Algorytmy Grafowe

dr hab. Bożena Woźna-Szcześniak, prof. UJD

Uniwersytet Jana Długosza w Częstochowie

b.wozna@ujd.edu.pl

Wykład 1 i 2

Cel przedmiotu

- Celem wykładów jest zapoznanie studentów z podstawowymi algorytmami grafowymi. Przedstawiona zostanie również ich poprawność i złożoność.
- Celem laboratorium jest implementacja algorytmów wprowadzonych na wykładzie.

Program przedmiotu

- Reprezentacje grafu.
- Wyszukiwanie wszerz i wyszukiwanie w głąb.
- (Silnie) spójne komponenty.
- Sortowanie topologiczne.
- Minimalne drzewo rozpinające - Algorytmy Kruskala i Prima.
- Znajdowanie cyklu lub ścieżki Eulera. Algorytm Fleury'ego.
- Znajdowanie cyklu lub ścieżki Hamiltona
- Problem najkrótszej ścieżki: Algorytm Floyda-Warshalla
- Problem najkrótszej ścieżki: Algorytm Dijkstry
- Problem najkrótszej ściezki: Algorytm Bellmana-Forda

Efekty uczenia się

E1 Student rozpoznaje złożoność obliczeniową wybranych algorytmów grafowych prezentowanych na wykładzie – efekt weryfikowany na egzaminie.

E2 Student wyjaśnia wybrane algorytmy grafowe przedstawione na wykładzie – efekt weryfikowany na egzaminie.

E3 Student implementuje algorytmy grafowe prezentowane na wykładzie w wybranym języku programowania – efekt weryfikowany na laboratorium.

E4 Student potrafi samodzielnie pozyskiwać informacje z różnych źródeł oraz wykorzystywać je do rozwiązania postawionego. problemu – efekt weryfikowany na laboratorium.

E5 Student dostrzega potrzebę ciągłego aktualizowania i poszerzania wiedzy z zakresu algorytmiki – efekt weryfikowany na laboratorium.

Literatura

- Cormen T.H., Leiserson Ch.E., Rivest R.L. Wprowadzenie do algorytmów. WNT, 1997 i późniejsze.
- Undirected Graphs:
  `https://algs4.cs.princeton.edu/41graph/`
- Directed Graphs:
  `https://algs4.cs.princeton.edu/42digraph/`

Wprowadzenie

Cel wykładu

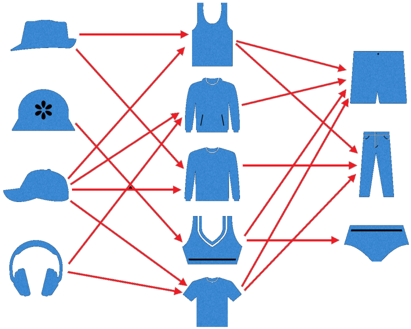
Celem wykładu jest zapoznanie z **teorią grafów** i ich zastosowaniem do rozwiązywania wybranych problemów praktycznych.

Teoria grafów

Dział matematyki zajmujący się badaniem własności grafów oraz ich zastosowań do rozwiązywania rzeczywistych problemów.

Przykład problemu z teorii grafów

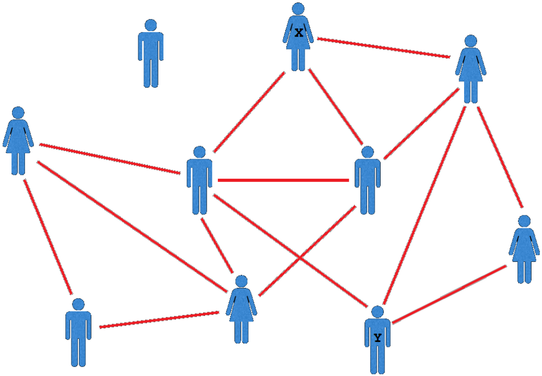
Na podstawie poniższej grafiki, ile różnych zestawów odzieży można skomponować, wybierając po jednym artykule z każdej kategorii?



Sieć społecznościowa jako problem z teorii grafów

Reprezentacja grafowa pozwala odpowiedzieć na interesujące pytania, takie jak:

- ilu znajomych ma osoba X?
- ile stopni separacji jest między osobą X a osobą Y?



