Artur Piękosz

Matematyka dyskretna

Kraków, 5 marca 2009

Rozdział 1

Podstawowe pojęcia matematyki

1.1. Co to jest matematyka dyskretna?

matematyka dyskretna	matematyka ciągła
zbiory skończone lub przeliczalne, nie ma przejścia do granicy	zbiory nieprzeliczalne, przejście do granicy, pochodne, całki, ...
algebra, logika, teoria liczb, kryptografia, teoria grafów, arytmetyka, teoria relacji, prawdopodobieństwo dyskretne	analiza matematyczna, równania różniczkowe, równania całkowe, topologia, prawdopodobieństwo ciągłe

1.2. Notacja logiczna

$\wedge$	koniunkcja	$\Longrightarrow$, $\rightarrow$	implikacja
$\vee$	alternatywa	$\Longleftrightarrow$, $\leftrightarrow$	równoważność
$\neg$, $\sim$	negacja	$\top$, 1	zdanie prawdziwe
$\oplus$	alternatywa wykluczająca	$\bot$, 0	zdanie fałszywe

$\forall$	kwantyfikator uniwersalny
$\exists$	kwantyfikator egzystencjalny
$\exists!$	„istnieje dokładnie jeden”

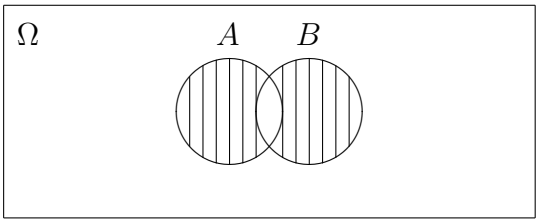
1.3. Notacja teoriomnogościowa

$\cup$	suma mnogościowa	$\cap$	część wspólna, przecięcie
$\setminus$	różnica	$\div$, $\triangle$	różnica symetryczna
$'$ C	dopełnienie	$\times$	iloczyn kartezjański
$\subset$	relacja zawierania	$\subsetneq$	relacja zawierania właściwego
$=$	relacja równości	(a, b)	para uporządkowana o pierwszym elemencie a i drugim elemencie b.

$\{x \in Z \mid W(x)\}$, $\{x_1, \ldots, x_n\}$	sposoby definiowania zbioru
$\mathcal{P}(X) = 2^X = \{B \mid B \subset X\}$	**zbiór potęgowy** zbioru X

Zbiory liczbowe

$\mathcal{P}$ = zbiór liczb pierwszych
$\mathbb{N}$ = zbiór liczb naturalnych = $\{0, 1, 2, \ldots\}$
$\mathbb{Z}$ = zbiór liczb całkowitych
$\mathbb{Q}$ = zbiór liczb wymiernych
$\mathbb{R}$ = zbiór liczb rzeczywistych
$\mathbb{C}$ = zbiór liczb zespolonych



$A \triangle B$

Diagramy Venna

1.4. Relacje

Relacją n-argumentową nazywamy dowolny podzbiór R zadanego iloczynu kartezjańskiego n-zbiorów $X_1 \times \cdots \times X_n$. Bardziej formalnie: relacją jest system $\mathcal{R} = (X_1, \ldots, X_n, R)$, gdzie $R \subset X_1 \times \cdots \times X_n$. Jeżeli jasne jest jakie zbiory $X_1, \ldots, X_n$ bierzemy pod uwagę, to relację $\mathcal{R}$ utożsamiamy z jej **wykresem** R.

Relacją (2-argumentową) **w zbiorze** X nazywamy podzbiór kwadratu kartezjańskiego $X^2 = X \times X$. Mamy np. **relację pustą** $\varnothing \subset X \times X$ oraz **relację uniwersalną** $X \times X \subset X \times X$.

Mówimy, że relacja $R \subset X \times X$ jest:
(a) **zwrotna**, jeśli $\Delta_X = \{(x, x) \mid x \in X\} \subset R$,
(b) **symetryczna**, jeśli $R \subset R^{-1} = \{(y, x) \in X \times X \mid (x, y) \in R\}$,
(c) **przechodnia**, jeśli $R \circ R \stackrel{\text{def}}{=} \{(x, z) \in X \times X \mid \exists_{y \in X}\ (x, y) \in R \wedge (y, z) \in R\} \subset R$,
(d) **słabo antysymetryczna**, jeśli $R \cap R^{-1} \subset \Delta_X$,
(e) **mocno antysymetryczna**, jeśli $R \cap R^{-1} = \varnothing$,
(f) **spójna**, jeśli $X \times X = R \cup R^{-1} \cup \Delta_X$,
(g) **przeciwzwrotna**, jeśli $R \cap \Delta_X = \varnothing$,
(h) **mocno spójna**, jeśli $X \times X = R \cup R^{-1}$.

Przykłady 1.4.1.
(1) *Relacja pusta jest przeciwzwrotna, symetryczna, przechodnia, mocno antysymetryczna i nie jest spójna o ile X ma co najmniej 2 elementy.*
(2) *Relacja uniwersalna jest spójna, zwrotna, symetryczna, przechodnia i nie jest nawet słabo antysymetryczna o ile X ma co najmniej 2 elementy.*
(3) *Przekątna Δ_X jest wykresem relacji równości $=$.*

Jeśli R jest zwrotna, symetryczna i przechodnia, to nazywamy ją **relacją równoważności** (krócej: **równoważnością**). Jeśli R jest zwrotna, słabo antysymetryczna i przechodnia,

to mówimy, że R jest **relacją częściowego porządku** (krócej: **częściowym porządkiem**). Jeśli relacja częściowego porządku jest spójna, to nazywamy ją **relacją liniowego porządku** (krócej: **liniowym porządkiem**). Relacja zwrotna i przechodnia nazywa się **preporządkiem**.

Jeśli $S_1 \subset X \times Y$ i $S_2 \subset Y \times Z$ są dowolnymi relacjami 2-argumentowymi, to możemy utworzyć ich **złożenie**

$$S_2 \circ S_1 = \{(x,z) \in X \times Z \mid \exists_{y \in Y}\ (x,y) \in S_1 \wedge (y,z) \in S_2\}.$$

1.5. Odwzorowania (funkcje)

Aby zdefiniować jednoznacznie **odwzorowanie** (**funkcję**) należy określić jej dziedzinę, przeciwdziedzinę i wykres. Formalnie będziemy więc utożsamiać funkcję z uporządkowaną trójką $f = (A, B, \operatorname{graf} f)$, gdzie A, B są zbiorami oraz $\operatorname{graf} f$ jest podzbiorem $A \times B$ spełniającym następujący warunek jednoznaczności

$$\forall_{a \in A} \exists!_{b \in B}\ (a,b) \in \operatorname{graf} f. \tag{$*$}$$

Mówimy wtedy, że f przekształca zbiór A w zbiór B i zapisujemy $f\colon A \longrightarrow B$. Natomiast dla jakiegokolwiek $a \in A$ piszemy $f(a) = b$ wtedy i tylko wtedy, gdy $(a,b) \in \operatorname{graf} f$. Zbiór A nazywamy **dziedziną** funkcji f i oznaczamy $\operatorname{dom} f$ lub D_f. Zbiór B nazywamy **przeciwdziedziną** funkcji f i oznaczamy $\operatorname{ran} f$ lub D^f. Natomiast zbiór

$$f(A) = \{b \in B \mid \exists_{a \in A}\ f(a) = b\} = \{f(a) \mid a \in A\}$$

nazywamy **obrazem** funkcji f i oznaczamy $\operatorname{im} f$. Podobnie dla dowolnego podzbioru C dziedziny A definiujemy jego **obraz**

$$f(C) = \{b \in B \mid \exists_{a \in C}\ b = f(a)\} = \{f(a) \mid a \in C\}$$

oraz dla dowolnego podzbioru D przeciwdziedziny B jego **przeciwobraz**

$$f^{-1}(D) = \{a \in A \mid f(a) \in D\}.$$

Odwzorowanie f nazywamy **różnowartościowym** (**injekcją**), jeśli zachodzi warunek

$$\forall_{a_1, a_2 \in A}\ a_1 \neq a_2 \Longrightarrow f(a_1) \neq f(a_2).$$

Odwzorowanie f nazywamy odwzorowaniem **na** (**surjekcją**), jeśli $\operatorname{im} f = B$. Odwzorowanie f nazywamy **wzajemnie jednoznacznym** (**bijekcją**), jeśli jest injekcją i surjekcją.

