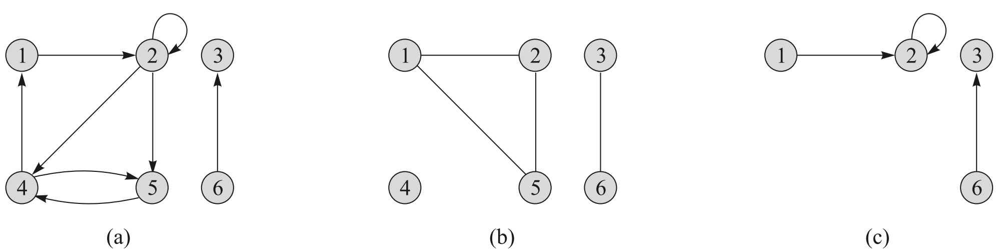
B.4 Grafy

W tym dodatku przedstawimy dwa rodzaje grafów: skierowane (zorientowane) i nieskierowane (niezorientowane). Pewne definicje, z którymi spotykamy się w literaturze matematycznej, mogą się różnić od podanych tutaj, lecz w większości przypadków różnice te są niewielkie. W podrozdziale 20.1 pokazujemy, jak można reprezentować grafy w pamięci komputera.

Graf skierowany (lub ***digraf***) (ang. *directed graph*) G jest opisany jako para (V, E), gdzie V jest zbiorem skończonym, a E jest relacją dwuargumentową w V. Zbiór V jest nazywany ***zbiorem wierzchołków*** G, a jego elementy są nazywane ***wierzchołkami***. Zbiór E jest nazywany ***zbiorem krawędzi*** G, a jego elementy nazywamy ***krawędziami***. Na rysunku B.2(a) jest pokazany graf skierowany o zbiorze wierzchołków $\{1, 2, 3, 4, 5, 6\}$. Wierzchołki są przedstawione jako kółka, a krawędzie jako strzałki. Zauważmy, że możliwe jest istnienie ***pętli*** od danego wierzchołka do niego samego.

W ***grafie nieskierowanym*** $G = (V, E)$ zbiór krawędzi E to zbiór *nieuporządkowanych* par wierzchołków. Oznacza to, że krawędź jest zbiorem $\{u, v\}$, gdzie $u, v \in V$ i $u \neq v$. Do oznaczenia krawędzi będziemy używać notacji (u, v) zamiast zapisu $\{u, v\}$ przeznaczonego dla zbiorów; zapisy (u, v) i (v, u) oznaczają tę samą krawędź. W grafie nieskierowanym nie mogą występować pętle, więc każda krawędź zawiera dokładnie dwa różne wierzchołki. Na rysunku B.2(b) jest pokazany graf nieskierowany o zbiorze wierzchołków $\{1, 2, 3, 4, 5, 6\}$.

Wiele definicji dotyczących grafów skierowanych i nieskierowanych jest takich samych, chociaż niektóre terminy mają nieco inne znaczenie w tych dwóch kontekstach. Jeśli (u, v) jest krawędzią grafu skierowanego $G = (V, E)$, to mówimy, że krawędź (u, v) jest ***wychodząca*** z wierzchołka u i jest ***wchodząca*** do wierzchołka v. Na przykład krawędzie wychodzące z wierzchołka 2 na rys. B.2(a) to $(2, 2)$, $(2, 4)$ i $(2, 5)$. Krawędzie wchodzące do wierzchołka 2 to $(1, 2)$ i $(2, 2)$. Jeśli (u, v) jest krawędzią grafu nieskierowanego $G = (V, E)$, to mówimy, że (u, v) jest ***incydentna*** z wierzchołkami u i v. Na rysunku B.2(b) krawędzie incydentne z wierzchołkiem 2 to $(1, 2)$ i $(2, 5)$.



Rysunek B.2 Grafy skierowane i nieskierowane. (a) Graf skierowany $G = (V, E)$, gdzie $V = \{1, 2, 3, 4, 5, 6\}$ i $E = \{(1, 2), (2, 2), (2, 4), (2, 5), (4, 1), (4, 5), (5, 4), (6, 3)\}$. Krawędź $(2, 2)$ jest pętlą. (b) Graf nieskierowany $G = (V, E)$, gdzie $V = \{1, 2, 3, 4, 5, 6\}$ i $E = \{(1, 2), (1, 5), (2, 5), (3, 6)\}$. Wierzchołek 4 jest izolowany. (c) Podgraf grafu z części (a) indukowany przez zbiór wierzchołków $\{1, 2, 3, 6\}$

