDMT sprawdza czy dla każdych dwu kolejnych wierzchołków $v_{i[j]}, v_{i[j+1]}$ jest $\{v_{i[j]}, v_{i[j+1]}\} \in E$ oraz czy wierzchołki powtarzają się.

Ponieważ te czynności można wykonać w czasie wielomianowym, a więc $\pi_1 \in NP$.



Twierdzenie

Problem najdłuższej ścieżki jest NP-zupełny.

Dowód, że $\pi_2 \propto \pi_1$

π_1 - Problem najdłuższej ścieżki,

π_2 - Problem ścieżki Hamiltona.



Problem ścieżki Hamiltona

Dane:

Graf nieskierowany $G = \langle V, E \rangle$, $n = |V|$

Pytanie:

Czy G zawiera ścieżkę Hamiltona, tzn. czy istnieje takie uporządkowanie wierzchołków grafu $\langle v_{i[1]}, v_{i[2]}, \dots, v_{i[n]} \rangle$, że $\{v_{i[j]}, v_{i[j+1]}\} \in E$ dla $j \in \overline{\{1, n-1\}}$?

Problem najdłuższej ścieżki

Graf nieskierowany $G = \langle V, E \rangle$, liczba naturalna $K < |V|$.

Pytanie:

Czy G zawiera ścieżkę prostą (ścieżkę przechodzącą przez każdy z jej wierzchołków dokładnie jeden raz) zawierającą K lub więcej krawędzi?

Czy jeden z powyższych problemów jest ogólniejszy od drugiego?

Jak dobrać K ?



Twierdzenie

Problem najdłuższej ścieżki jest NP-zupełny.

Dowód, że $\pi_2 \propto \pi_1$

π_1 - Problem najdłuższej ścieżki,

π_2 - Problem ścieżki Hamiltona.

W celu wykazania powyższego należy π_1 ograniczyć tylko do tych instancji $I_1 \in D_{\pi_1}$, takich, że $K = |V| - 1$. Tak ograniczony π_1 jest problemem π_2 .



Graf pełny

W grafie pełnym każda para wierzchołków jest połączona.

Izomorfizm grafów

Dane:

Grafy $G_1 = \langle V_1, E_1 \rangle$, $G_2 = \langle V_2, E_2 \rangle$.

Pytanie:

Czy istnieje funkcja jeden na jeden $f: V_1 \rightarrow V_2$ taka, że $\{u, v\} \in E_1$ wtedy i tylko wtedy gdy $\{f(u), f(v)\} \in E_2$?



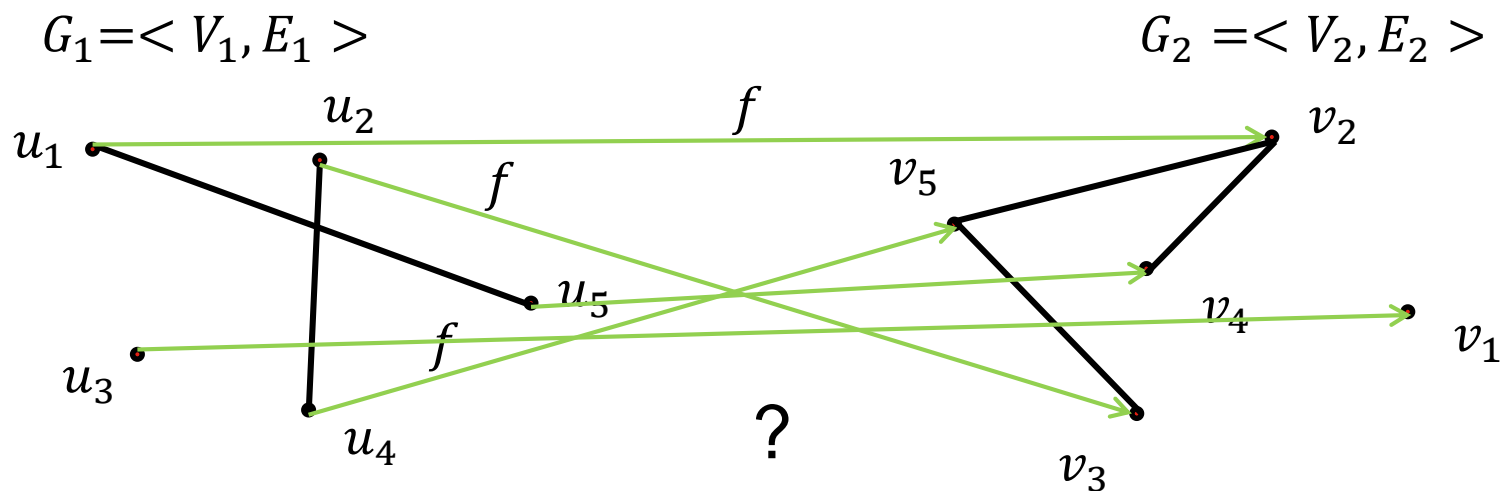
Izomorfizm grafów

Dane:

Grafy $G_1 = \langle V_1, E_1 \rangle$, $G_2 = \langle V_2, E_2 \rangle$.