Jeśli $f\colon A \longrightarrow B$, $g\colon B \longrightarrow C$, to określamy **złożenie** $g \circ f\colon A \longrightarrow C$ tak, że

$$\begin{aligned}\operatorname{graf}(g \circ f) &= (\operatorname{graf} g) \circ (\operatorname{graf} f)\\ &= \{(a,c) \in A \times C \mid \exists_{b \in B}\ b = f(a) \wedge c = g(b)\} = \{(a, g(f(a))) \mid a \in A\}.\end{aligned}$$

Składanie funkcji jest łączne: dla dowolnych funkcji f, g, h zachodzi $(h \circ g) \circ f = h \circ (g \circ f)$ o ile obie strony mają sens. Zwykle składanie odwzorowań nie jest przemienne.

Odwzorowania identycznościowe $\mathrm{id}_A\colon A \longrightarrow A$ zadane wzorem $\mathrm{id}_A(a) = a$ dla dowolnego zbioru A są „elementami neutralnymi" dla składania odwzorowań, tzn. jeśli $f\colon A \longrightarrow B$, to $f = \mathrm{id}_B \circ f$ oraz $f = f \circ \mathrm{id}_A$.

Odwzorowanie $f\colon A \longrightarrow B$ jest **odwracalne** (tzn. istnieje odwzorowanie **odwrotne** $f^{-1}\colon B \longrightarrow A$ takie, że $f \circ f^{-1} = \mathrm{id}_B$ i $f^{-1} \circ f = \mathrm{id}_A$) wtedy i tylko wtedy, gdy f jest bijekcją. W takim przypadku $\mathrm{graf}(f^{-1})$ jest relacją odwrotną do relacje $\mathrm{graf}(f)$.

Restrykcją (**zawężeniem**, **obcięciem**) odwzorowania $f\colon A \longrightarrow B$ do zbioru $A_1 \subset A$ nazywamy odwzorowanie $f|_{A_1}\colon A_1 \longrightarrow B$ takie, że $\mathrm{graf}(f|_{A_1}) = \mathrm{graf}(f) \cap (A_1 \times B)$. Czasem stosuje się też **restrykcję w przeciwdziedzinie**: jeśli $\operatorname{im} f \subset B_1 \subset B$, to $f|^{B_1}\colon A \longrightarrow B_1$ jest takim odwzorowaniem, że $\mathrm{graf}(f|^{B_1}) = \mathrm{graf}(f) \cap (A \times B_1)$.

Przykłady 1.5.1. *Odwzorowanie $f\colon \mathbb{R} \longrightarrow \mathbb{R}$ zadane wzorem $f(x) = x^2$ nie jest ani injekcją ani surjekcją. Jego restrykcja $f|_{[0,+\infty)}$ jest injekcją, a jego restrykcja w przeciwdziedzinie $f|^{[0,+\infty)}$ jest surjekcją. Bijekcją jest restrykcja w dziedzinie i przeciwdziedzinie $f|^{[0,+\infty)}_{[0,+\infty)}\colon [0,+\infty) \longrightarrow [0,+\infty)$ i ta funkcja ma odwrotną $g\colon [0,+\infty) \longrightarrow [0,+\infty)$, $g(x) = \sqrt{x}$.*

1.6. Relacje równoważności

Relację równoważności często zapisujemy przy użyciu symbolu $\sim$. Jeśli $x \sim y$, to mówimy, że x jest **równoważne** y (w sensie relacji $\sim$). Wtedy przez $[x]_\sim$ ($[x]$ lub $\bar{x}$) oznaczamy **klasę równoważności** elementu x, czyli zbiór

$$[x]_\sim = \{y \in X \mid x \sim y\}.$$

Zbiór wszystkich klas równoważności $\{[x]_\sim \mid x \in X\}$ nazywamy **zbiorem ilorazowym** zbioru X przez relację $\sim$ i oznaczamy $X/\!\sim$. Klasy równoważności stanowią **rozkład** zbioru X (tzn. jest to rodzina niepustych, parami rozłącznych podzbiorów X o sumie mnogościowej będącej całym X). Odwrotnie: zadanie rozkładu zbioru X wyznacza jednoznacznie relację równoważności w X. Odwzorowanie

$$X \in x \longmapsto [x]_\sim \in X/\!\sim$$

jest wyznaczone przez $\sim$; nazywamy je **odwzorowaniem naturalnym** lub **kanonicznym** (lub **surjekcją kanoniczną**).

Jeśli w X określone są działania (np. $+, \cdot$), to relację równoważności $\sim$ na X nazywamy **kongruencją**, gdy relacja ta jest zgodna z działaniami, czyli np.

$$x_1 \sim y_1,\ x_2 \sim y_2 \Longrightarrow x_1 + x_2 \sim y_1 + y_2,\ x_1 \cdot x_2 \sim y_1 \cdot y_2.$$

Przykłady 1.6.1. *Niech $n \in \mathbb{N} \setminus \{0\}$. Definiujemy $\sim\, \subset \mathbb{Z} \times \mathbb{Z}$ przy pomocy warunku:*

$$k_1 \sim k_2 \iff n \mid k_1 - k_2 \iff \exists_{l \in \mathbb{Z}}\ l \cdot n = k_1 - k_2.$$

Wtedy $k_1 \sim k_2$ i $k_3 \sim k_4$ implikuje $k_1 + k_3 \sim k_2 + k_4$ i $k_1 \cdot k_3 \sim k_2 \cdot k_4$, bo jeśli $k_1 - k_2 = l_1 \cdot n$, $k_3 - k_4 = l_2 \cdot n$, to

$$(k_1 + k_3) - (k_2 + k_4) = (k_1 - k_2) + (k_3 - k_4) = (l_1 + l_2) \cdot n,$$

$$k_1 \cdot k_3 - k_2 \cdot k_4 = k_1 \cdot (k_3 - k_4) + (k_1 - k_2) \cdot k_4 = (k_1 \cdot l_2 + k_2 \cdot l_1) \cdot n.$$

Relacja $\sim$ jest równoważnością, a więc $\mathbb{Z}$ jest kongruencją na $\mathbb{Z}$. Jej klasy równoważności można utożsamiać z resztami z dzielenia przez n. Zbiór ilorazowy oznaczamy $\mathbb{Z}_n$, jest to

zbiór n-elementowy. Interesujące jest, że kongruencja pozwala przenieść działania na zbiór ilorazowy:

$$[k] + [l] = [k + l],$$
$$[k] \cdot [l] = [k \cdot l]$$

(wynik działania nie zależy od wyboru reprezentanta z klasy równoważności). Na przykład w $\mathbb{Z}_5$ *mamy:*

$+$	$[0]$	$[1]$	$[2]$	$[3]$	$[4]$
$[0]$	$[0]$	$[1]$	$[2]$	$[3]$	$[4]$
$[1]$	$[1]$	$[2]$	$[3]$	$[4]$	$[0]$
$[2]$	$[2]$	$[3]$	$[4]$	$[0]$	$[1]$
$[3]$	$[3]$	$[4]$	$[0]$	$[1]$	$[2]$
$[4]$	$[4]$	$[0]$	$[1]$	$[2]$	$[3]$

$\cdot$	$[0]$	$[1]$	$[2]$	$[3]$	$[4]$
$[0]$	$[0]$	$[0]$	$[0]$	$[0]$	$[0]$
$[1]$	$[0]$	$[1]$	$[2]$	$[3]$	$[4]$
$[2]$	$[0]$	$[2]$	$[4]$	$[1]$	$[3]$
$[3]$	$[0]$	$[3]$	$[1]$	$[4]$	$[2]$
$[4]$	$[0]$	$[4]$	$[3]$	$[2]$	$[1]$

1.7. Równoliczność

Mówimy, że dwa zbiory A, B są **równoliczne**, jeśli istnieje bijekcja z A do B. „Relacja" równoliczności w **klasie wszystkich zbiorów** (!) jest „równoważnością". „Klasę równoważności" zbioru A nazywamy jego **licznością** lub **mocą** (i oznaczamy $|A|$, $\#A$ lub $\bar{\bar{A}}$).

Zbiór A nazywamy **nieskończonym**, jeśli jest on równoliczny ze swoim właściwym podzbiorem (podzbiorem nie będącym całym zbiorem). W przeciwnym razie mówimy, że zbiór A jest **skończony**.

