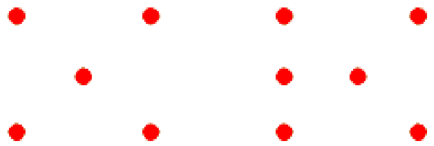
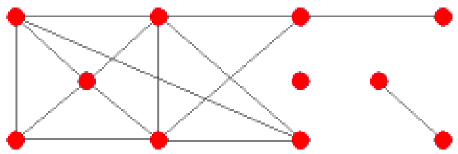
Wprowadzenie do teorii grafów

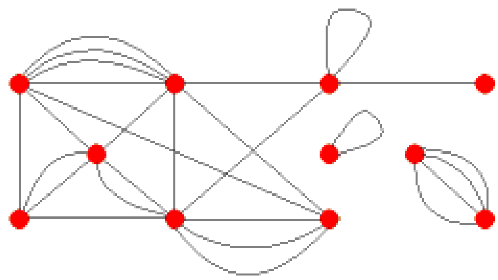
Grafem (ang. *graph*) nazywamy uporządkowaną parę $G = (V, E)$, gdzie: V (ang. *vertices*) oznacza zbiór wierzchołków (na poniższym rysunku z lewej strony zaprezentowane zostały jako czerwone kropki), a E (ang. *edges*) rodzinę krawędzi łączących wierzchołki (rysunek po prawej stronie), przy czym graf musi mieć co najmniej jeden wierzchołek ($V \neq \emptyset$). Jeśli graf posiada wierzchołki, a nie posiada żadnej krawędzi, nazywamy go **zerowym** (ang. *null graph*). **Rząd grafu** (ang. *graph order*), to liczba wierzchołków w grafie, zaś **rozmiar grafu** (ang. *graph size*), to liczba krawędzi w grafie.



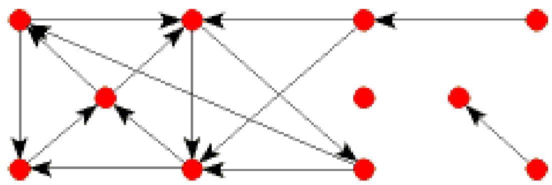


Liczbę wierzchołków grafu oznaczamy najczęściej symbolem n, a liczbę krawędzi symbolem m albo q. Grafy wykorzystywane są powszechnie, między innymi w elektrotechnice i elektronice, w chemii, biologii, ekonomii oraz we wszystkich obszarach informatyki.

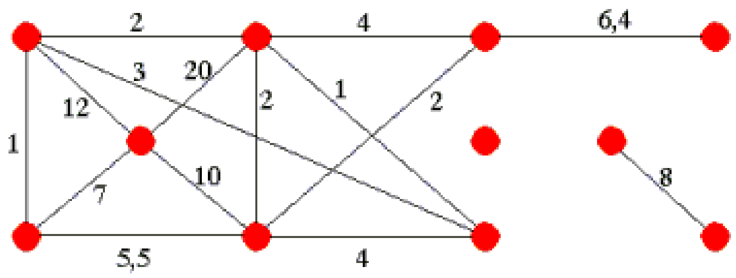
Krawędzią grafu $e \in E$ jest uporządkowana para wierzchołków $e = (v_1, v_2)$, $v_1 \in V$, $v_2 \in V$. Jeśli rodzina E zawiera powtarzające się elementy, to takie krawędzie nazywamy równoległymi lub wielokrotnymi (rysunek poniżej). Graf zawierający pętle lub krawędzie wielokrotne nazywamy **multigrafem** (ang. *multigraph*).