Pytanie:

Czy istnieje funkcja jeden na jeden $f: V_1 \rightarrow V_2$ taka, że $\{u, v\} \in E_1$ wtedy i tylko wtedy gdy $\{f(u), f(v)\} \in E_2$?





Testy izomorfizmu grafów

?



Testy izomorfizmu grafów

- Liczba wierzchołków
- Liczba krawędzi
- Stopnie wierzchołków



Niech grafy $G_1 = \langle V_1, E_1 \rangle$, $G_2 = \langle V_2, E_2 \rangle$
będą takie, że $|V_1| = |V_2| = n$.

Ile jest funkcji jeden na jeden $f: V_1 \rightarrow V_2$?



Niech grafy $G_1 = \langle V_1, E_1 \rangle$, $G_2 = \langle V_2, E_2 \rangle$
będą takie, że $|V_1| = |V_2| = n$.

Funkcji jeden na jeden $f: V_1 \rightarrow V_2$ jest $n!$.



Problem kliki

Dane:

Graf nieskierowany $G = \langle V, E \rangle$, liczba naturalna $k \leq |V|$.

Pytanie:

Czy graf zawiera klikę o rozmiarze k lub większym, tzn. podgraf pełny zawierający $|V'| \geq k$ wierzchołków.

Problem największego wspólnego podgrafu

Dane:

Grafy nieskierowane $G_1 = \langle V_1, E_1 \rangle$ i $G_2 = \langle V_2, E_2 \rangle$, liczba naturalna K .

Pytanie:

Czy istnieją podzbiory $E_1' \subset E_1$ i $E_2' \subset E_2$ takie, że $|E_1'| = |E_2'| \geq K$ i podgrafy $G_1' = \langle V_1, E_1' \rangle$ i $G_2' = \langle V_2, E_2' \rangle$ są izomorficzne?



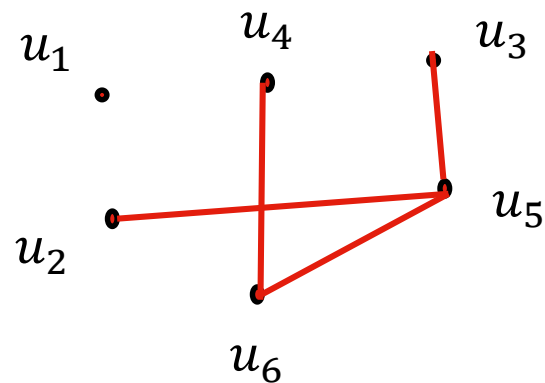
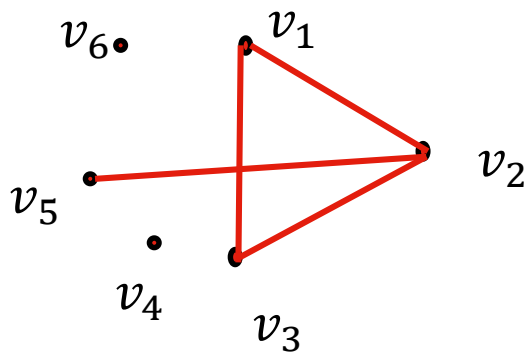
Problem największego wspólnego podgrafu

Dane:

Grafy nieskierowane $G_1 = \langle V_1, E_1 \rangle$ i $G_2 = \langle V_2, E_2 \rangle$, liczba naturalna K .

Pytanie:

Czy istnieją podzbiory $E_1' \subset E_1$ i $E_2' \subset E_2$ takie, że $|E_1'| = |E_2'| \geq K$ i podgrafy $G_1' = \langle V_1, E_1' \rangle$ i $G_2' = \langle V_2, E_2' \rangle$ są izomorficzne?





Twierdzenie

Problem największego wspólnego podgrafu π_1 jest NP-zupełny.