Jeśli (u, v) jest krawędzią grafu $G = (V, E)$, to mówimy, że wierzchołek v jest ***sąsiedni*** względem wierzchołka u. Kiedy graf jest nieskierowany, relacja sąsiedztwa jest symetryczna. Kiedy graf jest skierowany, relacja sąsiedztwa nie musi być symetryczna. Jeśli w grafie skierowanym v jest wierzchołkiem sąsiednim względem u, to czasem zapisujemy to jako $u \to v$. Na rysunku B.2(a) i (b) wierzchołek 2 jest sąsiedni względem wierzchołka 1, ponieważ krawędź $(1, 2)$ należy do obu grafów. Wierzchołek 1 *nie* jest sąsiedni względem wierzchołka 2 na rys. B.2(a), ponieważ krawędź $(2, 1)$ nie należy do grafu.

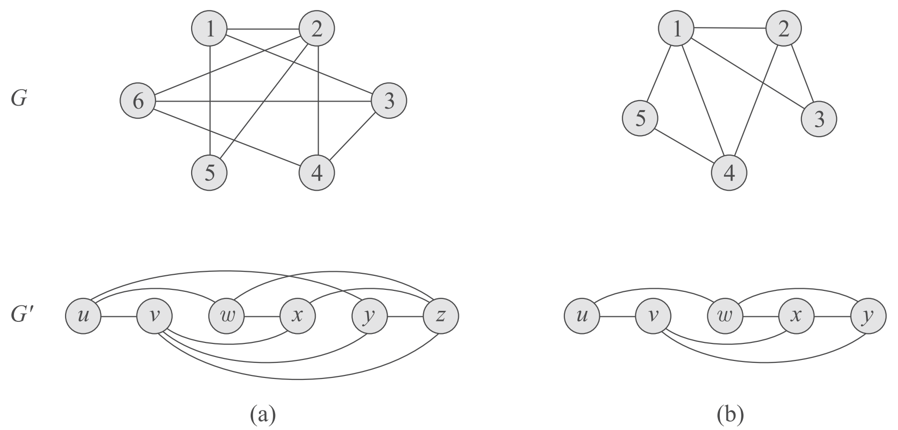
Stopniem wierzchołka w grafie nieskierowanym jest liczba incydentnych z nim krawędzi. Na przykład wierzchołek 2 na rys. B.2(b) ma stopień 2. Wierzchołek, którego stopień wynosi 0, taki jak wierzchołek 4 na rys. B2(b), jest nazywany **izolowanym**. W grafie skierowanym ***stopień wyjściowy*** wierzchołka jest liczbą krawędzi wychodzących z niego, a ***stopień wejściowy*** wierzchołka jest liczbą krawędzi do niego wchodzących. ***Stopniem*** wierzchołka w grafie skierowanym jest liczba będąca sumą jego stopni: wejściowego i wyjściowego. Wierzchołek 2 na rys. B.2(a) ma stopień wejściowy 2, stopień wyjściowy 3 i stopień 5.

Ścieżka (droga) ***długości*** k z wierzchołka u do wierzchołka u' w grafie $G = (V, E)$ jest ciągiem wierzchołków $\langle v_0, v_1, v_2, \ldots, v_k \rangle$ takich, że $u = v_0$, $u' = v_k$ i $(v_{i-1}, v_i) \in E$ dla $i = 1, 2, \ldots, k$. Długość ścieżki jest liczbą jej krawędzi, która jest o 1 mniejsza niż liczba wierzchołków na ścieżce. Ścieżka ***zawiera*** wierzchołki $v_0, v_1, \ldots, v_k$ i krawędzie $(v_0, v_1), (v_1, v_2), \ldots, (v_{k-1}, v_k)$. (Zawsze istnieje ścieżka o długości 0 z u do u). Jeśli istnieje ścieżka p z u do u', to mówimy, że u' jest ***osiągalny*** z u po ścieżce p, co zapisujemy: $u \overset{p}{\leadsto} u'$. Ścieżkę nazywamy ***prostą***[4], jeśli wszystkie jej wierzchołki są różne. Na rysunku B.2(a) ścieżka $\langle 1, 2, 5, 4 \rangle$ jest ścieżką prostą o długości 3. Ścieżka $\langle 2, 5, 4, 5 \rangle$ nie jest ścieżką prostą. ***Podścieżka*** ścieżki $p = \langle v_0, v_1, \ldots, v_k \rangle$ jest ciągiem jej kolejnych wierzchołków. To znaczy, że dla dowolnych $0 \le i \le j \le k$ podciąg wierzchołków $\langle v_i, v_{i+1}, \ldots, v_j \rangle$ jest podścieżką ścieżki p.

