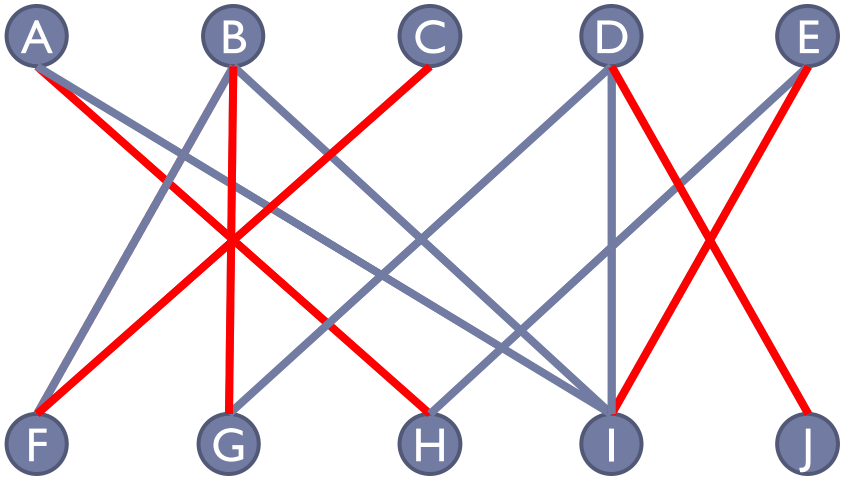


Maksymalne skojarzenia
Grafy dwudzielne

dr Anna Beata Kwiatkowska, UMK Toruń

Skojarzenie i problem maksymalnego skojarzenia

Definicja 11.1

Podzbiór krawędzi M w nieskierowanym grafie G=(V,E) nazywamy skojarzeniem, jeśli żadne dwie krawędzie w M nie mają wspólnego wierzchołka. Pojedyncza krawędź grafu również tworzy skojarzenie i jest to wtedy najmniej liczne skojarzenie.

Problem maksymalnego (najliczniejkszego) skojarzenia (ang. *cardinality matching problem*)

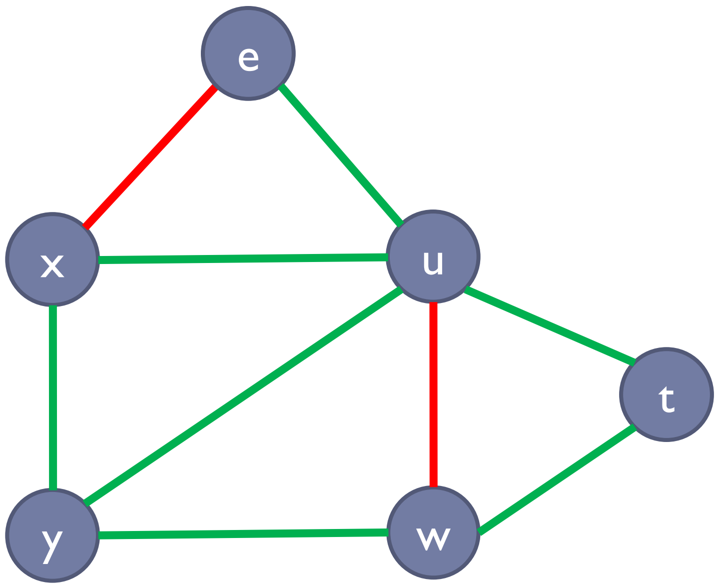
W danym grafie G wyznaczyć maksymalne skojarzenie, tzn. skojarzenie z największą liczbą krawędzi. (Claud i Berge, 1957)

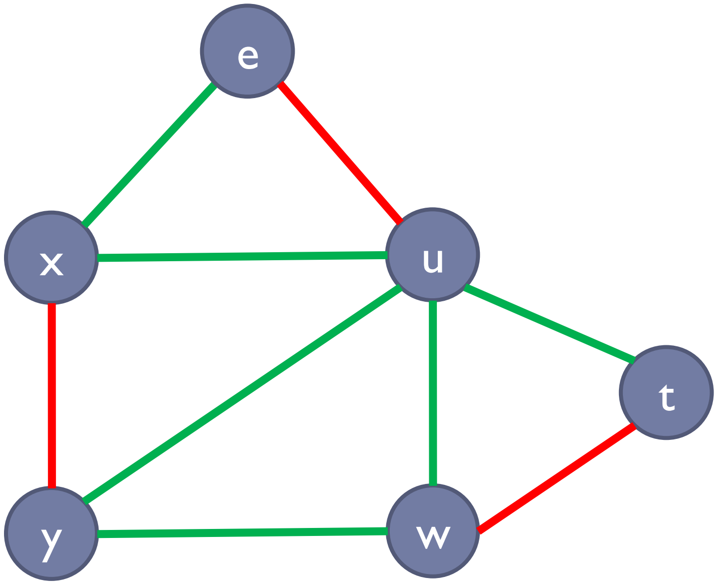
Przykłady zastosowań:

1. Obsada stanowisk kandydatami o określonych kwalifikacjach.
2. Wyznaczanie przydziału pracowników do maszyn, maksymalizującego sumaryczną wydajność (problem obciążonego skojarzenia).
3. Problem optymalizacji kodu komputerowego.