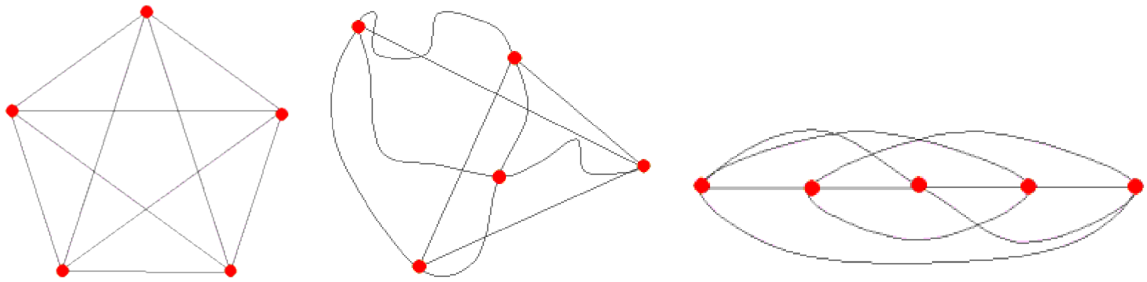
Jeśli $(v_1, v_2) \in E$ i $(v_2, v_1) \in E$, to taka para nazywana jest **krawędzią niezorientowaną** (nieskierowaną), a graf nazywamy jest najczęściej jako nieskierowany. Na powyższym rysunku wszystkie krawędzie są niezorientowane. W przeciwnym wypadku, gdy $(v_1, v_2) \in E$ i $(v_2, v_1) \notin E$, krawędź nazywamy **łukiem** albo **krawędzią zorientowaną** (skierowaną), przy czym jeśli łuk łączy wierzchołek v z w, to v jest poprzednikiem wierzchołka w, a w jest następnikiem wierzchołka v. Przykład grafu skierowanego (ang. *directed graph* albo *digraph*) znajduje się na poniższym rysunku.



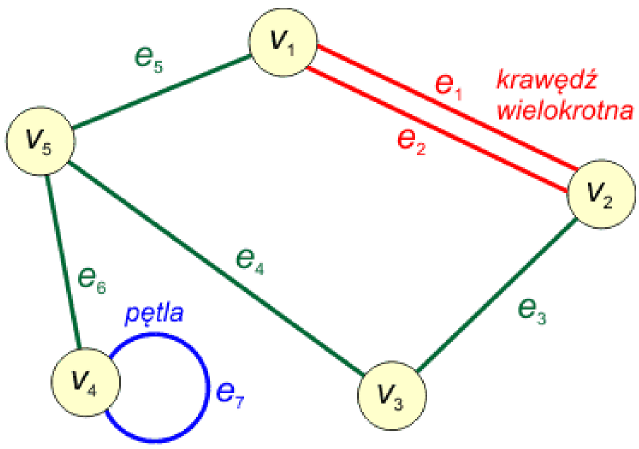
Odrębną klasę grafów stanowią **grafy ważone** (ang. *weighted graph*). Graf $G = (V, E)$ nazywamy ważonym, jeśli istnieje funkcja $w: E \to R^+$ przyporządkowująca każdej krawędzi dodatnią liczbę rzeczywistą. W zależności od kontekstu jest ona nazywana wagą albo kosztem krawędzi (rysunek poniżej). Waga (koszt) może oznaczać np. odległość między wierzchołkami.



Należy pamiętać, że rysunek grafu to tylko jedna z jego wielu reprezentacji graficznych, ponieważ każdy graf można narysować na wiele sposobów. Poniżej kilka reprezentacji graficznych tego samego grafu.

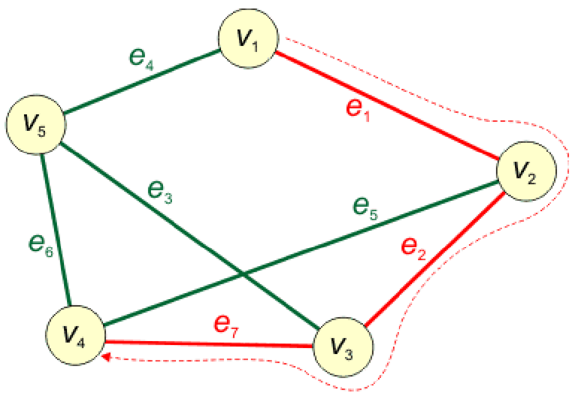


Stopniem wierzchołka grafu (ang. *degree*) nazywamy liczbę krawędzi, które łączą się z danym wierzchołkiem, przy czym, jeśli graf posiada pętle, to liczymy je za 2. Na poniższym rysunku grafu wierzchołki posiadają następujące stopnie: $deg(v_1) = 3$, $deg(v_2) = 3$, $deg(v_3) = 2$, $deg(v_4) = 3$, $deg(v_5) = 3$.