Dowód, że $\pi_1 \in NP$



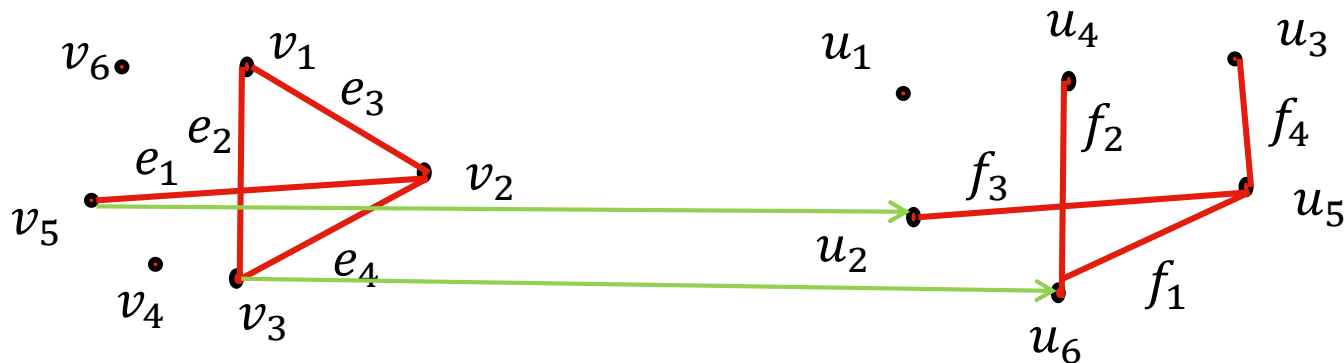
Dowód, że *Problem największego wspólnego podgrafu* $\pi_1 \in NP$

Dane:

Grafy nieskierowane $G_1 = \langle V_1, E_1 \rangle$ i $G_2 = \langle V_2, E_2 \rangle$, liczba naturalna K .

Pytanie:

Czy istnieją podzbiory $E_1' \subset E_1$ i $E_2' \subset E_2$ takie, że $|E_1'| = |E_2'| \geq K$ i podgrafy $G_1' = \langle V_1, E_1' \rangle$ i $G_2' = \langle V_2, E_2' \rangle$ są izomorficzne?



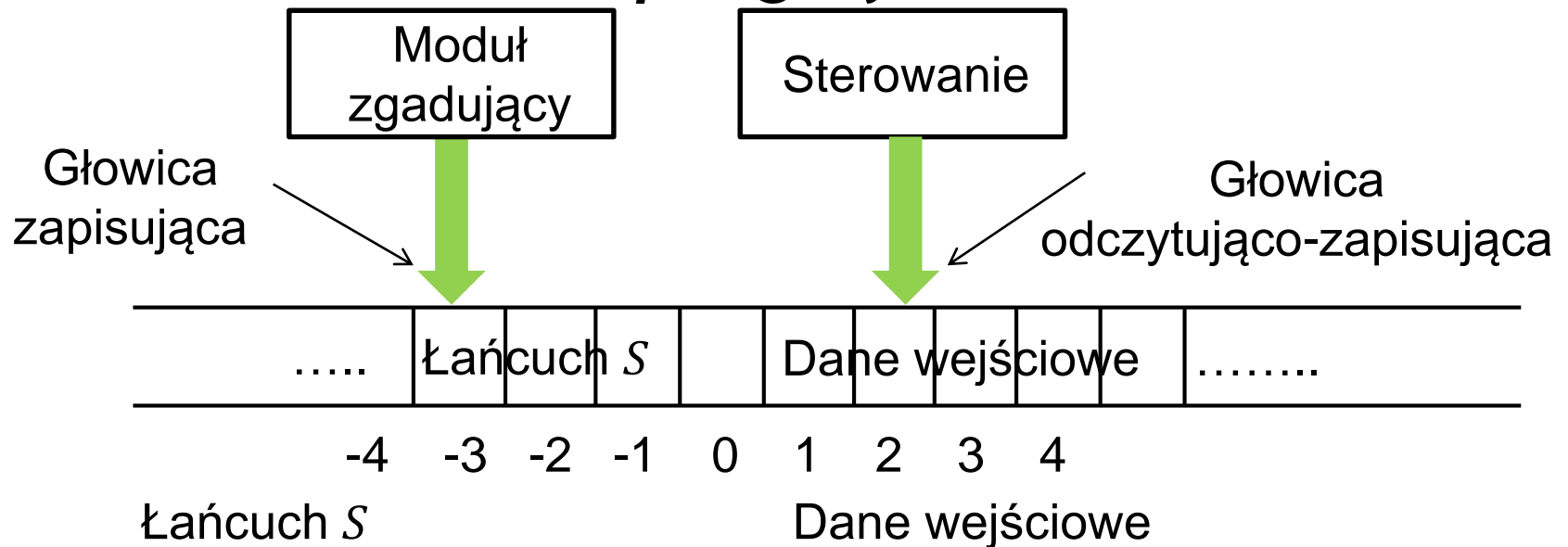
$$E_1' = \{v_2 v_5, v_1 v_3, v_2 v_1, v_2 v_3\}$$