Moce zbiorów nazywamy **liczbami kardynalnymi**. Moce skończone utożsamiamy z liczbami naturalnymi: mocą zbioru $\{0, 1, \ldots, n-1\}$ jest liczba $n \in \mathbb{N}$. Zbiór liczb naturalnych jest nieskończony (bo $\mathbb{N} \ni n \longmapsto n+1 \in \mathbb{N} \setminus \{0\}$ jest bijekcją). Jego moc **alef zero** ($\aleph_0$) jest najmniejszą nieskończoną liczbą kardynalną. Następne to $\aleph_1, \aleph_2, \ldots$ Istnieje nieskończenie wiele nieskończonych liczb kardynalnych. Zbiory mocy alef zero nazywamy **przeliczalnymi**. Zbiór skończony lub przeliczalny nazywamy **co najwyżej przeliczalnym**. Zbiory pozostałe nazywamy **nieprzeliczalnymi**.

Mówimy, że $|A| \leqslant |B|$, jeśli istnieje injekcja $i : A \hookrightarrow B$ ([1]). Równoważnie: istnieje surjekcja $s\colon B \twoheadrightarrow A$([2]). Dowodzi się, że „relacja" $\leqslant$ jest „liniowym porządkiem" w klasie wszystkich liczb kardynalnych.

Oś liczb kardynalnych

Twierdzenie 1.7.1.

(*a*) $|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}|$, *czyli* $\mathbb{Z}$ *i* $\mathbb{Q}$ *są przeliczalne.*

(*b*) $\mathbb{R}$ *jest zbiorem nieprzeliczalnym.*

(*c*) *Dla dowolnego zbioru* A *jest* $|A| < |\mathcal{P}(A)|$.

[1] Tak zapisujemy injekcję.

[2] A tak zapisujemy surjekcję.

Dowód.
(a) Łatwe, patrz [RW, str. 810].
(b) Widać z **metody przekątniowej Cantora** ([RW, str. 812]). Załóżmy, że istnieje surjekcja $f\colon \mathbb{N} \longrightarrow \mathbb{R}$, czyli wszystkie liczby rzeczywiste można ustawić w ciąg nieskończony:

$$\begin{array}{rcl}
f(1) & = & \cdots\, e_{1,3}\, e_{1,2}\, e_{1,1}\ ,\ \boldsymbol{d_{1,1}}\, d_{1,2}\, d_{1,3}\, \cdots \\
f(2) & = & \cdots\, e_{2,3}\, e_{2,2}\, e_{2,1}\ ,\ d_{2,1}\, \boldsymbol{d_{2,2}}\, d_{2,3}\, \cdots \\
f(3) & = & \cdots\, e_{3,3}\, e_{3,2}\, e_{3,1}\ ,\ d_{3,1}\, d_{3,2}\, \boldsymbol{d_{3,3}}\, \cdots \\
\multicolumn{3}{c}{\dotfill} \\
f(n) & = & \cdots\, e_{n,3}\, e_{n,2}\, e_{n,1}\ ,\ d_{n,1}\, d_{n,2}\, d_{n,3}\, \cdots\, \boldsymbol{d_{n,n}}\, \cdots \\
\multicolumn{3}{c}{\dotfill}
\end{array}$$

Niech x będzie liczbą o przedstawieniu dziesiętnym $0, d_1\, d_2\, d_3\, \cdots$, gdzie

$$d_i = \begin{cases} 1, & \text{jeśli } d_{i,i} \neq 1, \\ 2, & \text{jeśli } d_{i,i} = 1. \end{cases}$$

Wtedy x nie jest żadną z liczb $f(1), f(2), f(3), \ldots$, a jest liczbą rzeczywistą (korzystamy z tego, że x ma jednoznaczne przedstawienie dziesiętne). Sprzeczność z założeniem pokazuje, że nieprawdą jest $|\mathbb{R}| \leqslant |\mathbb{N}|$, czyli że $\mathbb{R}$ jest zbiorem nieprzeliczalnym.
(c) Nierówność $|A| \leqslant |\mathcal{P}(A)|$ jest oczywista, bo mamy injekcję

$$I\colon A \ni x \longmapsto \{x\} \in \mathcal{P}(A).$$

Załóżmy, że $|A| = |\mathcal{P}(A)|$, czyli że istnieje bijekcja $F\colon A \ni x \longmapsto F(x) \in \mathcal{P}(A)$. Weźmy zbiór

$$Z = \{x \in A \mid x \notin F(x)\}$$

oraz dowolny element $x_0 \in A$. Jeśli $x_0 \in Z$, to $x_0 \notin F(x_0)$. Zatem $Z \neq F(x_0)$. Jeśli $x_0 \notin Z$, to $x_0 \in F(x_0)$, zatem $Z \neq F(x_0)$. Czyli zbiór Z nie jest wartością bijekcji F. Sprzeczność dowodzi, że $|A| \neq |\mathcal{P}(A)|$, czyli $|A| < |\mathcal{P}(A)|$. □

Rozdział 2

Więcej o relacjach

2.1. Relacje porządku

Relacje częściowego lub liniowego porządku często zapisujemy przy użyciu symbolu $\leqslant$ lub $\preccurlyeq$. Jeśli $x \preccurlyeq y$, to mówimy, że x jest **niewiększy niż** y (w sensie relacji $\preccurlyeq$). Wtedy relacja $\prec$ określona warunkiem

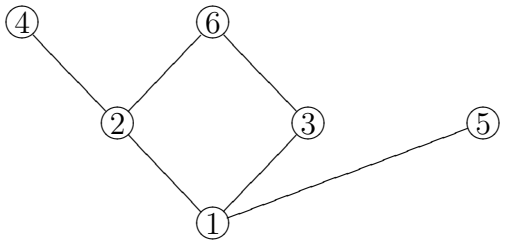
$$x \prec y \iff x \preccurlyeq y \wedge x \neq y$$

jest przeciwzwrotna i przechodnia. Takie relacje nazywamy **quasi-porządkami**. Odwrotnie: mając quasi-porządek $\prec$ możemy zdefiniować odpowiadającą mu relację częściowego porządku warunkiem

$$x \preccurlyeq y \iff x \prec y \vee x = y.$$

Jeśli X jest zbiorem skończonym, to relację częściowego porządku na X można przedstawić graficznie przy pomocy **diagramu Hassego**. Mówimy, że element y **nakrywa** element x, jeśli $x \prec y$ oraz nie istnieje element pośredni z (tzn. taki, że $x \prec z \prec y$). Wtedy y przedstawiamy jako punkt (wierzchołek) powyżej punktu x i prowadzimy odcinek (krawędź) od y do x.

Przykłady 2.1.1. *Niech $X = \{1, 2, 3, 4, 5, 6\}$ i niech $\mid$ będzie relacją podzielności w X. Wtedy diagramem Hassego jest*



Dla dowolnego zbioru częściowo uporządkowanego $(X, \preccurlyeq)$ mówimy, że:
(1) x jest **elementem maksymalnym**, jeśli $\forall_{y \in X}\ x \preccurlyeq y \Longrightarrow x = y$,
(2) x jest **elementem minimalnym**, jeśli $\forall_{y \in X}\ y \preccurlyeq x \Longrightarrow x = y$,
(3) x jest **elementem największym**, jeśli $\forall_{y \in X}\ y \preccurlyeq x$,
(4) x jest **elementem najmniejszym**, jeśli $\forall_{y \in X}\ x \preccurlyeq y$.
Jeżeli ponadto dany jest podzbiór $S \subset X$, to mówimy, że:
(5) x jest **ograniczeniem górnym** (**majorantą**) S, jeśli $\forall_{y \in S}\ y \preccurlyeq x$,
(6) x jest **ograniczeniem dolnym** (**minorantą**) S, jeśli $\forall_{y \in S}\ x \preccurlyeq y$,
(7) x jest **kresem górnym** (**supremum**) S, jeśli jest najmniejszą majorantą S,
(8) x jest **kresem dolnym** (**infimum**) S, jeśli jest największą minorantą S.

Jeśli zachodzi (7) lub (8), to piszemy $x = \sup S$ lub $x = \inf S$ odpowiednio.

Kratą (ang. lattice) nazywamy zbiór częściowo uporządkowany $(X, \leqslant)$, w którym dla dowolnych $x, y \in X$ istnieją kres górny i kres dolny zbioru $\{x, y\}$. Krata $(X, \leqslant)$ jest **zupełna** jeśli dla dowolnego $A \subset X$ istnieją $\sup A$ oraz $\inf A$.