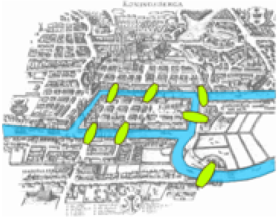
Leonhard Euler

- Za pierwszego teoretyka i badacza grafów uważa się Leonharda Eulera[1], który rozstrzygnął tzw. `zagadnienie mostów królewieckich`.

[1] Leonhard Euler (ur. 15 kwietnia 1707 r. w Bazylei - Szwajcaria, zm. 18 września 1783 r. w Petersburgu - Rosja) - szwajcarski matematyk, fizyk i astronom, jeden z twórców nowoczesnej matematyki.

Mosty królewieckie

Przez Królewiec przepływała rzeka Pregole, w której rozwidleniach znajdowały się dwie wyspy. Ponad rzeką przerzucono siedem mostów, z których jeden łączył obie wyspy, a pozostałe mosty łączyły wyspy z brzegami rzeki. Plan mostów pokazuje rysunek:

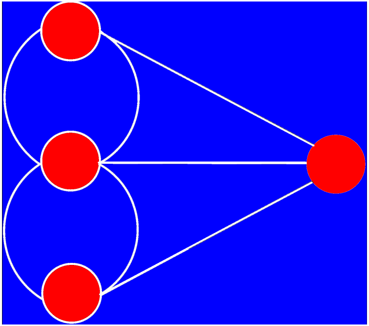


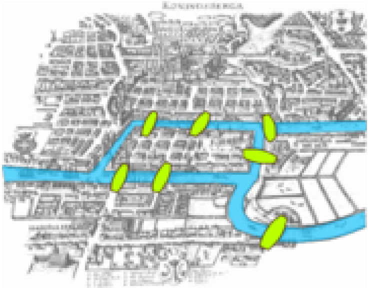
Mosty królewieckie

- Zwykłe spacerowanie szybko się znudziło mieszkańcom Królewca i zaczęli zastanawiać się, czy istnieje taka trasa spacerowa, która przechodzi przez każdy most dokładnie raz, żadnego nie omija, i pozwala wrócić do punktu wyjścia.
- Mieszkańcy nie potrafili rozwiązać postawionego problemu samodzielnie, więc postanowili napisać do matematyka **Leonharda Eulera**.

Mosty królewieckie

- Euler wykazał, że rozwiązanie problemu mieszkańców nie jest możliwe, a decyduje o tym nieparzysta liczba wylotów mostów zarówno na każdą z wysp, jak i na oba brzegi rzeki – jeśli wejdzie się po raz trzeci na wyspę, nie ma jak z niej wyjść.
- Sytuację tę można przedstawić za pomocą następującego grafu:





Rodzaje grafów

Istnieje wiele różnych rodzajów grafów. Ważne jest, aby móc rozpoznać, z jakim typem grafu pracujemy, zwłaszcza kiedy programujemy i rozwiązujemy przy pomocy grafów konkretny problem.

Rodzaje grafów rozważane na wykładach

- grafy nieskierowane (niezorientowane)
- grafy skierowane (zorientowane)
- grafy z wagami

Inne rodzaje grafów

- multigrafy
- hipergrafy

Graf nieskierowany

Definicja

Grafem nieskierowanym (grafem) nazywamy parę $G = (V, E)$, gdzie:

- V to zbiór **wierzchołków** oraz
- $E \subseteq \{\{u, v\} : u, v \in V, u \neq v\}$ to zbiór *nieuporządkowanych par* wierzchołków, zwanych **krawędziami.**

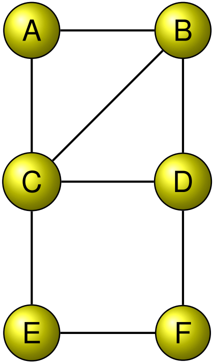
W grafie nieskierowanym nie mogą występować pętle.

Intuicyjnie

Graf nieskierowany to graf, w którym krawędzie nie mają orientacji. To znaczy, krawędź od węzła u do węzła v jest identyczna z krawędzią od węzła v do węzła u.

Graf nieskierowny $G = (V, E)$ - przykład

- $V = \{A, B, C, D, E, F\}$
- $E = \{\{A, B\}, \{A, C\}, \{B, C\}, \{B, D\}, \{C, D\}, \{E, F\}\}$
- $|V| = n = 6$, $|E| = m = 6$



Uwaga!
Na powyższym grafie węzły mogą reprezentować miasta, a krawędzie drogi dwukierunkowe.