W grafie skierowanym rozróżniamy **stopień wchodzący**, tj. liczbę krawędzi wchodzących do wierzchołka oraz **stopień wychodzący**, tj. liczbę krawędzi wychodzących z wierzchołka.

Ścieżka lub droga (ang. *path*) jest uporządkowanym ciągiem kolejnych krawędzi (lub kolejno mijanych wierzchołków), po których należy przejść, aby dotrzeć z wierzchołka startowego do wierzchołka końcowego. W grafie może istnieć wiele różnych ścieżek pomiędzy dwoma wybranymi wierzchołkami, zaś **najkrótszą ścieżką** (ang. *shortest path*) nazywamy ścieżkę, która zawiera najmniej krawędzi (wierzchołków). Mówimy przy tym, że ścieżka jest **prosta**, jeśli każdą krawędź (wierzchołek) przechodzimy tylko jeden raz.



W przykładowym grafie ścieżki P między wierzchołkami v_1 oraz v_4 mogą mieć następujący opis oraz przebieg:

$P(v_1, v_4) = \{e_1, e_2, e_7\}$

$P(v_1, v_4) = \{e_1, e_5\}$

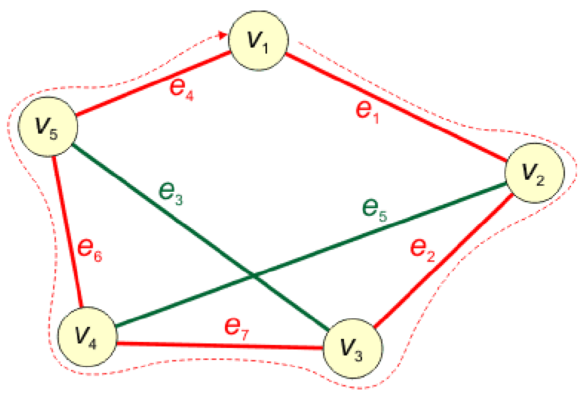
$P(v_1, v_4) = \{e_4, e_6\}$

...

$P(v_1, v_4) = \{v_1, v_2, v_3, v_4\}$

...

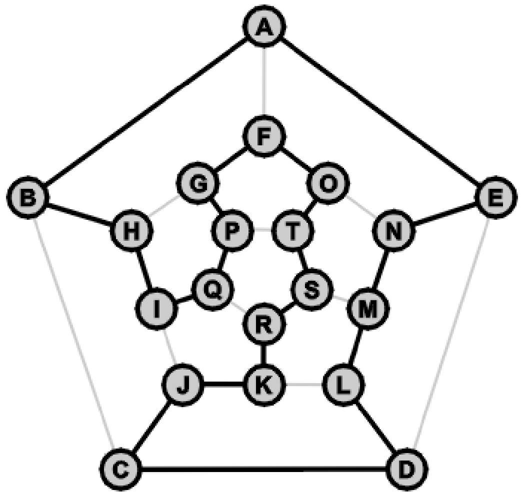
Cyklem (ang. *cycle*) nazywamy ścieżkę, która rozpoczyna się i kończy w tym samym wierzchołku, przy czym nie należy mylić cyklu z pętlą, czyli pojedynczą krawędzią. Na rysunku poniżej cykl dla wierzchołka v_1 można opisać przykładowo: $C(v_1) = \{e_1, e_2, e_7, e_6, e_4\}$; $C(v_1) = \{v_1, v_2, v_3, v_4, v_5, v_1\}$ itd.