Przykłady 2.1.2.
(1) *Każdy liniowy porządek jest kratą.*
(2) *Zbiór* $\mathbb{N}\setminus\{0\}$ *z relacją podzielności | jest kratą (wtedy* $\sup\{x, y\} = \mathrm{nww}(x, y)$, $\inf\{x, y\} = \mathrm{nwd}(x, y)$*).*
(3) *Przedział domknięty* $[0, 1] \subset \mathbb{R}$ *wraz z relacją* $\leqslant$ *jest kratą zupełną.*

Kratę zdefiniować można również jako strukturę $(X, \wedge, \vee)$, gdzie $\vee\colon X \times X \longrightarrow X$, $\wedge\colon X \times X \longrightarrow X$ są działaniami wewnętrznymi w X (czytanymi odpowiednio: sup i inf) spełniającymi aksjomaty:
$(K1)$ $\forall_{x,y\in X}\ x \wedge x = x,\ x \vee x = x$ (idempotentność),
$(K2)$ $\forall_{x,y\in X}\ x \wedge y = y \wedge x,\ x \vee y = y \vee x$ (przemienność),
$(K3)$ $\forall_{x,y,z\in X}\ x \wedge (y \wedge z) = (x \wedge y) \wedge z,\ x \vee (y \vee z) = (x \vee y) \vee z$ (łączność),
$(K4)$ $\forall_{x,y\in X}\ x \wedge (x \vee y) = x,\ x \vee (x \wedge y) = x$ (pochłanianie).
Jeśli dodatkowo zachodzi
(5) $\forall_{x,y,z\in X}\ x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z),\ x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$ (rozdzielność),
to mówimy, że krata jest **rozdzielna**.

Twierdzenie 2.1.3. *Powyższe definicje kraty są równoważne, tzn. istnieje wzajemnie jednoznaczna odpowiedniość pomiędzy kratami postaci* $(X, \leqslant)$ *oraz postaci* $(X, \wedge, \vee)$.

Dowód. Dla kraty w pierwszym sensie $(X, \leqslant)$ określamy:

$$x \wedge y = \inf\{x, y\},$$
$$x \vee y = \sup\{x, y\}.$$

Wtedy $(X, \wedge, \vee)$ jest kratą w drugim sensie, bo:

$$x = \inf\{x, x\} = \sup\{x, x\}, \quad \inf\{x, y\} = \inf\{y, x\}, \quad \sup\{x, y\} = \sup\{y, x\}.$$

Jeśli $x_1 = \inf\{\inf\{x, y\}, z\}$ oraz $x_2 = \inf\{x, \inf\{y, z\}\}$, to

$$x_1 \leqslant x_2, \quad \text{bo } x_1 \leqslant x,\ x_1 \leqslant \inf\{y, z\}, \quad \text{czyli } x_1 \text{ jest minorantą } \{x, \inf\{y, x\}\},$$

$$x_2 \leqslant x_1, \quad \text{bo } x_2 \leqslant z,\ x_2 \leqslant \inf\{x, y\}, \quad \text{czyli } x_2 \text{ jest minorantą } \{z, \inf\{x, y\}\},$$

stąd $x_1 = x_2$ (operacja $\wedge$ jest łączna). Analogicznie dowodzi się, że operacja $\vee$ jest łączna. Ponadto:

$$\inf\{x, \sup\{x, y\}\} = x, \quad \text{bo } x \leqslant \sup\{x, y\},$$
$$\sup\{x, \inf\{x, y\}\} = x, \quad \text{bo } x \geqslant \inf\{x, y\}.$$

Dla kraty w drugim sensie $(X, \wedge, \vee)$ określamy

$$x \preccurlyeq y \iff x \vee y = y \ (\iff x \wedge y = x \text{ wobec } (K4)).$$

Oczywiście $x \preccurlyeq x$, bo $x \vee x = x$. Jeśli $x \preccurlyeq y$ i $y \preccurlyeq x$, to $x = x \vee y = y$. Jeśli $x \preccurlyeq y$ i $y \preccurlyeq z$, to

$$z = z \vee y = z \vee (y \vee x) = (z \vee y) \vee x = z \vee x,$$

czyli $x \preccurlyeq z$. Zauważmy, że $x \preccurlyeq z$ i $y \preccurlyeq z$ oznacza $x \vee z = z$, $y \vee z = z$, czyli

$$(x \vee y) \vee z = x \vee (y \vee z) = x \vee z = z$$

i $x \vee y \preccurlyeq z$. Odwrotnie $x \vee y \preccurlyeq z$ oznacza, że $x \preccurlyeq z$ i $y \preccurlyeq z$, bo zawsze $x \preccurlyeq x \vee y$, $y \preccurlyeq y \vee x = x \vee y$, zatem

$$\sup_{\preccurlyeq}\{x, y\} = x \vee y, \qquad \inf_{\preccurlyeq}\{x, y\} = x \wedge y.$$

Przejścia te są wzajemnie odwrotne, bo dla przejścia $(X, \leqslant) \longrightarrow (X, \wedge, \vee) \longrightarrow (X, \preccurlyeq)$ mamy

$$x \preccurlyeq y \iff x \vee y = y \iff \sup_{\leqslant}\{x, y\} = y \iff x \leqslant y,$$

czyli $(X, \leqslant) = (X, \preccurlyeq)$; natomiast dla przejścia $(X, \wedge, \vee) \longrightarrow (X, \preccurlyeq) \longrightarrow (X, \curlywedge, \curlyvee)$ jest

$$x \curlywedge y = \inf_{\preccurlyeq}\{x, y\} = x \wedge y,$$

$$x \curlyvee y = \sup_{\preccurlyeq}\{x, y\} = x \vee y,$$

czyli $(X, \wedge, \vee) = (X, \curlywedge, \curlyvee)$. □

Zbiór liniowo uporządkowany nazywamy **dobrze uporządkowanym**, gdy każdy jego niepusty podzbiór ma element najmniejszy.

Przykłady 2.1.4. *Każdy skończony zbiór liniowo uporządkowany jest dobrze uporządkowany. Również dobrze uporządkowany jest* $(\mathbb{N}, \leqslant)$.

Niech teraz $(X_1, \leqslant_1), \dots, (X_n, \leqslant_n)$ będą zbiorami częściowo uporządkowanymi.

Porządkiem produktowym $\leqslant_p$ na $X_1 \times \cdots \times X_n$ nazywamy porządek częściowy określony warunkiem

$$(x_1, \dots, x_n) \leqslant_p (y_1, \dots, y_n) \iff (x_1 \leqslant_1 y_1) \wedge (x_2 \leqslant_2 y_2) \wedge \cdots \wedge (x_n \leqslant_n y_n).$$

Porządkiem słownikowym (**leksykograficznym**) $\leqslant_l$ na $X_1 \times \cdots \times X_n$ nazywamy częściowy porządek określony warunkiem

$$\begin{aligned}(x_1, \dots, x_n) \leqslant_l (y_1, \dots, y_n) \iff {} & (x_1 <_1 y_1) \vee (x_1 = y_1 \wedge x_2 <_2 y_2) \\ & \vee \cdots \vee (x_1 = y_1 \wedge \cdots \wedge x_{n-1} = y_{n-1} \wedge x_n \leqslant_n y_n)\end{aligned}$$

Twierdzenie 2.1.5. *Relacje* $\leqslant_p$ *i* $\leqslant_l$ *są rzeczywiście częściowymi porządkami, przy czym porądek leksykograficzny jest mocniejszy od produktowego:* $\leqslant_p \subset \leqslant_l$.

Dowód. Ćwiczenie częściowo omówione w [RW]. □

Twierdzenie 2.1.6. *Jeśli* $(X_1, \leqslant_1), \dots, (X_n, \leqslant_n)$ *są liniowo uporządkowane, to* $(X_1 \times \cdots \times X_n, \leqslant_l)$ *również.*

Dowód. Ćwiczenie ([RW, str. 657]). □

Dowolny niepusty zbiór Σ będziemy nazywać **alfabetem**. Elementy alfabetu nazywamy **literami**. **Słowem** alfabetu Σ dowolny skończony (być może pusty) ciąg liter. Zbiór wszystkich słów alfabetu Σ oznaczamy Σ^* ($\Sigma^* = \coprod_{n \in \mathbb{N}} \Sigma^n$). **Długością** $d(w)$ słowa w jest długość ciągu, czyli ilość liter słowa liczona z krotnościami (powtarzające się litery liczymy

za każdym razem). **Językiem nad alfabetem** Σ nazywamy dowolny podzbiór J zbioru wszystkich słów Σ^* alfabetu Σ.

