Definicje i konwencje

Zebrane definicje i konwencje z wykładów, niekoniecznie w kolejności omawiania (patrz też: stosowne wykłady i rozdziały skryptu).
Część „Adnotacje” stanowią uwagi „na boku”.
Uwaga 1: Definicje 8, 9, 16 i 17 dotyczą wykładów, które jeszcze się nie odbyły (stan sprzed wykładu 8) i nie obowiązują na kolokwium.

1 Notacja asymptotyczna

Definicja 1. *Niech $f\colon \mathbb{N} \to \mathbb{R}_{\geqslant 0}$. Definiujemy następujące rodziny funkcji:*

(a) $O(f) = \{g \in \mathbb{R}_{\geqslant 0}^{\mathbb{N}} : \exists M \in \mathbb{R}_+ \; \exists n_0 \in \mathbb{N} \; \forall n > n_0 \; g(n) < Mf(n)\}$.

(b) $\Omega(f) = \{g \in \mathbb{R}_{\geqslant 0}^{\mathbb{N}} : \exists m \in \mathbb{R}_+ \; \exists n_0 \in \mathbb{N} \; \forall n > n_0 \; mf(n) < g(n)\}$.

(c) $\Theta(f) = O(f) \cap \Omega(f)$.

$O(f)$ to zbiór funkcji ograniczonych asymptotycznie z góry przez (pewne przeskalowanie) funkcji f. $\Omega(f)$ to zbiór funkcji ograniczonych asymptotycznie z dołu przez (pewne przeskalowanie) funkcji f.

Dla funkcji $f, g \in \mathbb{R}_{\geqslant 0}^{\mathbb{N}}$ możemy napisać $g \in O(f)$ etc. Możemy też explicite napisać argumenty funkcji: $g(n) \in O(f(n))$ etc.[1]

2 Grafy

Definicja 2. ***Graf (skierowany)*** *to para (V, E) gdzie:*

*(a) V jest zbiorem, którego elementy nazywamy **wierzchołkami** (lub **węzłami**),*

*(b) E jest relacją na V. Parę $(u, v) \in E$ nazywamy **krawędzią** z u do v.*

Od teraz $G = (V, E)$ jest grafem.

Uwaga 3. *Konwencja na wykładzie: V jest zawsze **skończony** (w konsekwencji zbiór krawędzi też jest skończony).*

Definicja 4. *Graf G nazywamy **nieskierowanym**[2], gdy relacja E jest symetryczna, czyli*

$$\forall u, v \in V \; \Big((u, v) \in E \iff (v, u) \in E\Big).$$

Rysując grafy nieskierowane, zamiast rysować parę krawędzi pomiędzy połączonymi wierzchołkami, możemy rysować jedną kreskę pomiędzy parą połączonych węzłów i nie zaznaczać jej kierunku. Kreska ta oznacza parę krawędzi.

[1] por. Uwaga 18.
[2] por. Uwaga 19.

Definicja 5. *(a) Jeśli $(u, v) \in E$, to v nazywamy* ***sąsiadem*** *u.*

(b) ***Ścieżka (skierowana)***[3] *z wierzchołka u do v to ciąg krawędzi $(u_1, v_1), \ldots, (u_n, v_n)$ takich, że $(\forall i < n)\ v_i = u_{i+1}, u = u_1, v_n = v$.*

(c) ***Ścieżka nieskierowana*** *z wierzchołka u do v to ciąg krawędzi $(u_1, v_1), \ldots, (u_n, v_n)$ takich, że $(\forall i < n)\ \{u_i, v_i\} \cap \{u_{i+1}, v_{i+1}\} \neq \emptyset$ (tzn. każde kolejne dwie krawędzie mają co najmniej jeden wspólny wierzchołek), oraz $u \in \{u_1, v_1\}, v \in \{u_n, v_n\}$.*

(d) ***Odległość***[4] *między wierzchołkami u i v to długość najkrótszej ścieżki z u do v, jeśli taka istnieje. W przeciwnym wypadku jako odległość kładziemy ∞.*

Przyjmujemy przy tym, że każdy wierzchołek v jest połączony sam ze sobą pustą ścieżką. Wtedy też odległość z v do v wynosi 0. W definicji odległości możemy zastąpić ścieżkę skierowaną przez ścieżkę nieskierowaną: wtedy odległość wierzchołków jest funkcją symetryczną.

Definicja 6. ***Cyklem*** *nazywamy niepustą ścieżkę z wierzchołka v do v.*

3 Grafy ważone

Definicja 7. ***Graf ważony***[5] *to trójka (V, E, f) taka, że para (V, E) jest grafem, a $f \colon E \to \mathbb{R}$ jest funkcją przypisującą każdej krawędzi jej* ***wagę****.*

Wagi nie muszą być dodatnie. W grafie ważonym mogą istnieć krawędzie $(u, v), (v, u) \in E$ takie, że $f((u, v)) \neq f((v, u))$[6].

W grafach ważonych inaczej definiujemy długość ścieżki:

Definicja 8.