W grafie skierowanym ścieżka $\langle v_0, v_1, \ldots, v_k \rangle$ tworzy ***cykl***, jeśli $v_0 = v_k$, a ścieżka zawiera co najmniej jedną krawędź. Cykl nazywamy ***prostym***, jeśli dodatkowo wierzchołki $v_1, v_2, \ldots, v_k$ są różne. Cykl składający się z k wierzchołków ma ***długość*** k. Pętla jest cyklem o długości 1. Dwie ścieżki $\langle v_0, v_1, v_2, \ldots, v_{k-1}, v_0 \rangle$ i $\langle v'_0, v'_1, v'_2, \ldots, v'_{k-1}, v'_0 \rangle$ tworzą ten sam cykl, jeśli istnieje liczba całkowita j taka, że $v'_i = v_{(i+j) \bmod k}$ dla $i = 0, 1, \ldots, k-1$. Na rysunku B.2(a) ścieżka $\langle 1, 2, 4, 1 \rangle$ tworzy ten sam cykl co ścieżki $\langle 2, 4, 1, 2 \rangle$ i $\langle 4, 1, 2, 4 \rangle$. Jest to cykl prosty, natomiast cykl $\langle 1, 2, 4, 5, 4, 1 \rangle$ nie jest prosty. Cykl $\langle 2, 2 \rangle$ złożony z krawędzi $(2, 2)$ jest pętlą. Graf skierowany niezawierający pętli nazywamy ***prostym***. W grafie nieskierowanym ścieżka $\langle v_0, v_1, \ldots, v_k \rangle$ tworzy ***cykl***, jeśli $k > 0$, $v_0 = v_k$ i wszystkie krawędzie na ścieżce są różne. Cykl jest ***prosty***, jeśli wierzchołki $v_1, v_2, \ldots, v_k$ są różne. Na przykład na rys. B.2(b) ścieżka $\langle 1, 2, 5, 1 \rangle$ tworzy cykl prosty. Graf niezawierający cykli prostych nazywamy ***acyklicznym***.

Graf nieskierowany jest ***spójny*** (ang. *connected*), jeśli każdy wierzchołek jest osiągalny ze wszystkich innych wierzchołków. ***Spójne składowe*** (ang. *connected components*) grafu nieskierowanego to klasy abstrakcji określonej w zbiorze wierzchołków relacji „jest osiągalny z". Graf na rys. B.2(b) ma trzy spójne składowe: $\{1, 2, 5\}$, $\{3, 6\}$ i $\{4\}$. Każdy wierzchołek ze zbioru

[4] Niektórzy autorzy o ścieżce mówią „marszruta", a słowo „ścieżka" rezerwują dla ścieżki prostej. W tej książce terminów „ścieżka" i „ścieżka prosta" używamy zgodnie z podanymi tu definicjami.



Rysunek B.3 (a) Para izomorficznych grafów. Wierzchołki górnego grafu można odwzorować na wierzchołki dolnego grafu za pomocą funkcji $f(1) = u$, $f(2) = v$, $f(3) = w$, $f(4) = x$, $f(5) = y$, $f(6) = z$. (b) Dwa nieizomorficzne grafy: górny graf ma wierzchołek stopnia 4, natomiast dolny takiego nie ma

$\{1, 2, 5\}$ jest osiągalny z każdego innego wierzchołka w tym zbiorze. Graf nieskierowany jest spójny, jeśli ma dokładnie jedną spójną składową. Krawędzie spójnej składowej to te, które są incydentne tylko z wierzchołkami tej składowej; inaczej mówiąc, krawędź (u, v) należy do spójnej składowej tylko wtedy, gdy zarówno u, jak i v są wierzchołkami tej składowej.

Graf skierowany jest ***silnie spójny*** (ang. *strongly connected*), jeśli każde dwa wierzchołki są osiągalne jeden z drugiego. ***Silnie spójne składowe*** grafu skierowanego są klasami abstrakcji relacji „są wzajemnie osiągalne” w zbiorze wierzchołków. Graf skierowany jest silnie spójny, jeśli ma tylko jedną silnie spójną składową. Graf na rys. B.2(a) ma trzy silnie spójne składowe: $\{1, 2, 4, 5\}$, $\{3\}$ i $\{6\}$. Wszystkie pary wierzchołków w zbiorze $\{1, 2, 4, 5\}$ są wzajemnie osiągalne. Wierzchołki $\{3, 6\}$ nie tworzą silnie spójnej składowej, ponieważ wierzchołek 6 nie jest osiągalny z wierzchołka 3.