Skojarzenia

Przykład: Graf ze wskazanym dowolnym i maksymalnym skojarzeniem.





Droga naprzemienna i powiększająca

Rozważamy skojarzenie M w grafie G=(V,E).

Definicja 11.2

Drogę $P = (v_1, v_2, \dots, v_k)$ nazywamy drogą naprzemienną względem skojarzenia M, jeśli jej krawędzie na przemian należą i nie należą do M.

Definicja 11.3

Drogę naprzemienną zaczynającą się i kończącą w wierzchołku nieskojarzonym nazywamy drogą powiększającą.

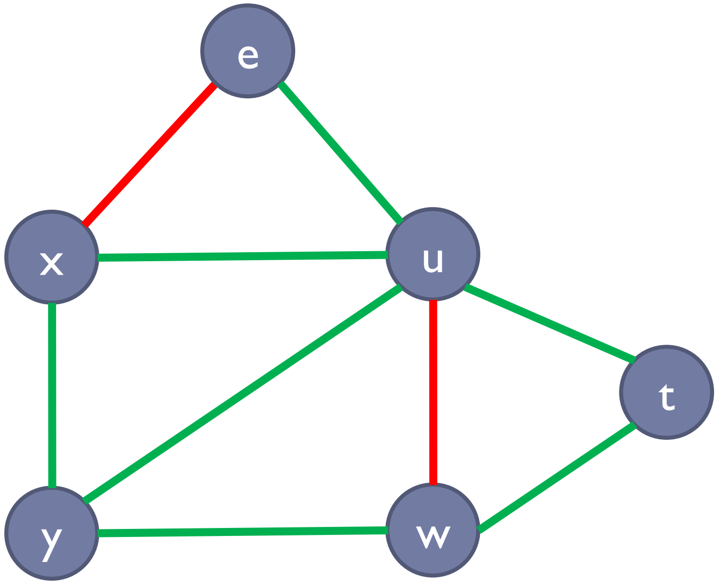
Przykład: W przedstawionym grafie

P= (t, w, u, e, x, y) jest drogą powiększającą.

Wierzchołki x, e, u, w są skojarzone,

pozostałe wierzchołki nie są skojarzone

(są wolne).



Algorytm szukania maksymalnego skojarzenia

Twierdzenie 11.1 Berge'a

Skojarzenie M w grafie G jest maksymalne wtedy i tylko wtedy, gdy G nie zawiera drogi powiększającej względem M.

Dowód

← Niech P będzie drogą powiększającą względem maksymalnego skojarzenia M w grafie G. Zamieniając rolami krawędzie w drodze P, tzn. usuwając ze skojarzenia M te krawędzie, które należą do drogi P, a wstawiając do skojarzenia M te krawędzie, które nie należą do P, otrzymujemy skojarzenie M', które ma o jedną krawędź więcej. Przeczy to założeniu, że M było maksymalnym skojarzeniem w grafie G.

→ dowód do własnego rozważenia.