Niech $G = (V, E, f)$ będzie grafem ważonym i niech $S = (e_1, e_2, \ldots, e_n)$ będzie (skierowaną) ścieżką. ***Długością*** *ścieżki S nazywamy sumę wag jej krawędzi: $\sum_{i=1}^{n} f(e_i)$.*

Zwracamy uwagę, że w grafie ważonym może istnieć ścieżka z u do v, lecz może nie istnieć ścieżka najkrótsza (gdy wystąpi cykl ujemnej długości). Dlatego należy ostrożnie stosować poniższą definicję:

Definicja 9.

Niech $G = (V, E, f)$ będzie grafem ważonym i $u, v \in V$. ***Odległość*** *z u do v to długość najkrótszej ścieżki z u do v lub ∞, gdy taka nie istnieje.*

Ważony graf skierowany, w którym wszystkie krawędzie mają wagę dodatnią nazywamy **dodatnio ważonym**. W takim grafie istnienie ścieżki między wierzchołkami implikuje istnienie najkrótszej ścieżki. W takim przypadku stosuje się algorytm Dijkstry.

[3] por. Definicja 10(e).
[4] por. Definicja 9
[5] por. Uwaga 20.
[6] por. Uwaga 21.

4 Drzewa

Drzewo to szczególny przypadek grafu. Rozważmy skończony zbiór węzłów V z relacją S. Opisując drzewa używamy specyficznego nazewnictwa:

Definicja 10. *(a) Dla $(x, y) \in S$ mówimy, że y jest* ***dzieckiem*** *x, a x jest* ***rodzicem*** *y.*

(b) ***Korzeń*** *to węzeł, który nie ma rodzica.*

(c) ***Liść*** *to węzeł, który nie ma dziecka.*

(d) Węzeł jest ***wewnętrzny****, gdy nie jest liściem.*

(e) ***Ścieżka*** *z węzła x do węzła y to ciąg węzłów $x = x_1, x_2, \ldots x_n = y$ taki, że $x_1 S x_2 S \ldots S x_n$.*

(f) Jeśli istnieje ścieżka z x do y, to x nazywamy ***przodkiem*** *y, a y* ***potomkiem*** *x.*

Definicja 11. *Parę $T = (V, S)$ nazywamy drzewem, gdy spełnia następujące dwa warunki:*

(a) W zbiorze V istnieje dokładnie jeden korzeń r.

(b) Dla każdego węzła $x \in V$ istnieje dokładnie jedna ścieżka z r do x.

Uwaga 12. *(a) Para (V, S) jest w szczególności grafem. Relacja S nie jest symetryczna (rodzic dziecka nie jest jego dzieckiem). Jest to zatem graf skierowany (por. directed rooted tree).*

(b) W drzewie jest dokładnie jeden korzeń, czyli w tej definicji drzewo jest niepuste. Pasuje to do implementacji drzewa binarnego z wykładu (`BinaryTree`*).*

Rysując drzewo, możemy pominąć kierunki krawędzi z rodzica do dziecka umieszczając dziecko poniżej rodzica. W szczególności, najwyżej rysujemy korzeń drzewa.

Od teraz para $T = (V, S)$ jest drzewem. Kilka dalszych terminów:

Definicja 13. *(a)* ***Gałąź*** *to każda maksymalna ścieżka w drzewie. Każda gałąź zaczyna się w korzeniu i kończy w liściu.*

(b) ***Wysokość*** *drzewa to maksymalna długość gałęzi.*

W szczególności, drzewo zawierające tylko jeden wierzchołek (korzeń, jednocześnie liść) ma wysokość 1.

Definicja 14. *Dla $x \in V$, zbiór $V_x = \{y \in V : y \text{ jest potomkiem } x\}$ z relacją S obciętą do V_x nazywamy* ***poddrzewem*** *T.*

V_x z obciętą relacją jest drzewem o korzeniu x.

Definicja 15. *Drzewo $T = (V, S)$ nazywamy* ***drzewem binarnym****, gdy każdy węzeł ma (co najwyżej) dwa dziecka[7]. Wyróżniamy wtedy dziecko* ***lewe*** *i* ***prawe****, uzupełniając (V, S) o relacje L i R na V takie, że $x \; S \; y \iff x \; L \; y \veebar x \; R \; y$ i dla każdego x istnieje co najwyżej jeden y taki, że $x \; L \; y$ (oraz analogicznie dla R).*

[7] Liczba mnoga rzeczownika nieżywotnego - por. „królowie” na tronach, lecz „króle” na szachownicy.

Węzeł y jest lewym dzieckiem x dokładnie wtedy, gdy $xSy \wedge xLy$, a prawym dokładnie wtedy, gdy $xSy \wedge xRy$.

Gdy r jest korzeniem T, a x jego lewym (względnie prawym) dzieckiem, poddrzewo V_x nazywamy lewym (względnie prawym) poddrzewem T. Używamy tych terminów również w odniesieniu do węzłów: dla dowolnego $t \in V$ i jego lewego dziecka x, poddrzewo V_x nazywamy lewym poddrzewem t, analogicznie dla poddrzew prawych (w szczególności V_x jest tutaj lewym/prawym poddrzewem drzewa V_t).