Dwa grafy $G = (V, E)$ i $G' = (V', E')$ są ***izomorficzne***, jeśli istnieje bijekcja $f : V \to V'$ taka, że $(u, v) \in E$ wtedy i tylko wtedy, gdy $(f(u), f(v)) \in E'$. Innymi słowy, możemy przenumerować wierzchołki G tak, aby były wierzchołkami G', z zachowaniem odpowiadających sobie krawędzi w G i G'. Na rysunku B.3(a) widać parę izomorficznych grafów G i G' ze zbiorami wierzchołków, odpowiednio, $V = \{1, 2, 3, 4, 5, 6\}$ i $V' = \{u, v, w, x, y, z\}$. Odwzorowanie z V w V' dane przez $f(1) = u$, $f(2) = v$, $f(3) = w$, $f(4) = x$, $f(5) = y$, $f(6) = z$ jest wymaganą bijekcją. Grafy na rys. B.3(b) nie są izomorficzne. Mimo że obydwa grafy mają po 5 wierzchołków i po 7 krawędzi, tylko górny graf ma wierzchołek stopnia 4.

Mówimy, że graf $G' = (V', E')$ jest ***podgrafem*** grafu $G = (V, E)$, jeśli $V' \subseteq V$ i $E' \subseteq E$. Dla danego zbioru $V' \subseteq V$ podgrafem G ***indukowanym*** przez V' jest graf $G' = (V', E')$, gdzie

$$E' = \{(u, v) \in E : u, v \in V'\} .$$

Podgraf indukowany przez zbiór wierzchołków $\{1, 2, 3, 6\}$ z rys. B.2(a) jest pokazany na rys B.2(c), a jego zbiór krawędzi to $\{(1, 2), (2, 2), (6, 3)\}$.

Dla danego grafu nieskierowanego $G = (V, E)$ ***wersją skierowaną*** grafu G nazywamy graf skierowany $G' = (V, E')$, w którym $(u, v) \in E'$ wtedy i tylko wtedy, gdy $(u, v) \in E$. Oznacza to, że każda krawędź nieskierowana (u, v) grafu G zostaje w wersji skierowanej zastąpiona przez dwie krawędzie skierowane (u, v) i (v, u). Dla grafu skierowanego $G = (V, E)$ jego ***wersją nieskierowaną*** nazywamy graf nieskierowany $G' = (V, E')$, w którym $(u, v) \in E'$ wtedy i tylko wtedy, gdy $u \neq v$ i E zawiera co najmniej jedną z krawędzi (u, v) lub (v, u). Oznacza to, że wersja nieskierowana zawiera krawędzie grafu G „z usuniętym skierowaniem" oraz usuniętymi pętlami. (Ponieważ (u, v) i (v, u) są w grafie nieskierowanym tą samą krawędzią, wersja nieskierowana grafu skierowanego zawiera ją tylko raz, nawet jeśli graf skierowany zawiera obie krawędzie (u, v) i (v, u)). W grafie skierowanym $G = (V, E)$ ***sąsiadem*** wierzchołka u jest dowolny wierzchołek sąsiedni względem u w nieskierowanej wersji G. To znaczy, że v jest sąsiadem u, jeśli $(u, v) \in E$ lub $(v, u) \in E$. W grafie nieskierowanym wierzchołki u i v są sąsiadami, jeśli są sąsiednie.

Niektóre rodzaje grafów mają nadane specjalne nazwy. ***Graf pełny*** (ang. *complete graph*) jest to graf nieskierowany, w którym każda para wierzchołków to wierzchołki sąsiednie. ***Graf dwudzielny*** (ang. *bipartite graph*) jest to graf nieskierowany $G = (V, E)$, w którym zbiór V można podzielić na dwa zbiory V_1 i V_2 takie, że jeśli $(u, v) \in E$, to $u \in V_1$ i $v \in V_2$ lub $u \in V_2$ i $v \in V_1$. Oznacza to, że każda krawędź ma końce w dwóch różnych zbiorach V_1 i V_2. Acykliczny graf nieskierowany nazywa się ***lasem***, a spójny, acykliczny graf nieskierowany nazywa się ***(wolnym) drzewem*** (patrz podrozdz. B.5). Acykliczny graf skierowany określamy czasami w skrócie jako ***dag*** – od pierwszych liter nazwy angielskiej *directed acyclic graph*.