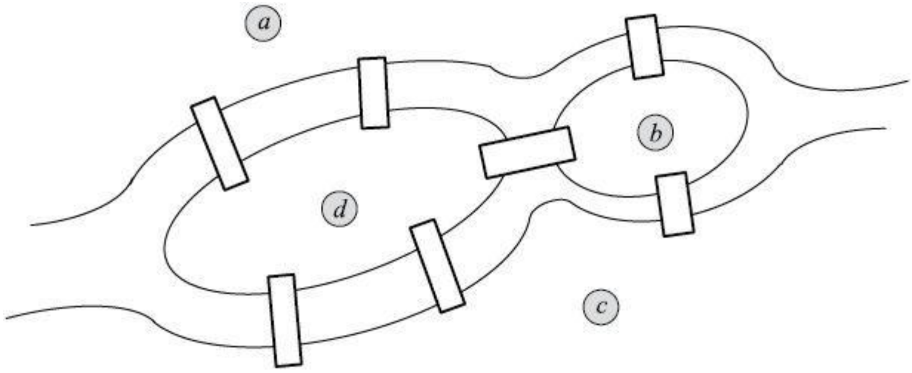
Cykl nazywamy **prostym** (ang. *simple cycle*), jeśli każda jego krawędź (wierzchołek) jest przechodzona dokładnie jeden raz. Oczywiście nie uwzględniamy przy tym faktu, że w cyklu wierzchołek startowy i końcowy jest tożsamy. Inaczej mówiąc, ścieżka musi być zamknięta.

Ścieżka prosta zawierająca wszystkie krawędzie grafu nosi nazwę **ścieżki Hamiltona** (ang. *Hamiltonian path*), zaś cykl prosty zawierający wszystkie wierzchołki grafu nazywa się **cyklem Hamiltona** (ang. *Hamiltonian cycle*). Natomiast ścieżka prosta, która przechodzi przez wszystkie krawędzie grafu, nazywana jest **ścieżką (cyklem) Eulera** (ang. *Eulerian path*). Cykl Eulera i cykl Hamiltona nie są tożsame, ponieważ w cyklu Hamiltona ważne jest przejście przez wszystkie wierzchołki dokładnie jeden raz (niektóre krawędzie grafu mogą być w ogóle nie przechodzone), natomiast w cyklu Eulera musimy z kolei przejść przez każdą krawędź, zatem niektóre wierzchołki mogą zostać kilkakrotnie odwiedzone wtedy, gdy łączą się kilkoma krawędziami.

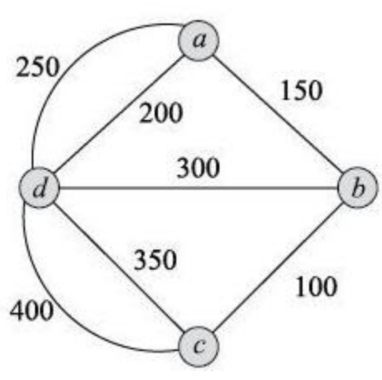
Rysunek poniżej pokazuje przykładowy graf z cyklem Hamiltona:



Klasycznym przykładem początku teorii grafów jest problem mostów królewieckich, sformułowany i rozwiązany właśnie przez Leonarda Eulera. Uczony zastanawiał się na spacerze nad rzeką Pregoła w Królewcu, czy da się tak zaprojektować trasę przejścia przez mosty, żeby startując z dowolnego punktu przejść wszystkie siedem mostów przekraczając każdy z nich tylko jeden raz (literami b i d oznaczono wyspy na rzece).