W drzewach binarnych mamy trzy naturalne kolejności odwiedzania węzłów, zadane w następujący, rekurencyjny sposób (gdzie poddrzew nie odwiedzamy, jeśli nie istnieją):

(a) Kolejność *preorder*: odwiedzamy korzeń drzewa; następnie odwiedzamy lewe poddrzewo w kolejności preorder; następnie odwiedzamy prawe poddrzewo w kolejności preorder.

(b) Kolejność *inorder*: odwiedzamy lewe poddrzewo w kolejności inorder; następnie odwiedzamy korzeń drzewa; następnie odwiedzamy prawe poddrzewo w kolejności inorder.

(c) Kolejność *postorder*: odwiedzamy lewe poddrzewo w kolejności postorder; następnie odwiedzamy prawe poddrzewo w kolejności postorder; następnie odwiedzamy korzeń drzewa.

Odwiedzanie wierzchołków w kolejności pre-, in-, postorder zadaje nam odpowieni porządek na węzłach drzewa. Nietrudno zauważyć, że w porządku preorder korzeń drzewa jest pierwszy, a w porządku postorder ostatni.

Definicje poniżej służą tylko do opisu kopców binarnych. Pojawiają się dla kompletności:

Definicja 16. *W drzewie T każdemu węzłowi w przypisujemy rekurencyjnie liczbę naturalną, zwaną* ***poziomem****:*

(a) Korzeń ma poziom 0.

(b) Jeśli węzeł ma poziom n, to każde jego dziecko ma poziom $n+1$.

W drzewie wysokości h, węzły mają poziomy $0, 1, \ldots, h-1$ (por. `range(h)`). W drzewie binarnym na poziomie n znajduje się co najwyżej 2^n węzłów. Z kolei na poziomach mniejszych niż n znajduje się w sumie co najwyżej $\sum_{i<n} 2^i = 2^n - 1$ węzłów.

Definicja 17. *Drzewo binarne T jest* ***zupełne****, gdy istnieje $n \in \mathbb{N}$ takie, że:*

(a) Na poziomach mniejszych niż n znajduje się dokładnie $2^n - 1$ węzłów.

(b) Wszystkie pozostałe węzły znajdują się na poziomie n i są umieszczone możliwie najbardziej na lewo.

Ściślej, warunek (b) można opisać następująco: odwiedzając węzły na poziomie $n-1$ w kolejności inorder, odwiedzamy: najpierw węzły, które mają dwa dziecka; potem co najwyżej jeden węzeł, który ma tylko lewe dziecko; a potem liście.

5 Adnotacje

Uwaga 18. *Często notacji „wielkiego O” nie wprowadza się przez zdefiniowanie $O(f(n))$ jako rodziny funkcji, a przez czysto formalny zapis „$f(n) = O(g(n))$” (nadużywający symbolu równości) oznaczający asymptotyczne ograniczenie f przez (przeskalowanie) g.*

Uwaga 19. *W innej formalizacji grafów nieskierowanych, E jest zbiorem nieuporządkowanych par postaci $\{u, v\} \subseteq V$. Czasami zakładamy wtedy, że $u \neq v$ (wykluczamy nieskierowaną krawędź z u do u).*

Uwaga 20. *W innej formalizacji grafów ważonych, $G = (V, E)$, gdzie E jest funkcją częściową z $V \times V$ w $\mathbb{R}$, to znaczy: dziedziną E jest $X \subseteq V \times V$ i $E \colon X \to \mathbb{R}$. Wtedy mówimy, że $u, v \in V$ są połączone krawędzią wtedy, gdy (u, v) leży w dziedzinie E. Wagą tej krawędzi jest $E(u, v)$.*

Uwaga 21. *Ogólnie, w grafie ważonym mogą istnieć krawędzie $(u, v), (v, u) \in E$ takie, że $f((u, v)) \neq f((v, u))$. Graf ważony (V, E, f) nazywamy **nieskierowanym**, gdy (V, E) jest grafem nieskierowanym i dla wszystkich $(u, v) \in E$ zachodzi $f((u, v)) = f((v, u))$. W szczególności może zajść sytuacja, w której (V, E, f) nie jest nieskierowanym grafem ważonym, ale (V, E) jest (nieważonym) grafem nieskierowanym.*

Uwaga 22. *Inne, częste konwencje definiowania obiektów:*

(a) *W literaturze ścieżkę w drzewie bardzo często definiuje się tak, jak dla ogólnie dla grafów (jako ciąg krawędzi, a nie węzłów). Wysokość drzewa składającego się z pojedynczego węzła jest wtedy równa 0.*

(b) *Drzewa są też definiowane jako grafy nieskierowane bez cykli. Tę definicję stosuje się głównie, gdy w drzewie nie wyróżniamy korzenia. Wyróżnienie korzenia zadaje naturalnie kierunki krawędzi relacji rodzic-dziecko (por. (directed) rooted tree).*