Graf skierowany

Definicja

Grafem skierowanym (lub **digrafem**) nazywamy parę $G = (V, E)$:

- V to skończony zbiór **wierzchołków** oraz
- $E \subseteq \{(u, v) : u, v \in V\}$ to zbiór uporządkowanych par wierzchołków ze zbioru V, zwanych **krawędziami**.

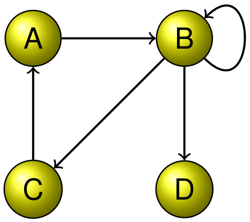
Możliwe jest istnienie **pętli** od danego wierzchołka do niego samego.

Intuicyjnie

Graf skierowany to graf, w którym krawędzie mają określoną orientację. Oznacza to, że krawędź prowadząca od węzła u do węzła v różni się od krawędzi biegnącej od węzła v do węzła u, co podkreśla kierunek relacji między węzłami.

Graf skierowny $G = (V, E)$ - przykład

- $V = \{A, B, C, D\}$
- $E = \{(A, B), (B, B), (B, C), (B, D), (C, A), (C, D), (D, C)\}$



W przedstawionym grafie:

- Węzły mogą reprezentować osoby.
- Krawędź (u, v) oznacza, że osoba u wręczyła prezent osobie v.
- Krawędź przychodząca oznacza otrzymanie prezentu.
- Krawędź wychodząca symbolizuje wręczenie prezentu.

Przykład:

- Osoba A wręczyła prezent osobie B.
- Osoba B wręczyła prezent samej sobie oraz osobom C i D.

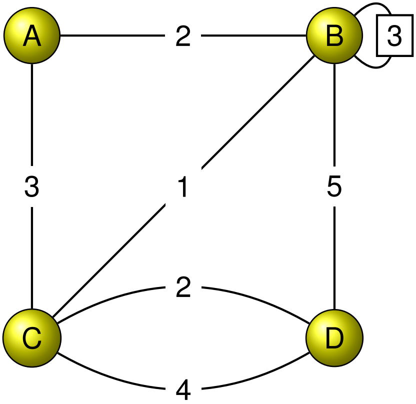
Etykietowany graf nieskierowany

Definicja

Etykietowanym grafem nieskierowanym nazywamy strukturę $G = (V, E, w : E \to R)$, gdzie

- V to zbiór wierzchołków,
- $E \subseteq \{\{u, v\} : u, v \in V\}$ to zbiór par wierzchołków ze zbioru V, zwanych krawędziami.
- $w : E \to R$ to funkcja **wagi**; wagi reprezentują pewne wielkości (np. koszt, odległość, ilość, itp.).

Etykietowany graf nieskierowany - przykład



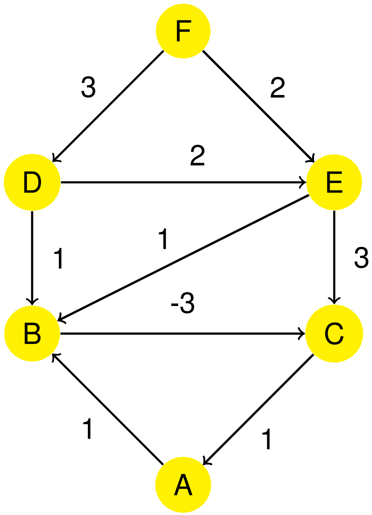
Etykietowany graf skierowany

Definicja

Etykietowanym grafem skierowanym nazywamy strukturę $G = (V, E, w : E \to R)$, gdzie

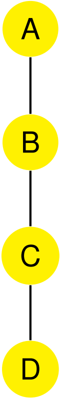
- V to zbiór wierzchołków,
- $E \subseteq \{(u, v) : u, v \in V\}$ to zbiór uporządkowanych par wierzchołków ze zbioru V, zwanych krawędziami.
- $w : E \to R$ jest funkcją **wagi**; wagi reprezentują pewne wielkości (np. koszt, odległość, ilość, itp.).