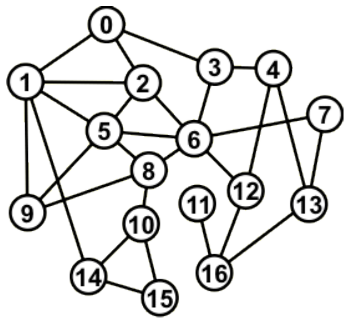
Problem ten można przedstawić za pomocą grafu ważonego (rysunek poniżej). Wierzchołki grafu odpowiadają częściom lądowym parku, krawędzie mostom, a wagi można interpretować jako długości odcinków drogi. Trzeba znaleźć w tym grafie cykl Eulera, czyli cykl przechodzący przez wszystkie wierzchołki i wszystkie krawędzie grafu, ale przez każdą krawędź tylko jeden raz. Łatwo sprawdzić, że odpowiedź na tak postawiony problem jest negatywna.

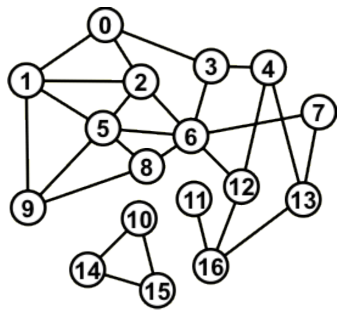


Leonard Euler uogólnił ten problem dla dowolnych grafów, zaś od tamtego czasu grafy, w których istnieje droga zamknięta spełniająca omówione warunki, noszą nazwę **grafów Eulera**. W związku z tym pierwsze twierdzenie grafów, sformułowane przez Eulera w roku 1736, brzmi:

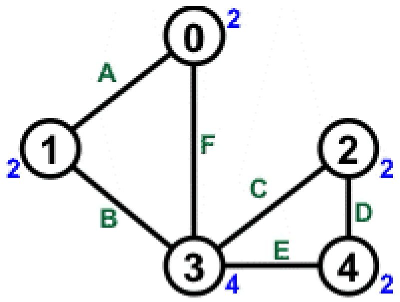
W grafie można znaleźć cykl Eulera wtedy i tylko wtedy, gdy graf jest spójny i każdy jego wierzchołek ma parzysty stopień.

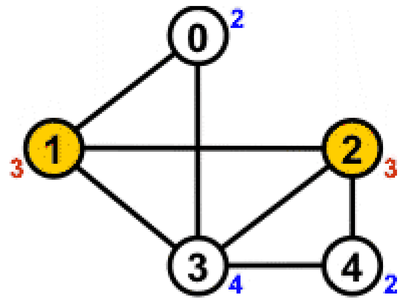
Graf jest **spójny** (ang. *connected graph*), jeśli dla każdych dwóch jego wierzchołków istnieje ścieżka, które je łączy. Na rysunkach pokazano graf spójny (z lewej) oraz graf niespójny (z prawej).





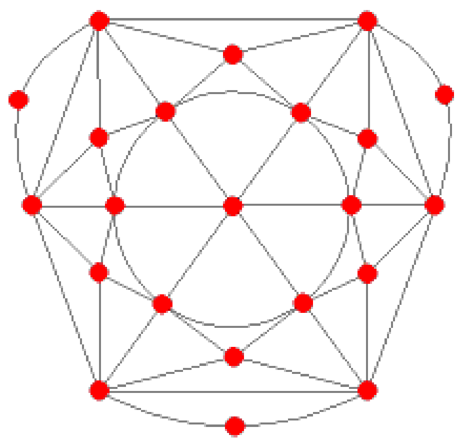
Rysunki poniżej przedstawiają graf z cyklem Eulera (z lewej) oraz bez cyklu Eulera (z prawej):





Jak widać, w grafie spójnym z cyklem Eulera każdy wierzchołek musi mieć parzysty stopień.

Zadanie 1.
Samodzielnie znajdź cykl Eulera w grafie przedstawionym na rysunku. Uwaga: Zadanie ma wiele rozwiązań.