Istnieją dwie struktury podobne do grafów, z którymi możemy się czasem spotkać. ***Multigraf*** jest podobny do grafu nieskierowanego, lecz może mieć zarówno wielokrotne krawędzie między wierzchołkami (takie jak dwie różne krawędzie (u, v) i (u, v)), jak i pętle. ***Hipergraf*** różni się od grafu nieskierowanego tym, że każda ***hiperkrawędź***, zamiast łączyć dwa wierzchołki, łączy pewien dowolny podzbiór wierzchołków. Wiele algorytmów zaprojektowanych dla zwykłych grafów skierowanych i nieskierowanych można zaadaptować do tego typu struktur grafopodobnych.

Ściągnięcie (ang. *contraction*) grafu nieskierowanego $G = (V, E)$ wzdłuż krawędzi $e = (u, v)$ to graf $G' = (V', E')$, gdzie $V' = V - \{u, v\} \cup \{x\}$, przy czym x jest nowym wierzchołkiem. Zbiór krawędzi E' tworzymy z E, usuwając krawędź (u, v) oraz dla każdego wierzchołka w sąsiedniego względem u lub v, usuwając krawędzie (u, w) i (v, w), o ile tylko należą do E, i dodając nową krawędź (x, w). W rezultacie u i v są „ściągnięte" w jeden wierzchołek.

Zadania

B.4-1

Uczestnicy wydziałowego przyjęcia podają sobie ręce na powitanie, przy czym każdy z nich zapamiętuje, ile razy podał komuś rękę. Na koniec przyjęcia dziekan sumuje, ile razy każdy z pracowników podał komuś rękę. Wykaż, że wynik jest parzysty, dowodząc ***lematu o podawaniu***

rąk: jeśli $G = (V, E)$ jest grafem nieskierowanym, to

$$\sum_{v \in V} \text{stopień}(v) = 2\,|E|\,.$$

B.4-2
Pokaż, że jeśli graf skierowany lub nieskierowany zawiera ścieżkę między dwoma wierzchołkami u i v, to zawiera ścieżkę prostą między tymi wierzchołkami. Pokaż, że jeśli graf skierowany zawiera cykl, to zawiera cykl prosty.

B.4-3
Pokaż, że dowolny spójny graf nieskierowany $G = (V, E)$ spełnia zależność $|E| \geq |V| - 1$.

B.4-4
Sprawdź, że w grafie nieskierowanym relacja „jest osiągalny z" jest relacją równoważności w zbiorze wierzchołków grafu. Która z trzech własności relacji równoważności zachodzi dla tej relacji w grafach skierowanych?

B.4-5
Co jest wersją nieskierowaną grafu skierowanego z rys. B.2(a)? Co jest wersją skierowaną grafu nieskierowanego z rys. B.2(b)?

★ ***B.4-6***
Pokaż, że hipergraf można reprezentować jako graf dwudzielny, jeśli przyjmiemy, że relacja incydencji dla hipergrafu odpowiada relacji sąsiedztwa w grafie dwudzielnym. (*Wskazówka:* Niech jeden zbiór wierzchołków w grafie dwudzielnym odpowiada wierzchołkom hipergrafu, a drugi zbiór wierzchołków – hiperkrawędziom).

B.5 Drzewa

Podobnie jak w przypadku grafów, istnieje wiele pokrewnych, ale nieco różniących się od siebie interpretacji pojęcia drzewa. W tej części książki podajemy definicje i matematyczne własności kilku typów drzew. W podrozdziałach 10.3 i 20.1 pokazujemy, jak można reprezentować drzewa w pamięci komputera.

B.5.1 Drzewa wolne

Zdefiniowane już w dodatku B.4 ***drzewo wolne*** jest spójnym, acyklicznym grafem nieskierowanym. Często pomijamy określenie „wolne", mówiąc po prostu, że dany graf jest drzewem. Graf nieskierowany, który jest acykliczny, ale niekoniecznie spójny, nazywamy ***lasem***. Wiele z algorytmów dla drzew działa również dla lasów. Na rysunku B.4(a) widać drzewo wolne, a na rys. B.4(b) – las. Las na rys. B.4(b) nie jest drzewem, ponieważ jest niespójny. Graf na rys. B.4(c) nie jest ani drzewem, ani lasem, ponieważ zawiera cykl.