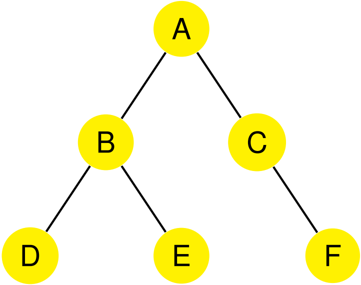
Etykietowany graf skierowany - przykład



Drzewa

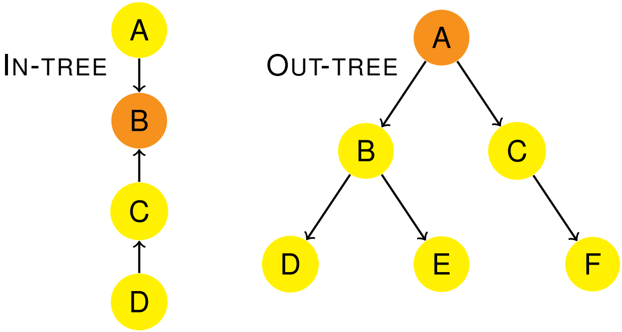
Drzewo - **graf nieskierowany bez cykli**.
Równoważnie, drzewo to **graf spójny** o n węzłach i $n - 1$ krawędziach.





Drzewa ukorzenione

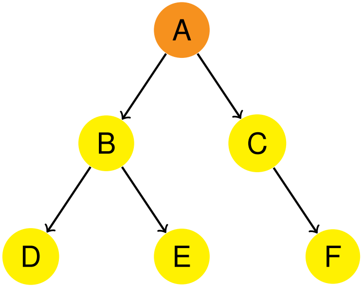
Drzewo ukorzenione to drzewo z wyróżnionym węzłem, zwanym **korzeniem**, w którym każda krawędź jest skierowana albo w kierunku od albo w kierunku do węzła korzenia.



Gdy krawędzie są odwrócone od korzenia, to mówimy, że graf jest typu **out-tree**, w przeciwnym przypadku mówimy, że graf jest typu **in-tree**.

Skierowane grafy acykliczne - ang. Directed Acyclic Graphs (DAG)

DAG to grafy skierowane bez cykli. Grafy te odgrywają ważną rolę w reprezentacji struktur danych z zależnościami. Wszystkie out-trees są DAG-ami. Odwrotna zależność nie zachodzi.



Grafy dwudzielne I

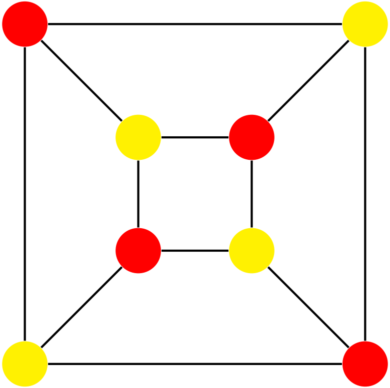
Graf dwudzielny (ang. bipartite graph lub bigraph) to graf, którego wierzchołki możemy podzielić na dwa rozłączne zbiory U i V.

Wierzchołki należące do zbioru U mogą się łączyć krawędziami tylko z wierzchołkami ze zbioru V i na odwrót.

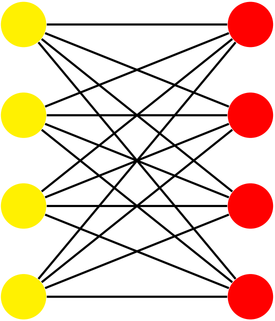
Wierzchołki należące do zbioru U możemy traktować jako pokolorowane, np. na żółto, a wierzchołki należące do zbioru V jako pokolorowane, np. na czerwono.

Graf jest dwudzielny, jeśli możemy pokolorować wszystkie jego wierzchołki w taki sposób, aby żaden z sąsiadów danego wierzchołka nie miał tego samego koloru co on.

Grafy dwudzielne II

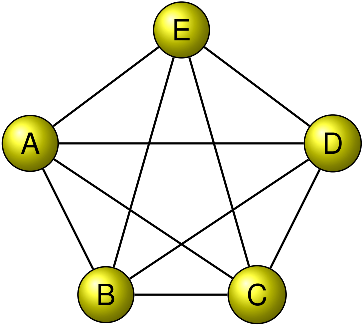


Grafy dwudzielne III



Grafy pełne (ang. complete graph)

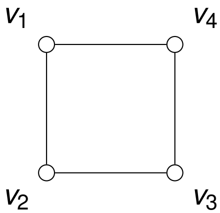
Graf pełny to graf nieskierowany, w którym każda para wierzchołków jest połączona krawędzią.
Graf pełny o n wierzchołkach oznacza się przez K_n (poniżej K_5)

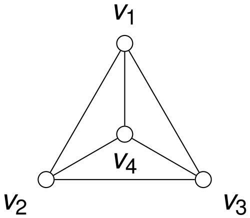


Graf pełny o n wierzchołkach posiada $\frac{n \cdot (n-1)}{2}$ krawędzi – przykładowo graf K_5 posiada 10 krawędzi.