Przykład 1.
Mamy do rozwiązania za pomocą grafu zmodyfikowaną wersję starej zagadki z kupcem, kozą, wilkiem i kapustą, który miał się przeprawić przez rzekę za pomocą małej łódki. Naraz mógł się do niej zmieścić tylko kupiec i tylko jedno zwierzę (albo tylko kapusta), a przy każdej przeprawie w łódce musiał znajdować się kupiec. Nie wolno przy tym zostawić ze względów bezpieczeństwa na tym samym brzegu ani wilka z kozą, ani kozy z kapustą. Dodatkowo każda przeprawa przez rzekę wymaga uiszczenia pewnej opłaty, które wynoszą: $c_{wi} = 8$, $c_{ko} = 5$, $c_{ka} = 1$, $c_{ku} = 10$. Pytanie problemowe brzmi, jaką strategię powinien przyjąć kupiec, aby bezpiecznie przewieźć na drugi brzeg kozę, wilka i kapustę, minimalizując przy tym łączną cenę przeprawy?

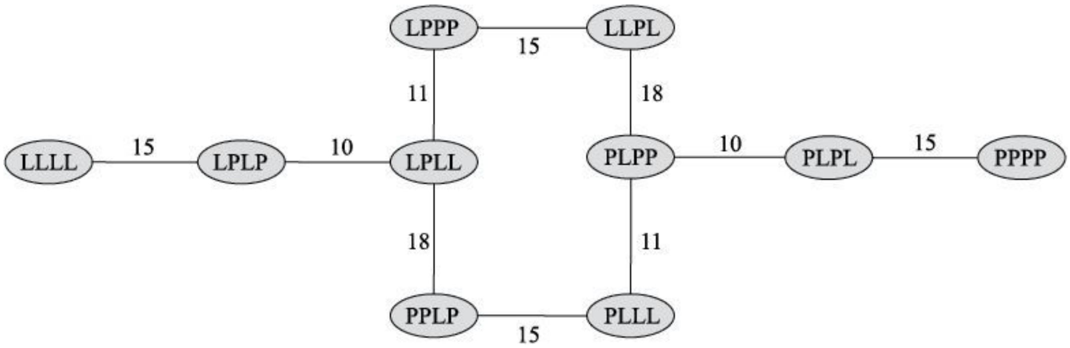
Rozwiązanie:
W naszej analizie zakładamy sztywny porządek kolejnych elementów: wilk, koza, kapusta, kupiec. Zatem można tutaj wyróżnić $2^4 = 16$ stanów binarnych, przykładowo skrót LPLL oznacza, że wilk, kapusta i kupiec są na lewym brzegu, zaś koza na prawym. Zakładamy przy tym, że przemieszczenie następuje z lewego brzegu na prawy.

Z warunkach zadania wynika, że niektóre stany są zabronione, na przykład zabroniona jest kombinacja LLLP, która oznacza, że wilk, koza i kapusta są na jednym (lewym) brzegu, a kupiec na drugim, ponieważ koza może zjeść kapustę, a wilk kozę. W sumie mamy sześć kombinacji zabronionych: LLLP, PLLP, PPLL, LPPL, LLPP, PPPL. Do rozważenia pozostaje dziesięć, które umieścimy najpierw w tabeli dopuszczalnych stanów.

	LLLL	PLLL	PPLP	LPPP	LPLL	LPLP	LLPL	PLPL	PLPP	PPPP
LLLL	0	0	0	0	0	1	0	0	0	0
PLLL	0	0	1	0	0	0	0	0	1	0
PPLP	0	1	0	0	1	0	0	0	0	0
LPPP	0	0	0	0	1	0	1	0	0	0
LPLL	0	0	1	1	0	1	0	0	0	0
LPLP	1	0	0	0	1	0	0	0	0	0
LLPL	0	0	0	1	0	0	0	0	1	0
PLPL	0	0	0	0	0	0	0	0	1	1
PLPP	0	1	0	0	0	0	1	1	0	0
PPPP	0	0	0	0	0	0	0	1	0	0

Symbolem „1” oznaczono w tabeli dozwolone przejście. Przykładowo, dozwolone jest przejście od kombinacji LLLL do LPLP (i odwrotnie), bo oznacza ono, że koza z kupcem (który musi być zawsze w łódce) przeprawia się na drugi brzeg, a wilk i kapusta pozostają. Tak skonstruowana tabela jest przykładem tzw. macierzy sąsiedztwa grafu.