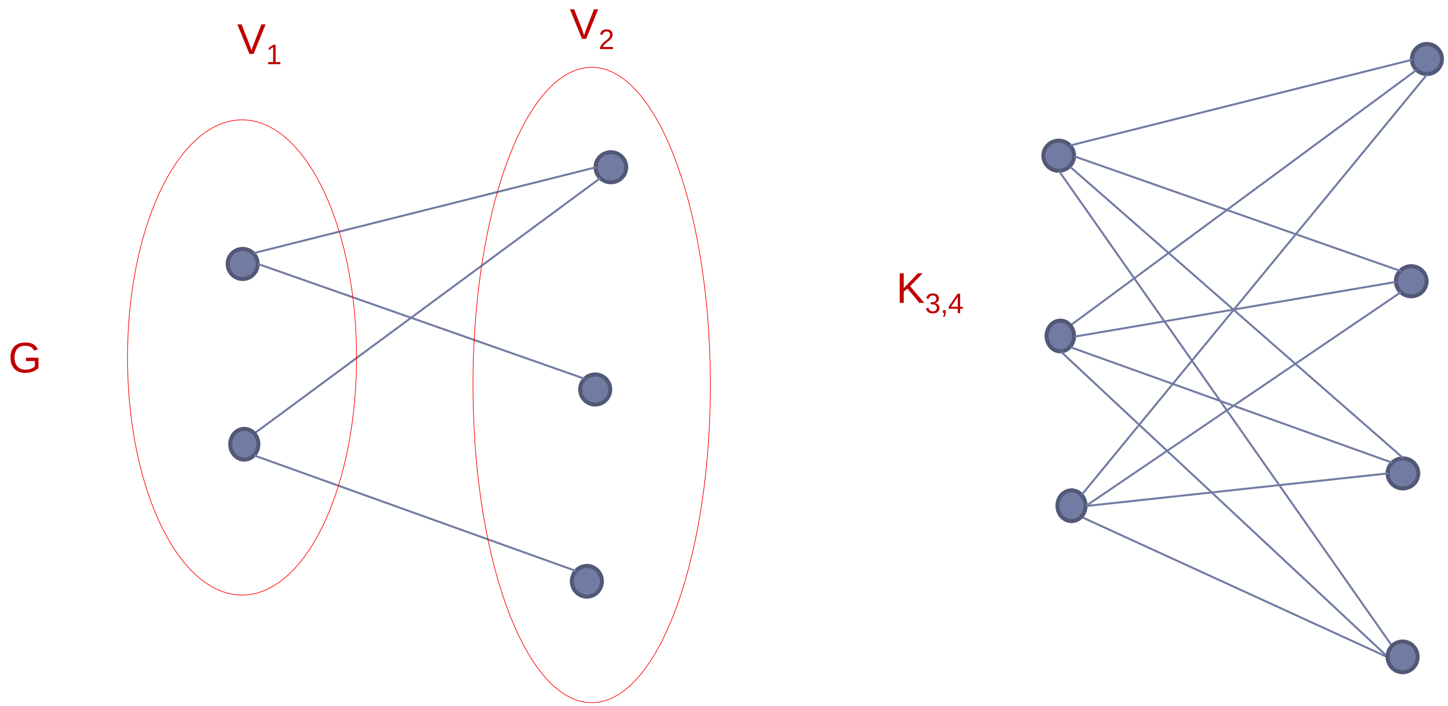
□

Powyższe twierdzenie jest m. in. podstawą algorytmu szukania mksymalnego skojarzenia w grafie dwudzielnym.

Patrz – problem kojarzenia małżeństw w dalszej części prezentacji.

Graf dwudzielny - przypomnienie

Graf G=(V, E) nazywamy grafem dwudzielnym jeśli zbiór V jest sumą dwóch niepustych zbiorów rozłącznych V_1 i V_2 takich, że każda krawędź w grafie G łączy wierzchołek ze zbioru V_1 z wierzchołkiem ze zbioru V_2. Zbiory V_1 i V_2 nazywamy klasami dwudzielności grafu G.



Badanie dwudzielności grafu - kolorowanie

Rozpatrujemy graf G=(V,E) o n wierzchołkach i m krawędziach.

Algorytm kolorowania do badania dwudzielności grafu

Algorytm wykorzystuje przeszukiwanie grafu, wierzchołki oznaczamy naprzemiennie kolorami np. 1 i -1.

Graf jest dwudzielny, jeśli nie natrafimy na wierzchołek o różnych kolorach. Otrzymujemy wtedy podział na dwa zbiory wierzchołków: do V_1 mogą należeć wszystkie wierzchołki o kolorze 1 do V_2 wszystkie wierzchołki o kolorze -1.

Złożoność obliczeniowa jest taka sama jak przeszukiwań, czyli O(n+m).

Badanie dwudzielności grafu – pseudokod algorytmu kolorowania dla przeszukiwania wszerz

```
L – tablica odwiedzeń
Wyzeruj L; wynik :=true; L[v]:=1; Q:={v}
While (Q nie jest pusta) and (wynik) do
begin
    u:=head[Q];
    for v z N(u) do
        if L[v]=0 then
        begin
            L[v]:=-L[u]; Enqueue(Q,v)
        end
        else if L[v]<>-L[u] then wynik:=false
        Dequeue (Q)
end
if wynik then TAK else NIE
```

Własność grafów dwudzielnych

Twierdzenie 11.2

Graf jest dwudzielny wtedy i tylko wtedy, gdy każdy jego cykl ma parzystą długość.

Dowód: $\rightarrow$ Załóżmy najpierw, że graf G=(V,E) jest dwudzielny czyli, że V można podzielić na dwa rozłączne zbiory wierzchołków V_1 oraz V_2 zgodnie z definicją. Rozważmy cykl $v_1 \rightarrow v_2 \rightarrow \ldots \rightarrow v_{k-1} \rightarrow v_k \rightarrow v_1$ o k elementach. Bez straty ogólności możemy założyć, że $v_1 \in V_1$. Ponieważ pomiędzy wierzchołkami z V_1 nie ma krawędzi, to $v_2 \in V_2$. Z kolei $v_3 \in V_1$, a $v_4 \in V_2$ i tak dalej. Tak więc każdy v_i o nieparzystym indeksie i należy do V_1. W konsekwencji v_k musi mieć parzysty indeks k, aby mógł być połączony z v_1. W rezultacie otrzymujemy, że cykle muszą być parzystej długości.