$$E_2' = \{u_6 u_4, u_5 u_2, u_6 u_5, u_5 u_3\}$$

$$g: V_1 \rightarrow V_2: v_5, u_2 \sqcup v_3, u_6 \sqcup \dots \sqcup$$



NDMT dla *Problemu największego wspólnego podgrafu*



Pary zbiorów krawędzi E'_1, E'_2
 Pary reprezentujące odwzorowanie
 $g: V_1 \rightarrow V_2: v_0, u_j \sqcup v_k, u_l \sqcup \dots \sqcup$

Liczby wierzchołków obu grafów z kraw.
 $|V_1| \sqcup v_0, v_j \sqcup v_k, v_l \sqcup \dots \sqcup v_p, v_r \sqcup |V_2| \sqcup$
 $u_0, u_k \sqcup u_k, u_m \sqcup \dots \sqcup u_r, u_s \sqcup$



Generacja rozwiązania

Wybór krawędzi podgrafu E_1' z K krawędziami

Wybór krawędzi podgrafu E_2' z K krawędziami

Wybór funkcji g

Weryfikacja czy podgrafy E_1' i E_2' są izomorficzne?

Dla każdej krawędzi $\{v_k, v_l\}$ ze zbioru E_1' wykonać

wyznaczyć zbiór $\{g(v_k), g(v_l)\}$;

sprawdzić czy $\{g(v_k), g(v_l)\}$ jest krawędzią w E_2' ;



Dowód, że *Problem największego wspólnego podgrafu* $\pi_1 \in NP$

NDMT musi sprawdzić, czy:

Czy istnieją podzbiory $E_1' \subset E_1$ i $E_2' \subset E_2$ takie, że $|E_1'| = |E_2'| \geq K$
i podgrafy $G_1' = \langle V_1, E_1' \rangle$ i $G_2' = \langle V_2, E_2' \rangle$ są izomorficzne?

Generacja jednego rozwiązania i sprawdzenie warunków mogą być
wykonane w czasie ograniczonym przez wielomian od rozmiaru
problemu.

Zatem $\pi_1 \in NP$.



Problem klikli

Dane:

Graf nieskierowany $G = \langle V, E \rangle$, liczba naturalna $k \leq |V|$.

Pytanie:

Czy graf zawiera klikę o rozmiarze k lub większym, tzn. podgraf pełny zawierający $|V'| \geq k$ wierzchołków.

Problem największego wspólnego podgrafu

Dane:

Grafy nieskierowane $G_1 = \langle V_1, E_1 \rangle$ i $G_2 = \langle V_2, E_2 \rangle$, liczba naturalna K .

Pytanie:

Czy istnieją podzbiory $E_1' \subset E_1$ i $E_2' \subset E_2$ takie, że $|E_1'| = |E_2'| \geq K$ i podgrafy $G_1' = \langle V_1, E_1' \rangle$ i $G_2' = \langle V_2, E_2' \rangle$ są izomorficzne?



Dowód, że $\pi_2 \propto \pi_1$

Problem kliki π_2

Problem największego wspólnego podgrafu π_1

Punktem wyjścia jest twierdzenie

Problem kliki jest NP-zupełny.

Niech w Problemie kliki dane będzie $k \leq |V|$.

W Problemie największego wspólnego podgrafu, jako G_1 przyjmujemy graf, dla którego należy rozwiązać problem kliki, natomiast

jako G_2 przyjmijmy graf zawierający graf pełny o k wierzchołkach.

Jego pozostałe $|V_2| - k$ wierzchołki mają stopnie równe 0. W G_2 jest krawędzi w liczbie $\sum_{i=k}^2 (i - 1) = (k - 1) \cdot k / 2$



Dowód, że $\pi_2 \propto \pi_1$

Problem kliki π_2

Problem największego wspólnego podgrafu π_1

Ograniczamy π_1 do postaci, gdy graf G_2 jest grafem jak poprzednio określony, a $K = |E_2| = (k - 1) \cdot k/2$. Tak zdefiniowany ***Problem największego podgrafu*** polega na szukaniu klikki o k wierzchołkach w grafie G_1 . Zatem tak ograniczony ***Problem największego wspólnego podgrafu*** daje rozwiązanie dla ***Problemu klikki***.