Grafy planarny

- Graf **planarny** - graf nieskierowany, który można narysować na płaszczyźnie tak, by krawędzie grafu nie przecinały się ze sobą. Planarność ma duże zastosowanie w informatyce, m.in., w graficznej reprezentacji różnego rodzaju układów (np. scalonych, bramek, etc.).





Reprezentacja grafów w komputerze

- Macierz sąsiedztwa
- Listy sąsiedztwa
- Lista krawędzi

Macierz sąsiedztwa

Definicja

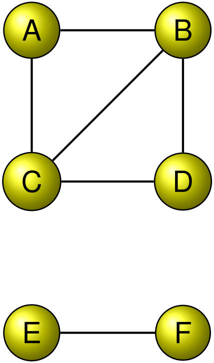
Dany jest graf $G = (V, E)$, gdzie $V = \{v_1, \ldots, v_n\}$ (tj. zakłada się, że wierzchołki są ponumerowane w pewien dowolny sposób) oraz $|V| = n$. **Macierz sąsiedztwa** grafu G to macierz $M(G) = (a_{ij})$ o wymiarze $n \times n$ taka, że

$$a_{ij} = \begin{cases} 1, & \text{jeśli } (v_i, v_j) \in E \\ 0, & \text{w przeciwnym razie} \end{cases}$$

Macierz sąsiedztwa - graf nieskierowany, przykład

- $V = \{A, B, C, D, E, F\}$
- $E = \{\{A, B\}, \{A, C\}, \{B, C\}, \{B, D\}, \{C, D\}, \{E, F\}\}$
- M(G) =

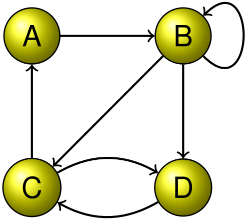
$$
\begin{array}{cc}
 & \begin{array}{cccccc} A & B & C & D & E & F \end{array} \\
\begin{array}{c} A \\ B \\ C \\ D \\ E \\ F \end{array} &
\begin{pmatrix}
0 & 1 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 0
\end{pmatrix}
\end{array}
$$



Macierz sąsiedztwa - Graf skierowny, przykład

- $V = \{A, B, C, D\}$
- $E = \{(A, B), (B, B), (B, C), (B, D), (C, A), (C, D), (D, C)\}$
- M(G) =

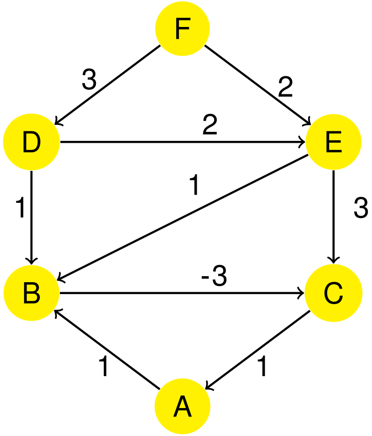
$$\begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$



Macierz sąsiedztwa - Etykietowany graf skierowany, przykład

M(G) =

$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \\ F \end{array} \begin{array}{c} \begin{array}{cccccc} A & B & C & D & E & F \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 2 & 0 \\ 0 & 1 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 2 & 0 \end{pmatrix} \end{array}$$



Macierz sąsiedztwa - własności

- Macierz sąsiedztwa dla grafów bez wag jest macierzą binarną.
- Macierz sąsiedztwa wymaga $\Theta(|V|^2)$ pamięci, niezależnie od liczby krawędzi w tym grafie.
- Dla grafów nieskierowanych macierz sąsiedztwa jest symetryczna, dlatego w pewnych zastosowaniach opłaca się pamiętać tylko elementy na i powyżej głównej przekątnej macierzy sąsiedztwa.