Pozostaje teraz przedstawienie tabeli w postaci grafu ważonego, w którym z każdą krawędzią związana jest waga w postaci ceny przejścia z jednego wierzchołka do drugiego. Celem jest przejście od stanu LLLL do PPPP przy jak najmniejszej opłacie łącznej (jak najmniejszym kosztem).

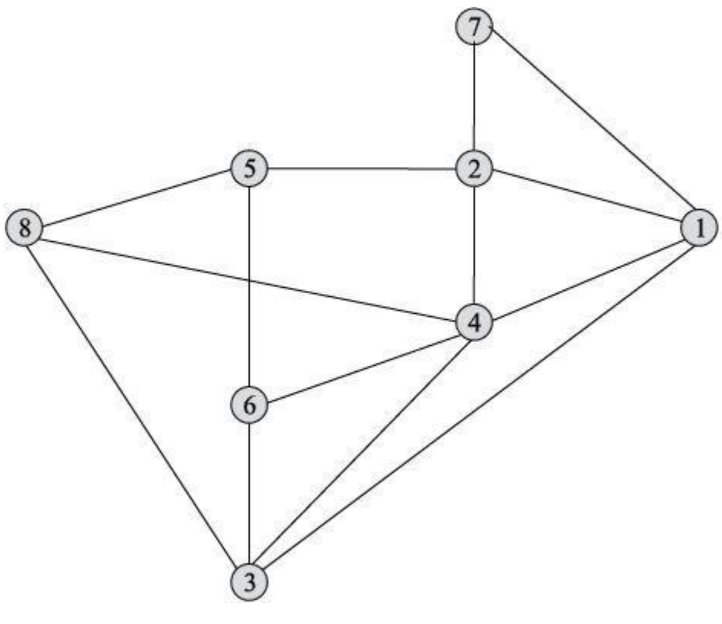


Jak wynika z grafu, możliwe są dwa rozwiązania przedstawionego problemu, które mają identyczny koszt równy 94.

Przykład 2.
Należy przygotować rozkład terminów sesji egzaminacyjnej dla ośmiu przedmiotów, oznaczonych kolejnymi numerami $1 \cdots 8$. Stan zabroniony zachodzi tutaj wtedy, gdy w danym terminie chociaż jeden student uczęszcza na dwa przedmioty objęte egzaminem. W tabeli pokazano przekazane przez dziekanat wszystkie stany zabronione, oznaczone jako X.

	1	2	3	4	5	6	7	8
1		X	X	X			X	
2	X			X	X		X	
3	X			X		X		X
4	X	X	X			X		X
5		X				X		X
6			X	X	X			
7	X	X						
8			X	X	X			

Na podstawie danych z tabeli możemy narysować graf, w którym przedmioty egzaminacyjne reprezentowane są przez wierzchołki. Krawędź łączy tutaj dwa wierzchołki wtedy, gdy z danych przedmiotów nie mogą być przeprowadzone egzaminy w tym samym terminie.



Przedstawiony problem jest mały, więc można go rozwiązać metodą intuicyjną. Przy większych problemach rozwiązanie wymaga użycia tzw. zbiorów niezależnych albo kolorowania wierzchołków grafu. W każdym przypadku należy podzielić zbiór wierzchołków grafu na możliwie najmniejszą liczbę podzbiorów takich, że elementami podzbioru są wierzchołki niepołączone w grafie krawędzią. Tutaj analiza intuicyjna prowadzi do wniosku, że niezbędne są trzy terminy, a w każdym z terminów mogą odbywać się egzaminy: $\{1, 6, 8\}$, $\{4, 5, 7\}$, $\{2, 3\}$.