Załóżmy teraz, że alfabet Σ jest częściowo uporządkowany przez $\leqslant$. Wtedy w każdej potędze kartezjańskiej Σ^k mamy porządek leksykograficzny $\leqslant^k$. **Porządkiem standardowym** $\leqslant^*$ w Σ^* nazywamy porządek określony warunkiem

$$w_1 \leqslant^* w_2 \iff d(w_1) < d(w_2) \vee (d(w_1) = d(w_2) \wedge w_1 \leqslant^{d(w_1)} w_2).$$

Twierdzenie 2.1.7. *Jeżeli $\leqslant$ jest liniowym porządkiem na skończonym alfabecie Σ, to $\leqslant^*$ jest dobrym porządkiem na Σ^*.*

Dowód. Porządek w Σ^* jest liniowy, bo jest sklejeniem liniowych porządków leksykograficznych $\leqslant^k$ na odpowiednich Σ^k:

$$\Sigma^* = \Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup \cdots .$$

Skoro Σ jest skończony, to każdy ze zbiorów Σ^k jest skończony ($|\Sigma^k| = |\Sigma|^k$). Weźmy dowolny niepusty język $J \subset \Sigma^*$. Skoro $(\mathbb{N}, \leqslant)$ jest dobrze uporządkowany, to istnieje najmniejsze takie $n \in \mathbb{N}$, że $J \cap \Sigma^n \neq \varnothing$. Nazwijmy je n_0. Zbiór $J \cap \Sigma^{n_0}$ jest skończony i liniowo uporządkowany, zatem jest dobrze uporządkowany. Istnieje najmniejsze słowo $w_0 \in J \cap \Sigma^{n_0}$. Słowo w_0 jest najmniejszym elementem J, bo dla porządku $\leqslant^*$ jeśli $n_1 < n_2$ i $w_1 \in \Sigma^{n_1}$, $w_2 \in \Sigma^{n_2}$, to $w_1 <^* w_2$. Wykazaliśmy, że każdy niepusty język ma najmniejsze słowo. □

Porządek leksykograficzny (**słownikowy**) $\leqslant_L$ na Σ^* określamy w następujący sposób: jeśli $a = (a_1, \ldots, a_n)$, $b = (b_1, \ldots, b_m)$, to

$$a \leqslant_L b \iff (a_1, \ldots, a_k) <^k (b_1, \ldots, b_k) \text{ dla } k = \min\{n, m\}$$
$$\text{lub } (a_1, \ldots, a_k) = (b_1, \ldots, b_k) \text{ i } d(a) \leqslant d(b).$$

(To jest tak, jakby dodać spację jako najmniejszą literę.)

Twierdzenie 2.1.8. *Jeśli $\leqslant$ jest liniowym porządkiem na Σ, to $\leqslant_L$ jest liniowym porządkiem na Σ^**

Dowód. [RW, str. 660]
(a) Zwrotność: $(a_1, \ldots, a_n) = (a_1, \ldots, a_n)$ i $n \leqslant n$.
(b) Słaba antysymetria: jeśli dla $k = \min\{n, m\}$ jest

$$\big[(a_1, \ldots, a_k) <^k (b_1, \ldots, b_k) \vee \big((a_1, \ldots, a_k) = (b_1, \ldots, b_k) \wedge n \leqslant m\big)\big]$$
$$\wedge \big[(b_1, \ldots, b_k) <^k (a_1, \ldots, a_k) \vee \big((b_1, \ldots, b_k) = (a_1, \ldots, a_k) \wedge m \leqslant n\big)\big],$$

to $(a_1, \ldots, a_k) = (b_1, \ldots, b_k)$ i $n = m = k$, czyli $a = b$.
(c) Przechodniość: niech $a \leqslant_L b$ i $b \leqslant_L c$. Oznaczmy $k = \min\{d(a), d(b)\}$, $l = \min\{d(b), d(c)\}$, $p = \min\{k, l\}$, $q = \min\{d(a), d(c)\}$. Mamy

$$\big[(a_1, \ldots, a_k) <^k (b_1, \ldots, b_k) \vee \big((a_1, \ldots, a_k) = (b_1, \ldots, b_k) \wedge d(a) \leqslant d(b)\big)\big]$$
$$\wedge \big[(b_1, \ldots, b_l) <^l (c_1, \ldots, c_l) \vee \big((b_1, \ldots, b_l) = (c_1, \ldots, c_l) \wedge d(b) \leqslant d(c)\big)\big].$$

Na pewno $(a_1, \ldots, a_p) \leqslant^p (c_1, \ldots, c_p)$, czyli albo $(a_1, \ldots, a_p) <^p (c_1, \ldots, c_p)$ i wtedy mamy $(a_1, \ldots, a_q) <^q (c_1, \ldots, c_q)$, czyli $a \leqslant_L c$ albo $(a_1, \ldots, a_p) = (c_1, \ldots, c_p)$ i teraz

(1) albo $k = p$, $k \leqslant l$, $q = d(a) \leqslant d(c)$, $(a_1, \ldots, a_k) = (b_1, \ldots, b_k) = (c_1, \ldots, c_k)$, $p = d(a) \leqslant d(b)$ i $a \leqslant_L c$.
(2) albo $k > p$, $k > l = p$, $p = d(c) = l$, $q = d(c)$, $(b_1, \ldots, b_q) = (c_1, \ldots, c_q) = (a_1, \ldots, a_q)$, $d(b) \geqslant d(c)$, czyli $b >_L c$ i mamy sprzeczność.

(*d*) Spójność: Jeśli $(a_1, \ldots, a_k) = (b_1, \ldots, b_k)$, to $a \leqslant_L b$ lub $b \leqslant_L a$. Jeśli $(a_1, \ldots, a_k) <^k (b_1, \ldots, b_k)$, to $a \leqslant_L b$. Jeśli zaś $(a_1, \ldots, a_k) >^k (b_1, \ldots, b_k)$, to $a \geqslant_L b$. □

2.2. Składanie relacji

Przypomnijmy, że jeśli $R_1 \subset S \times T$ i $R_2 \subset T \times U$, to ich złożenie $R_2 \circ R_1 = R_1 R_2 \subset S \times U$ jest określone warunkiem

$$(s, u) \in R_2 \circ R_1 \iff \exists_{t \in T}\ (s, t) \in R_1 \wedge (t, u) \in R_2.$$

Twierdzenie 2.2.1. *Składanie relacji jest łączne: jeśli dana jest relacja $R_3 \subset U \times W$, to $R_3 \circ (R_2 \circ R_1) = (R_3 \circ R_2) \circ R_1$.*

Dowód.

$$\begin{aligned}
(s, w) \in R_3 \circ (R_2 \circ R_1) &\iff \exists_{u \in U}\ (s, u) \in R_2 \circ R_1 \wedge (u, w) \in R_2 \\
&\iff \exists_{u \in U} \exists_{t \in T}\ (s, t) \in R_1 \wedge (t, u) \in R_2 \wedge (u, w) \in R_3 \\
&\iff \exists_{t \in T} \exists_{u \in U}\ (s, t) \in R_1 \wedge (t, u) \in R_2 \wedge (u, w) \in R_3 \\
&\iff \exists_{t \in T}\ (s, t) \in R_1 \wedge \exists_{u \in U}\ (t, u) \in R_2 \wedge (u, w) \in R_3 \\
&\iff \exists_{t \in T}\ (s, t) \in R_1 \wedge (t, w) \in R_3 \circ R_2 \\
&\iff (s, w) \in (R_3 \circ R_2) \circ R_1.
\end{aligned}$$ □

„Elementem neutralnym" dla składania relacji są relacje równości, które można utożsamiać z odwzorowaniami identycznościowymi, a których wykresy są przekątnymi $E_X = \{(x, y) \in X \times X \mid x = y\} = \Delta_X$.

Twierdzenie 2.2.2. *Zachodzą wzory $(R_2 \circ R_1)^{-1} = R_1^{-1} \circ R_2^{-1}$ oraz $(R_1^{-1})^{-1} = R_1$.*

Dowód. Ćwiczenie. □

2.3. Macierze relacji

Załóżmy teraz, że zbiory S, T, U są skończone. Numerujemy elementy zbiorów S, T, U. **Macierzą relacji** $R \subset S \times T$ nazywamy macierz M_R o $|S|$ wierszach i $|T|$ kolumnach, w której na miejscu $[i, j]$ jest 1, gdy $(s_i, t_j) \in R$ oraz 0, gdy $(s_i, t_j) \notin R$, $i = 1, \ldots, |S|$, $j = 1, \ldots, |T|$. Odpowiedniość pomiędzy relacjami w $S \times T$ i macierzami zerojedynkowymi wymiaru $|S| \times |T|$ jest wzajemnie jednoznaczna.