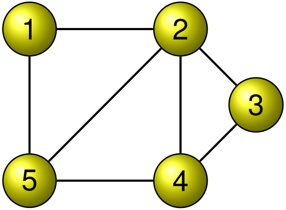
$$
\mathrm{M(G)} = \begin{array}{c c}
 & \begin{array}{cccccc} A & B & C & D & E & F \end{array} \\
\begin{array}{c} A \\ B \\ C \\ D \\ E \\ F \end{array} &
\begin{pmatrix}
0 & 1 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 0
\end{pmatrix}
\end{array}
$$

Listy sąsiedztwa

Definicja

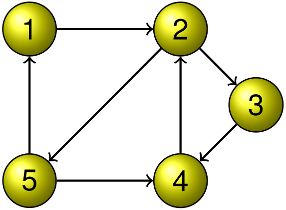
- W reprezentacji grafu $G = (V, E)$ za pomocą **list sąsiedztwa** dana jest tablica A zawierająca $|V|$ list, po jednej dla każdego wierzchołka z V.
- Dla każdego $u \in V$ elementami listy sąsiedztwa $A[u]$ są wszystkie wierzchołki v takie, że krawędź $(u, v) \in E$, tzn. w $A[u]$ przechowujemy zbiór wierzchołków połączonych krawędzią z u.

Listy sąsiedztwa - przykład



$1- >$[(2),(5)]
$2- >$[(1),(3),(4),(5)]
$3- >$[(2),(4)]
$4- >$[(2),(3),(5)]
$5- >$[(1),(2),(4)]

Listy sąsiedztwa - przykład



$1- >[(2)]$
$2- >[(3),(5)]$
$3- >[(4)]$
$4- >[(2)]$
$5- >[(1),(4)]$

Listy sąsiedztwa

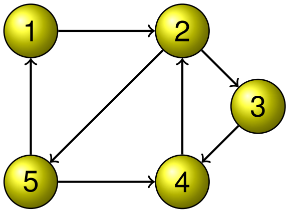
- Jeżeli G jest grafem skierowanym, to suma długości wszystkich list sąsiedztwa wynosi $|E|$, ponieważ krawędź postaci (u, v) jest reprezentowana przez wystąpienie v na liście $A[u]$.
- Jeżeli G jest grafem nieskierowanym, to suma długości wszystkich list sąsiedztwa wynosi $2 \cdot |E|$, ponieważ dla nieskierowanej krawędzi $\{u, v\}$ wierzchołek u występuje na liście sąsiedztwa v i odwrotnie, v występuje na liście u.
- Reprezentacja listowa do reprezentacji grafu wymaga $\Theta(|V| + |E|)$ pamięci.
- Reprezentacja listowa jest preferowana dla grafów **rzadkich** – tj. takich, dla których |E| jest dużo mniejsze niż $|V|^2$.
- Reprezentacja macierzy sąsiedztwa jest preferowana dla grafów **gęstych** – tj. takich, dla których |E| jest bliskie $|V|^2$.

Lista Krawędzi

Dany jest graf $G = (V, E)$, $|V| = n$, $|E| = m$

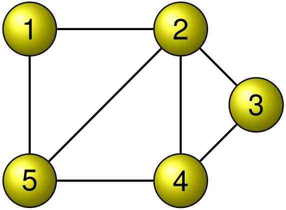
- Lista krawędzi to lista, na której przechowujemy wszystkie krawędzie występujące w grafie.
- Dla grafu skierowanego, może to być lista (tablica) zawierająca uporządkowane pary wierzchołków.
- Dla grafu nieskierowanego, może to być lista (tablica) zawierająca nieuporządkowane pary wierzchołków.
- Ponieważ każda krawędź zawiera tylko dwie liczby (tj. zakładamy, że wierzchołki sa ponumerowane), całkowita zajętość pamięci dla listy krawędzi jest rzędu $\Theta(|E|)$.

Lista krawędzi - graf skierowany



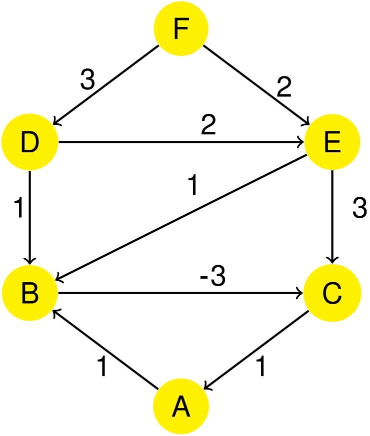
$L = [(1, 2), (2, 3), (3, 4), (4, 2), (2, 5), (5, 4), (5, 1)]$.

Lista krawędzi - graf nieskierowany



$L = [(1,2), (1,5), (2,3), (2,4),$
$(2,5), (3,4), (4,5)].$

Lista krawędzi - etykietowany graf skierowany, przykład



$L =$
$[(A, B, 1), (B, C, -3), (C, A, 1),$
$(D, B, 1), (E, B, 1), (E, C, 3),$
$(D, E, 2), (F, D, 3), (F, E, 2)].$