$\leftarrow$Niech graf G będzie spójny. Musimy podzielić zbiór V na dwa zbiory wierzchołków V_1,V_2, by, dla i=1,2, żadne dwa wierzchołki z V_i nie były ze sobą połączone. Wybierzmy z V dowolny wierzchołek v. Niech V_1 będzie zbiorem, do którego należy v oraz wszystkie wierzchołki, do których można dojść z v ścieżką parzystej długości, zaś V_2 niech składa się z pozostałych wierzchołków. Załóżmy, że $u_1,u_2 \in V_1$. Wtedy oczywiście istnieją ścieżki $v \rightarrow \ldots \rightarrow u_1$ oraz $u_2 \rightarrow \ldots \rightarrow v$ o parzystej długości. Gdyby u_1,u_2 były połączone krawędzią, to dostalibyśmy cykl $v \rightarrow \ldots \rightarrow u_1 \rightarrow u_2 \rightarrow \ldots \rightarrow v$ o nieparzystej długości. Dalej wystarczy zauważyć że V_2 składa się z tych wierzchołków grafu G, do których z początkowo wybranego wierzchołka v można dojść jedynie ścieżkami nieparzystej długości.

Niech graf G ma l>1 spójnych składowych $C_1,\ldots,C_l$. Wtedy każdą spójną składową C_i możemy podzielić na zbiory C_{1i},C_{2i} świadczące o dwudzielności grafu indukowanego G_{Ci}. W konsekwencji daje to podział V na $C_{11} \cup \ldots \cup C_{1l}$ oraz $C_{21} \cup \ldots \cup C_{2l}$ świadczący o dwudzielności całego grafu G. □

Badanie dwudzielności grafu – parzyste cykle

Rozpatrujemy graf G=(V,E) o n wierzchołkach i m krawędziach.

Algorytm sprawdzający czy każdy cykl ma parzystą długość

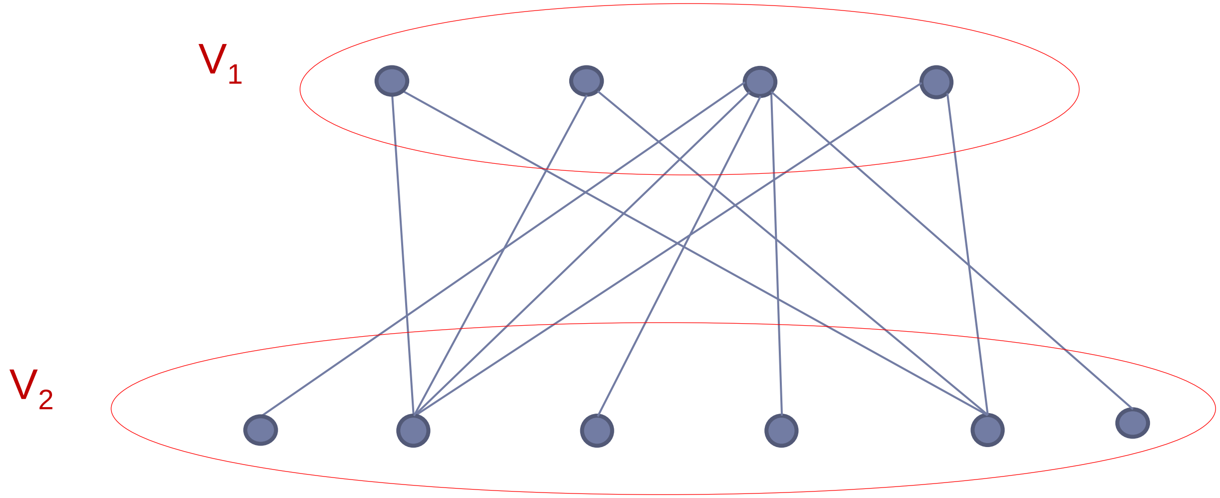
Algorytm wykorzystuje przeszukiwanie grafu w głąb. Przy natrafieniu na krawędź powracającą sprawdzamy czy wytyczony cykl jest parzystej długości. Gdy natrafimy na cykl nieparzystej długości, graf nie jest dwudzielny.

Złożoność obliczeniowa jest taka sama jak przeszukiwania w głąb, czyli O(n+m).

Skojarzenia w grafach dwudzielnych – obsadzanie stanowisk

Mamy czterech kandydatów $a_1, a_{2,} a_3, a_4$ oraz sześć stanowisk do obsadzenia $p_1, p_{2,} p_3, p_4, p_5, p_6$. Przy kandydatach podano kwalifikacje do obsadzenia stanowisk: $a_1 : p_{2,} p_5$, $a_2 : p_{2,} p_5$, $a_3, : p_{1,} p_2, p_3, p_4, p_6$, $a_4, : p_2 , p_5$.