Traktujemy teraz elementy 0 i 1 jako elementy dwuelementowej kraty $\mathbb{B} = (\{0, 1\}, \wedge, \vee)$, gdzie

$\wedge$	0	1
0	0	0
1	0	1

,

$\vee$	0	1
0	0	1
1	1	1

,

czyli $0 < 1$. Pozwala nam to wprowadzić **iloczyn boolowski macierzy**: dla danych macierzy $M \in M_{m,n}(\mathbb{B})$ i $N \in M_{n,p}(\mathbb{B})$ definiujemy

$$(M * N)[i,k] = \bigvee_{j=1}^{n} (M[i,j] \wedge N[j,k])$$

i wtedy $M * N \in M_{m,p}(\mathbb{B})$. Ponadto dla macierzy tego samego wymiaru $m \times n$ definiujemy operacje $\wedge, \vee$ wzorami:

$$(M_1 \wedge M_2)[i,j] = M_1[i,j] \wedge M_2[i,j],$$
$$(M_1 \vee M_2)[i,j] = M_1[i,j] \vee M_2[i,j]$$

oraz nierówność warunkiem

$$M_1 \leqslant M_2 \iff \forall_{i=1,\dots,m} \forall_{j=1,\dots,n}\ M_1[i,j] \leqslant M_2[i,j].$$

Twierdzenie 2.3.1. *Macierzą relacji $R_2 \circ R_1 = R_1 R_2$ jest $M_{R_1} * M_{R_2}$.*

Dowód.

$$\begin{aligned}
(s_i, u_k) \in R_2 \circ R_1 &\iff \exists_{t \in T}\ (s_i, t) \in R_1 \wedge (t, u_k) \in R_2 \\
&\iff \bigvee_{j=1}^{n} \big((s_i, t_j) \in R_1 \wedge (t_j, u_k) \in R_2\big) \\
&\iff \bigvee_{j=1}^{n} \big(M_{R_1}[i,j] = 1 \wedge M_{R_2}[j,k] = 1\big) \\
&\iff \bigvee_{j=1}^{n} \big(M_{R_1}[i,j] \wedge M_{R_2}[j,k] = 1\big) \\
&\iff \Big(\bigvee_{j=1}^{n} \big(M_{R_1}[i,j] \wedge M_{R_2}[j,k]\big)\Big) = 1 \\
&\iff (M_{R_1} * M_{R_2})[i,k] = 1,
\end{aligned}$$

czyli $M_{R_2 \circ R_1}[i,k] = (M_{R_1} * M_{R_2})[i,k]$ dla dowolnych i, k. Krócej $M_{R_2 \circ R_1} = M_{R_1} * M_{R_2}$ (to uzasadnia zapis $R_1 R_2$). □

Wniosek 2.3.2. *Iloczyn boolowski macierzy jest działaniem łącznym.*

Twierdzenie 2.3.3. *Niech dane będą relacje R, R_1, R_2 na zbiorze skończonym S. Wtedy:*

(a) R jest zwrotna $\iff M_R \wedge I = I$ ($\iff$ M_R ma same jedynki na przekątnej),
(b) R jest przeciwzwrotna $\iff M_R \wedge I = 0$ ($\iff$ M_R ma same zera na przekątnej),
(c) R jest symetryczna $\iff M_R = M_R^T$,
(d) R jest przechodnia $\iff M_R * M_R \leqslant M_R$,
(e) R jest słabo antysymetryczna $\iff M_R \wedge M_R^T \leqslant I$,
(f) R jest silnie antysymetryczna $\iff M_R \wedge M_R^T = 0$,
(g) $M_{R_1 \cup R_2} = M_{R_1} \vee M_{R_2}$,
(h) $M_{R_1 \cap R_2} = M_{R_1} \wedge M_{R_2}$,
(i) $R_1 \subset R_2 \iff M_{R_1} \leqslant M_{R_2}$.

Dowód. Przez ogląd. □

2.4. Algebry Boole'a

Algebrą Boole'a nazywamy strukturę $(X, \wedge, \vee, \sim, 0, 1)$, gdzie $(X, \wedge, \vee)$ jest kratą rozdzielną, stałe 0 i 1 oznaczają odpowiednio element najmniejszy i największy, a jednoelementowa operacja $\sim$, zwana **dopełnieniem boolowskim** spełnia warunki (zwane **prawami dopełnienia**)

$$\forall_{x\in X} \quad x \vee (\sim x) = 1, \quad x \wedge (\sim x) = 0. \tag{D}$$

Przykłady 2.4.1.
(1) *W* ***trywialnej*** *algebrze Boole'a jest* $0 = 1$*, czyli* $X = \{0\}$*.* ***Taką strukturę zwykle wyłącza się z rozważań, czyli zakłada się, że w algebrach Boole'a jest*** $0 \neq 1$*.*
(2) *Najmniejszą nietrywialną algebrą Boole'a jest* $\mathbb{B} = (\{0, 1\}, \wedge, \vee, \sim, 0, 1)$*, gdzie* $\wedge, \vee$ *są określone tak jak poprzednio oraz* $\sim 0 = 1$*,* $\sim 1 = 0$*.*

Przez analogię do dopełnienia, operację $\wedge$ w algebrze Boole'a nazywa się **iloczynem boolowskim**, a operację $\vee$ nazywamy **sumą boolowską**.

Twierdzenie 2.4.2. *W każdej algebrze Boole'a zachodzą następujące własności:*
(*a*) $\forall_x\ x \vee 0 = x,\ x \wedge 0 = 0,\ x \vee 1 = 1,\ x \wedge 1 = x,$
(*b*) $\forall_{x,y}\ \sim (x \vee y) = (\sim x) \wedge (\sim y),\ \sim (x \wedge y) = (\sim x) \vee (\sim y).$

Dowód. Prawa (*a*) mówią, że 0 jest elementem najmniejszym, a 1 elementem największym, więc wynikają z definicji. Wykażemy pierwsze z praw de Morgana (*b*).

$$\begin{aligned}(x \vee y) \vee \big((\sim x) \wedge (\sim y)\big) &= \big((x \vee y) \vee (\sim x)\big) \wedge \big((x \vee y) \vee (\sim y)\big)\\ &= \big((x \vee (\sim x)) \vee y\big) \wedge \big(x \vee (y \vee (\sim y))\big) = 1 \wedge 1 = 1,\end{aligned}$$

$$(x \vee y) \wedge \big((\sim x) \wedge (\sim y)\big) = (x \wedge (\sim x) \wedge (\sim y)) \vee (y \wedge (\sim x) \wedge (\sim y)) = 0 \vee 0 = 0.$$

Element $(\sim x) \wedge (\sim y)$ spełnia prawa dopełnienia dla $x \vee y$. Pozostaje wykazać, że prawa dopełnienia wyznaczają jednoznacznie dopełnienie boolowskie dowolnego elementu, czyli

$$\forall_{a,b}\ (a \vee b = 1,\ a \wedge b = 0) \Longrightarrow \sim a = b.$$

Otóż

$$\begin{aligned}b = b \vee 0 = b \vee (a \wedge (\sim a)) = (b \vee a) \wedge (b \vee (\sim a)) = 1 \wedge (b \vee (\sim a))\\ = (a \vee (\sim a)) \wedge (b \vee (\sim a)) = (a \wedge b) \vee (\sim a) = 0 \vee (\sim a) = \sim a,\end{aligned}$$

zatem $\sim (x \vee y) = (\sim x) \wedge (\sim y)$. Pozostałe prawo de Morgana dowodzi się analogicznie zamieniając $\vee$ na $\wedge$ i odwrotnie. □

Uwaga: W teorii krat i teorii algebr Boole'a zachodzi zasada dualności: Każde ogólne twierdzenie o kratach pozostanie prawdziwe jeśli symbol $\vee$ zamienimy na $\wedge$ i na odwrót (odpowiada to zamianie kierunku nierówności na przeciwny).

Atomem algebry Boole'a $(X, \vee, \wedge, \sim, 0, 1)$ nazywamy taki element $x \in X \setminus \{0\}$, że jeśli $x = y \vee z$, to $x = y$ lub $x = z$ (x jest nierozkładalny).

Twierdzenie 2.4.3. *Niech* x *będzie elementem algebry Boole'a. Następujące warunki są równoważne:*
(*a*) x *jest atomem,*
(*b*) x *jest większy od zera i nie ma elementu pośredniego (*$x > 0$ *oraz* $\neg\exists_{y\in X}\ 0 < y < x$*).*

Dowód.
$(a) \Longrightarrow (b)$ $x > 0$ z definicji. Załóżmy, że istnieje pośredni $y \in X$, tzn. $0 < y < x$. Wtedy

$$x = x \wedge 1 = (y \vee x) \wedge (y \vee (\sim y)) = y \vee (x \wedge (\sim y)).$$

Skoro x jest atomem, to $x = x \wedge (\sim y)$. Zatem

$$y = x \wedge y = (x \wedge (\sim y)) \wedge y = x \wedge 0 = 0.$$

Sprzeczność dowodzi, że x nakrywa 0.
$(b) \Longrightarrow (a)$ Oczywiście $x \neq 0$. Niech $x = y \vee z$. Gdyby x nie był atomem, to moglibyśmy wziąć $y \neq x$, $z \neq x$. Wtedy $y < x = y \vee z$, czyli $y = 0$. Ale to oznacza, że $x = z$. Sprzeczność dowodzi, że x jest atomem. □

Twierdzenie 2.4.4 (o reprezentacji w skończonych algebrach Boole'a)**.** *Niech* $(X, \vee, \wedge, \sim, 0, 1)$ *będzie skończoną algebrą Boole'a, której wszystkimi atomami są* $a_1, \ldots, a_n$*. Wtedy dowolny element* $x \in X$ *można przedstawić jednoznacznie jako sumę boolowską atomów* $x = a_{i_1} \vee \cdots \vee a_{i_k}$ *(*$i_1, \ldots, i_k \in \{1, \ldots, n\}$*). Jednoznaczność przedstawienia jest z dokładnością do kolejności składników.*

Dowód. **Teza 1:** Każdy element jest sumą (być może pustą) atomów.

Element 0 (neutralny dla $\vee$) jest sumą pustej rodziny atomów. Załóżmy, że zbiór

$$S = \{x \in X \mid x \text{ nie jest sumą atomów}\}$$

jest niepusty. Weźmy $x \in S$. Zatem $x \neq 0$ i x nie jest atomem. Istnieje rozkład $x = y \vee z$, w którym $x \neq y$ i $x \neq z$, a więc $y \neq 0$ i $z \neq 0$. Gdyby y i z były sumami atomów, to x byłby sumą atomów. Ale $x \in S$ więc $y \in S$ lub $z \in S$. Istnieje mniejszy od x element zbioru S. Ponawiając to rozumowanie otrzymujemy ciąg

$$x = x_1 > x_2 > x_3 > \cdots$$

różnych elementów należących do skończonego zbioru S. Sprzeczność dowodzi, że $S = \varnothing$.

Teza 2: Dowolny element $x \in X$ jest sumą tych atomów, które są niewiększe od x.

Z Tezy 1 wynika, że 1 jest sumą atomów, więc $1 = a_1 \vee \cdots \vee a_n$. Zatem

$$x = x \wedge 1 = x \wedge (a_1 \vee \cdots \vee a_n) = (x \wedge a_1) \vee \cdots \vee (x \wedge a_n).$$

Zawsze jest $0 \leqslant x \wedge a_i \leqslant a_i$. Skoro a_i są atomami, to jest $x \wedge a_i = 0$ lub $x \wedge a_i = a_i$ dla każdego i. Przy tym $x \wedge a_i = a_i \iff a_i \leqslant x$. Mamy $x = \bigvee\limits_{\substack{a_i \leqslant x \\ a_i \text{ atom}}} a_i$.

Teza 3: Przedstawienie dowolnego $x \in X$ jako suma atomów jest jednoznaczne z dokładnością do kolejności składników.

Weźmy dowolne takie przedstawienie $x = b_1 \vee \cdots \vee b_k$, gdzie b_i są atomami. Wtedy $\forall_{i=1,\ldots,k}$ jest $b_i \leqslant x$. Jeśli a jest atomem i $a \leqslant x$, to

$$0 \neq a = a \wedge x = a \wedge (b_1 \vee \cdots \vee b_k) = (a \wedge b_1) \vee \cdots \vee (a \wedge b_k).$$

Któryś ze składników musi być być różny od zera, czyli dla pewnego i mamy $0 \neq a = a \wedge b_i = b_i$, bo zarówno a jak i b_i są atomami. Zatem atomy $b_1, \ldots, b_k$ są to dokładnie te atomy w X, które są niewiększe od x. Stąd $b_1, \ldots, b_k$ są wyznaczone jednoznacznie z dokładnością do kolejności. □

Niech teraz $(X, \vee, \wedge, \sim, 0, 1)$ i $(X', \vee', \wedge', \sim', 0', 1')$ będą algebrami Boole'a. Odwzorowanie $\varphi\colon X \longrightarrow X'$ nazywamy **homomorfizmem algebr Boole'a** jeśli zachodzą prawa:

$$\begin{aligned}\varphi(x \vee y) &= \varphi(x) \vee' \varphi(y),\\ \varphi(x \wedge y) &= \varphi(x) \wedge' \varphi(y),\\ \varphi(\sim x) &= \sim' \varphi(x),\\ \varphi(0) &= 0',\\ \varphi(1) &= 1'.\end{aligned}$$

Jeśli homomorfizm $\varphi\colon X \longrightarrow X'$ jest bijekcją, to nazywamy go **izomorfizmem**. Odwzorowanie odwrotne do izomorfizmu algebr Boole'a jest izomorfizmem algebr Boole'a. Jeśli istnieje izomorfizm pomiędzy dwoma algebrami Boole'a, to mówimy, że są one **izomorficzne** (są wtedy takie same).

Twierdzenie 2.4.5 (o izomorfizmie)**.** *Niech $(X, \vee, \wedge, \sim, 0, 1)$ i $(X', \vee', \wedge', \sim', 0', 1')$ będą skończonymi algebrami Boole'a ze zbiorami atomów $A = \{a_1, \dots, a_n\}$ i $A' = \{a'_1, \dots, a'_n\}$. Wtedy istnieje izomorfizm algebr Boole'a $\varphi\colon X \longrightarrow X'$ taki, że $\varphi(a_i) = a'_i$ dla $i = 1, \dots, n$.*

Dowód. Odwzorowanie $\varphi\colon X \longrightarrow X'$ definiujemy następująco: jeśli $x = a_{i_1} \vee \dots \vee a_{i_k}$ jako suma atomów, to

$$\varphi(x) = a'_{i_1} \vee' \dots \vee' a'_{i_k}.$$

Z poprzedniego twierdzenia wynika, że φ jest poprawnie określone. Dla dowolnych $a \in A$ i $x \in X$ jest

$$a \leqslant x \iff \varphi(a) \leqslant' \varphi(x).$$

Zauważmy, że

$$\begin{aligned}\varphi(a) \leqslant' \varphi(x \vee y) &\iff a \leqslant x \vee y \iff (a \leqslant x) \vee (a \leqslant y)\\ &\iff (\varphi(a) \leqslant' \varphi(x)) \vee (\varphi(a) \leqslant' \varphi(y)).\end{aligned}$$

Ale $\varphi(A) = A'$, więc $\forall_{a' \in A'}$ jest $a' \leqslant' \varphi(x \vee y) \iff (a' \leqslant' \varphi(x)) \vee (a' \leqslant' \varphi(y))$. Z poprzedniego twierdzenia mamy

$$\varphi(x \vee y) = \varphi(x) \vee' \varphi(y).$$

Jeśli $x = a_{i_1} \vee \dots \vee a_{i_k}$ oraz $\{i_1, \dots, i_k, j_1, \dots, j_l\} = \{1, \dots, n\}$, to

$$\sim x = a_{j_1} \vee \dots \vee a_{j_l}.$$

Podobnie skoro $\varphi(x) = a'_{i_1} \vee' \dots \vee' a'_{i_k}$, to $\sim' \varphi(x) = a'_{j_1} \vee' \dots \vee' a'_{j_l} = \varphi(\sim x)$. Ponadto

$$\begin{aligned}\varphi(x \wedge y) &= \varphi(\sim((\sim x) \vee (\sim y))) = \sim' \varphi((\sim x) \vee (\sim y))\\ &= \sim'(\varphi(\sim x) \vee' \varphi(\sim y)) = (\sim' \varphi(\sim x)) \wedge' (\sim' \varphi(\sim y)) = \varphi(x) \wedge' \varphi(y).\end{aligned}$$