Sytuację możemy przedstawić jako graf dwudzielny, gdzie V_1 to kandydaci, a V_2 to stanowiska.



Szukamy najliczniejszego skojarzenia kandydatów ze stanowiskami.

Skojarzenia w grafach dwudzielnych – kojarzenie małżeństw

Dane:

m dziewczyn i pewna liczba chłopców, każda dziewczyna wskazuje co najmniej jednego chłopca, za którego chciałaby wyjść za mąż (preferencje).

Wynik: pary małżeńskie dla wszystkich dziewcząt

Przykład:

m=5, panny A, B, C, D, E, kawalerowie F, G, H, I , J

Preferencje:

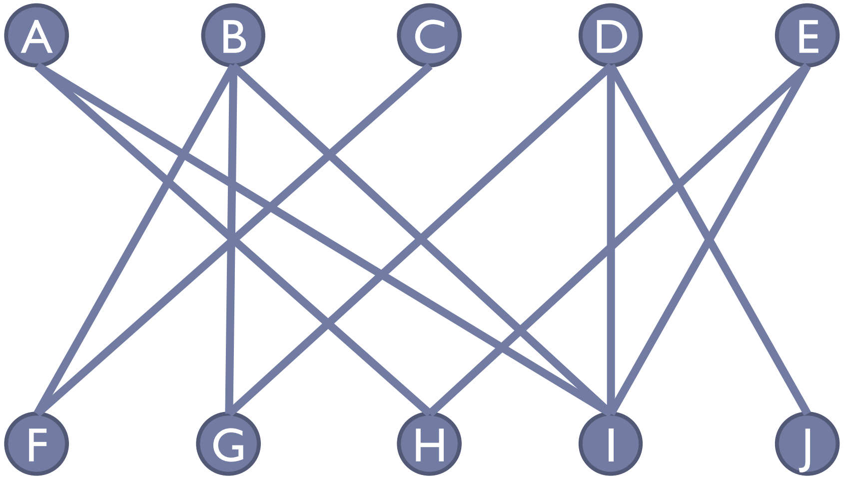
A: H, I;

B: F, G, I

C: F

D: G, I, J

E: H, I



Kojarzenie małżeństw – kiedy ma rozwiązanie?

Twierdzenie 11.3 (Hall, 1935).

Problem kojarzenia małżeństw z m dziewczynami ma rozwiązanie wtedy i tylko wtedy, gdy zachodzi warunek, że każde k dziewczyn, $0<k<=m$, zna łącznie nie mniej niż k chłopców.

Dowód. Indukcja po m. Dla $m = 1$ oczywiste.

Załóżmy teraz, że mamy $m > 1$ dziewcząt. Są możliwe tylko dwa przypadki.

1. Każde k dziewcząt, dla każdego $k < m$, zna łącznie przynajmniej $k+1$ chłopców (jeden jest zawsze w zapasie). Bierzemy wtedy pewną dziewczynę i wydajemy ją za mąż za pewnego jej znajomego. Dla pozostałych $m - 1$ dziewcząt i pozostałych chłopców warunek jest nadal spełniony (jeden chłopiec był zawsze w zapasie), więc kojarzymy je na mocy założenia indukcyjnego.

2. Pewien zbiór k dziewcząt zna dokładnie k chłopców, dla pewnego $k < m$. Wydajemy je za nich za mąż wobec założenia indukcyjnego. Dla pozostałej grupy $m - k$ dziewcząt i pozostałych chłopców warunek też zachodzi. Gdyby pewne $l<=m - k$ dziewcząt wśród tych pozostałych chłopców znało ich łącznie mniej niż l, to one wraz z już wybranymi k dziewczynami, czyli $l + k$ dziewczyn, znałoby łącznie mniej jak $l + k$ chłopców, co przeczy założeniu. Zatem pozostałe $m - k$ dziewczyn też się uda indukcyjnie wydać za mąż za pozostałych chłopców. □

Kojarzenie małżeństw – symulacja algorytmu

n=5, panny A, B, C, D, E, kawalerowie F, G, H, I , J

Preferencje:

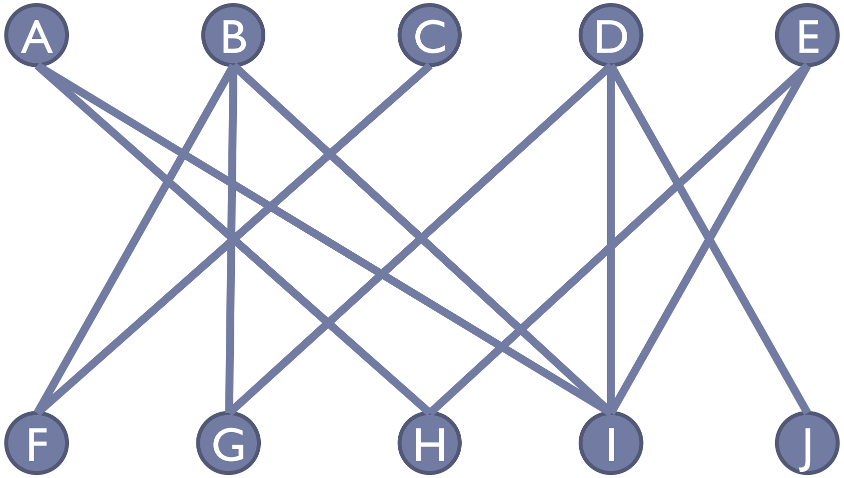
A: H, I;

B: F, G, I

C: F

D: G, I, J

E: H, I



Przykład – rozwiązanie – możliwe skojarzenie

n=5, panny A, B, C, D, E, kawalerowie F, G, H, I , J

Preferencje:

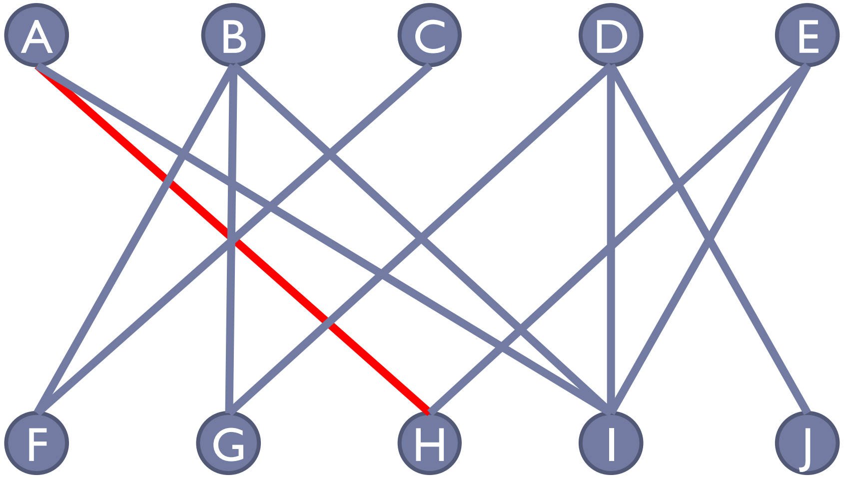
A: H, I;

B: F, G, I

C: F

D: G, I, J

E: H, I



Przykład – rozwiązanie – możliwe skojarzenie

n=5, panny A, B, C, D, E, kawalerowie F, G, H, I , J

Preferencje:

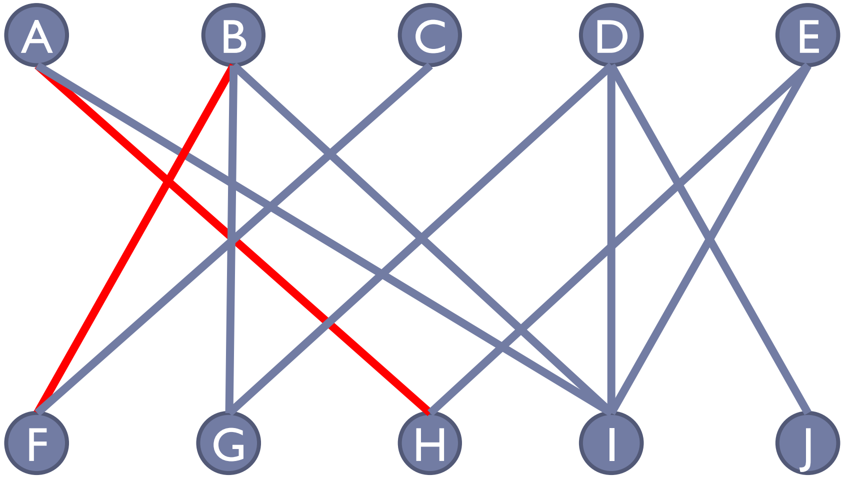
A: H, I;

B: F, G, I

C: F

D: G, I, J

E: H, I



Przykład – rozwiązanie – niemożliwe skojarzenie

n=5, panny A, B, C, D, E, kawalerowie F, G, H, I , J

Preferencje:

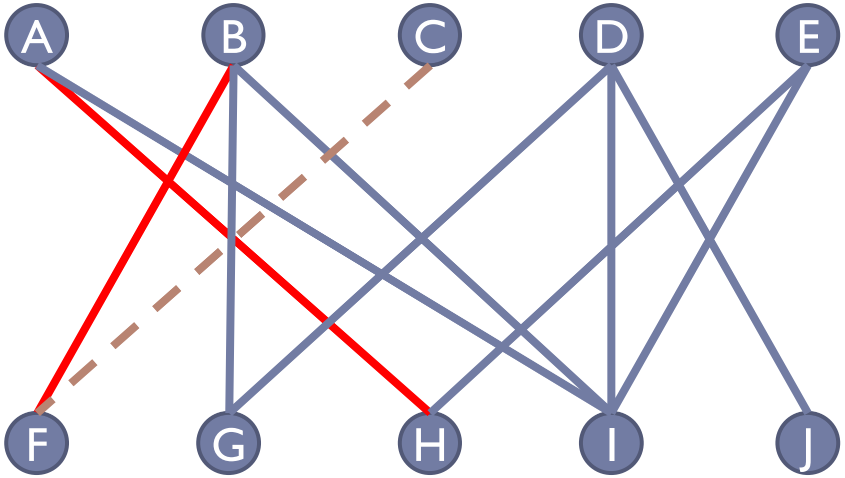
A: H, I;

B: F, G, I

C: F

D: G, I, J

E: H, I



Zamieniamy rolami krawędzie na drodze powiększającej CFBG.

Przykład – rozwiązanie – liczniejsze skojarzenie

n=5, panny A, B, C, D, E, kawalerowie F, G, H, I , J

Preferencje:

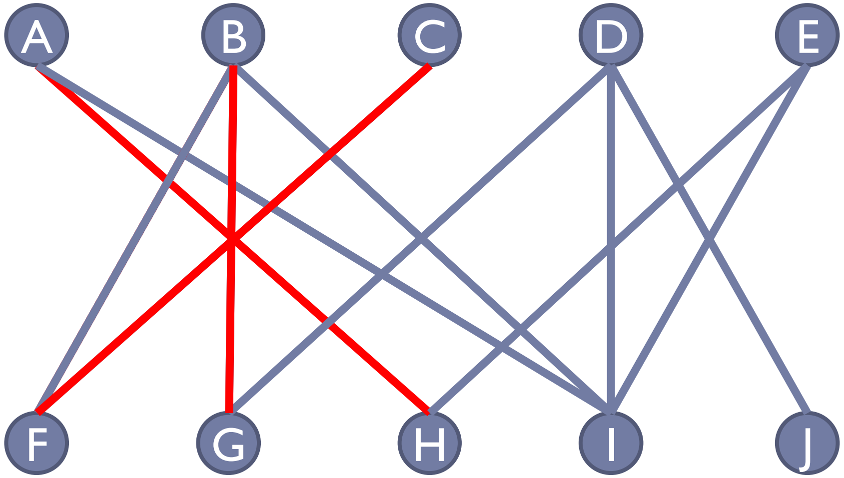
A: H, I;

B: F, G, I

C: F

D: G, I, J

E: H, I



.

Przykład – rozwiązanie

n=5, panny A, B, C, D, E, kawalerowie F, G, H, I , J

Preferencje:

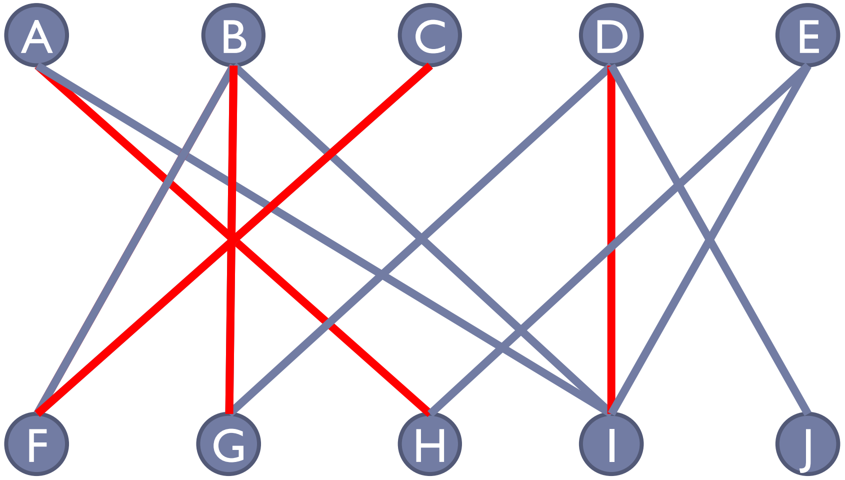
A: H, I;

B: F, G, I

C: F

D: G, I, J

E: H, I



Przykład – rozwiązanie

n=5, panny A, B, C, D, E, kawalerowie F, G, H, I , J

Preferencje:

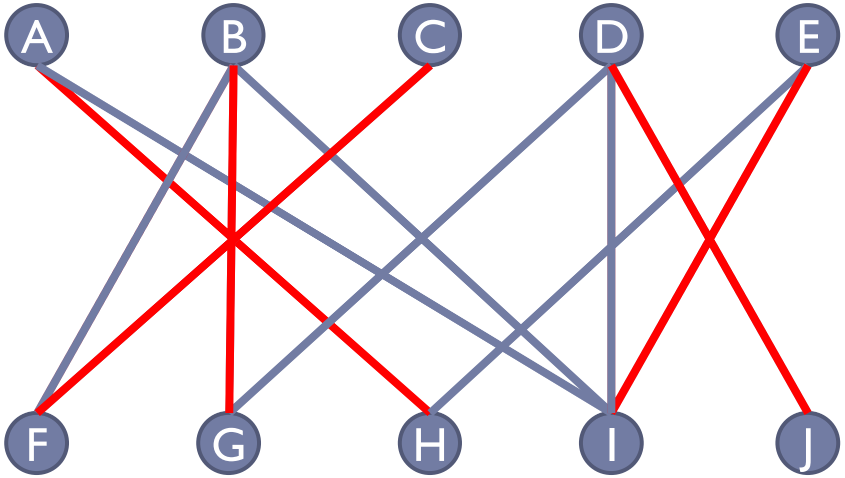
A: H, I;

B: F, G, I

C: F

D: G, I, J

E: H, I



Algorytm i struktury na potrzeby realizacji algorytmu

Dla każdej nieskojarzonej dziewczyny utwórz ścieżkę rozszerzającą prowadzącą do pierwszego napotkanego wolnego chłopca (BFS).

Ścieżka rozszerzająca na przemian ma ścieżki wolne i zajęte. Po znalezieniu jej zamieniamy wszystkie krawędzie wolne na zajęte i odwrotnie.

1. Drzewo BFS pamiętamy w tablicy ojców:
   - jeśli wierzchołek jest dziewczyną, dołączamy do niego wszystkich nieodwiedzonych kawalerów (krawędzie wolne),
   - jeśli wierzchołek jest chłopcem dodajemy do drzewa tylko krawędź skojarzoną.
2. Tablica n elementowa określająca rodzaj wierzchołka: kawaler (true), panna (false).
3. Tablica n elementowa pamiętająca skojarzenia – każdy element pamięta numer skojarzonego wierzchołka

Złożoność obliczeniowa

Dla grafu dwudzielnego G=(V,E) o n wierzchołkach i m krawędziach:

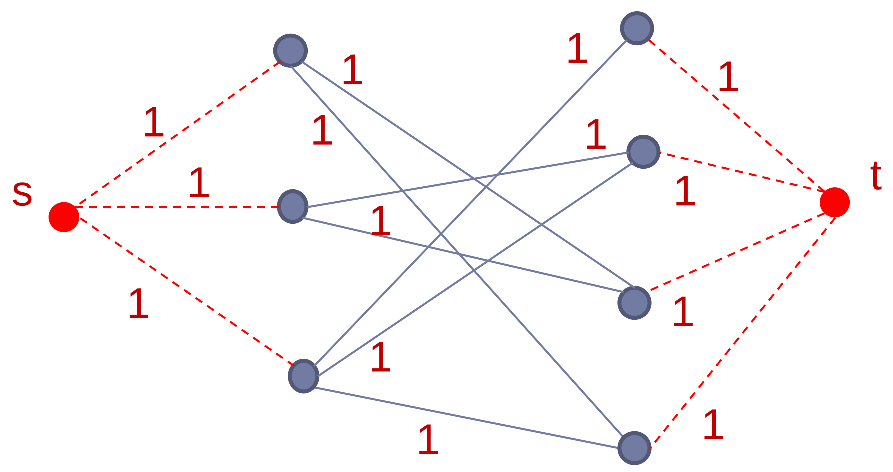
- Ograniczeniem górnym na rozmiar maksymalnego skojarzenia jest n/2.
- W każdym kroku pętli rozmiar skojarzenia rośnie o jeden.
- Pętla jest wykonywana co najwyżej O(n) razy.
- Wyszukanie jednaj drogi powiększającej zajmuje czas O(m)

Złożoność obliczeniowa algorytmu szukania maksymalnego skojarzenia w grafie dwudzielnym za pomocą dróg powiększajacych wynosi O(mn).

Maksymalne skojarzenie w grafie dwudzielnym – zastosowanie przepływu w sieciach

Rozważany wcześniej algorytm znajdowania maksymalnego skojarzenia w grafie dwudzielnym wykorzystywał metodę BFS do znajdowania naprzemiennych ścieżek rozszerzających.

Do tego samego wyniku możemy dojść stosując algorytmy znajdujące maksymalny przepływ w sieci. W tym celu musimy zmodyfikować graf dwudzielny dodając węzeł źródła s oraz węzeł ujścia t.



W celu znalezienia liczności maksymalnego skojarzenia wystarczy nadać wszystkim krawędziom wagi 1, rozważyć przepływ od źródła do ujścia i zastosować algorytm szukania maksymalnego przepływu w sieci.

Dziękuję z uwagę

dr Anna Beata Kwiatkowska

aba@mat.umk.pl

tel. 602 184 813