Tożsamość $\varphi(0) = 0'$, $\varphi(1) = 1'$ wynikają z poprzedniego twierdzenia. Odwzorowanie $\varphi\colon X \longrightarrow X'$ jest homomorfizmem algebr Boole'a. Podobnie odwzorowanie $\psi\colon X' \longrightarrow X$ zadane warunkiem $\psi(a'_{i_1} \vee' \dots \vee' a'_{i_k}) = a_{i_1} \vee \dots \vee a_{i_k}$ jest homomorfizmem algebr Boole'a. Ale $\psi \circ \varphi = \mathrm{id}_X$ oraz $\varphi \circ \psi = \mathrm{id}_{X'}$, więc φ i ψ są wzajemnie odwrotnymi izomorfizmami. □

Wniosek 2.4.6. *Każda skończona algebra Boole'a o n atomach jest izomorficzna z algebrą $(\mathcal{P}(\{1, \dots, n\}), \cup, \cap, ', \varnothing, \{1, \dots, n\})$ wszystkich podzbiorów zbioru n-elementowego. Ma więc 2^n elementów.*

Wniosek 2.4.7. *Każda skończona algebra Boole'a mająca n atomów jest izomorficzna z algebrą $\mathbb{B}^n$ (ciągów zero-jedynkowych o długości n)*

2.5. Notacja $O(\ )$ dla ciągów

Ciągiem (nieskończonym o wyrazach rzeczywistych) nazywamy dowolną funkcję $c\colon \mathbb{N} \longrightarrow \mathbb{R}$ (zbiór takich ciągów oznaczamy $\mathbb{R}^{\mathbb{N}}$). Mówimy, że ciąg a **jest** O **od** ciągu b ($a = O(b)$), jeśli istnieje taka stała $C \in \mathbb{R}$, że dla prawie wszystkich $n \in \mathbb{N}$ wartość bezwzględna z $a(n)$ jest mniejsza lub równa stałej C pomnożonej przez wartość bezwzględną z $b(n)$, czyli gdy

$$\exists_{C\in\mathbb{R}}\exists_{N_0\in\mathbb{N}}\forall_{n\geqslant N_0}\ |a(n)| \leqslant C|b(n)|.$$

Zauważmy, że stała C jest zawsze nieujemna. Ponadto, własność „jest O od" nie zależy od skończonej liczby wyrazów obu ciągów.

Twierdzenie 2.5.1. *Dla dowolnych ciągów $a, b, c, d \in \mathbb{R}^{\mathbb{N}}$ zachodzą własności:*
(a) $c = O(c)$,
(b) *jeśli* $a = O(b)$ *i* $b = O(c)$, *to* $a = O(c)$,
(c) *jeśli* $a = O(b)$ *i* $D \in \mathbb{R}$, *to* $D \cdot a = O(b)$,
(d) *jeśli* $a = O(c)$ *i* $b = O(c)$, *to* $a + b = O(c)$,
(e) *jeśli* $a = O(c)$ *i* $b = O(d)$, *to* $a \cdot b = O(c \cdot d)$.

Dowód.
(a) $C = 1$. Zawsze jest $|c(n)| \leqslant |c(n)|$.
(b) Jeśli $\forall_{n\geqslant N_1}\ |a(n)| \leqslant C_1|b(n)|$ oraz $\forall_{n\geqslant N_2}\ |b(n)| \leqslant C_2|c(n)|$, to

$$\forall_{n\geqslant \max\{N_1,N_2\}}\ |a(n)| \leqslant C_1 \cdot C_2|c(n)|.$$

(c) Jeśli $\forall_{n\geqslant N_0}\ |a_n| \leqslant C|b_n|$, to $\forall_{n\geqslant N_0}\ |D \cdot a_n| \leqslant C \cdot D \cdot |b_n|$.
(d) Jeśli $\forall_{n\geqslant N_1}\ |a_n| \leqslant C_1|c_n|$ oraz $\forall_{n\geqslant N_2}\ |b_n| \leqslant C_2|c_n|$, to

$$\forall_{n\geqslant \max\{N_1,N_2\}}\ |a_n + b_n| \leqslant |a_n| + |b_n| \leqslant (C_1 + C_2) \cdot |c_n|.$$

(e) Jeśli $\forall_{n\geqslant N_1}\ |a_n| \leqslant C_1|c_n|$ i $\forall_{n\geqslant N_2}\ |b_n| \leqslant C_2|d_n|$, to

$$\forall_{n\geqslant \max\{N_1,N_2\}}\ |a_n \cdot b_n| \leqslant C_1 \cdot C_2|c_n \cdot d_n|.$$

□

Wobec *(a)* i *(b)* relacja „jest O od" jest preporządkiem na zbiorze $\mathbb{R}^{\mathbb{N}}$.

Mówimy, że ciąg g **jest** Ω **od** f (piszemy $g = \Omega(f)$), jeśli $f = O(g)$. Ponadto mówimy, że h **jest** Θ **od** f, jeśli $h = O(f)$ oraz $h = \Omega(f)$. Wtedy mówimy, że g rośnie nie wolniej niż f, a h rośnie równie szybko co f.

Relacja „jest Θ od" jest równoważnością w $\mathbb{R}^{\mathbb{N}}$ i zbiór klas tej równoważności jest częściowo uporządkowany przez relację

$$[f]_\Theta \leqslant [g]_\Theta \iff f = O(g).$$

Piszemy czasem $f = o(g)$ gdy $[f]_\Theta < [g]_\Theta$.

Twierdzenie 2.5.2. *Dla dowolnych $k, l \in \mathbb{Q}$ (a nawet $\mathbb{R}$) jest:*
(a) $[n^k]_\Theta < [n^l]_\Theta \iff k < l$,
(b) $[n^k]_\Theta < [2^n]_\Theta$,
(c) $[2^n]_\Theta < [3^n]_\Theta$,
(d) $[3^n]_\Theta < [n!]_\Theta$,

(*e*) $[n!]_\Theta < [n^n]_\Theta$,
(*f*) $[\ln n]_\Theta < [n^k]_\Theta$, *jeśli* $k > 0$.

Dowód. Wynika to z równości

$$\lim_{n\to\infty} \frac{n^k}{n^l} \overset{k\le l}{=} \lim_{n\to\infty} \frac{n^k}{2^n} = \lim_{n\to\infty} \frac{2^n}{3^n} = \lim_{n\to\infty} \frac{3^n}{n!} = \lim_{n\to\infty} \frac{n!}{n^n} = \lim_{n\to\infty} \frac{\ln n}{n^k} \overset{k>0}{=} 0. \qquad \square$$

Relacja $\leqslant$ wprowadza na zbiorze $\mathbb{R}^{\mathbb{N}}/\Theta$ strukturę kraty, bo:

$$\begin{aligned} \sup\{[f]_\Theta, [g]_\Theta\} &= [\max\{|f|, |g|\}]_\Theta, \\ \inf\{[f]_\Theta, [g]_\Theta\} &= [\min\{|f|, |g|\}]_\Theta, \\ (\max\{|f|, |g|\}(n) &= \max\{|f(n)|, |g(n)|\}, \\ \min\{|f|, |g|\}(n) &= \min\{|f(n)|, |g(n)|\}). \end{aligned}$$

Relacja $\leqslant$ nie jest porządkiem liniowym, nie ma elementu największego, ale ma element najmniejszy $[0]_\Theta$.

Ćwiczenie: $(\mathbb{R}^{\mathbb{N}}/\Theta, \leqslant)$ jest kratą rozdzielną, ale nie jest algebrą Boole'a.

Rozdział 3

Więcej o liczbach

3.1. Zbiory liczbowe

Zbiorem liczb naturalnych nazywamy taki zbiór $\mathbb{N}$, że
$(A1)$ 0 jest liczbą naturalną ($0 \in \mathbb{N}$),
$(A2)$ każda liczba naturalna n ma dokładnie jeden następnik $S(n)$ w zbiorze liczb naturalnych,
$(A3)$ 0 nie jest następnikiem żadnej liczby naturalnej,
$(A4)$ jeśli $S(n) = S(m)$, to $n = m$,
$(A5)$ (zasada indukcji matematycznej) jeśli $A \subset \mathbb{N}$ jest taki, że $0 \in A$ oraz $\forall_{n\in\mathbb{N}}\ n \in A \Longrightarrow S(n) \in A$, to $A = \mathbb{N}$.
Są to **aksjomaty Peano** ([1]), ogłoszone około roku 1890. Krócej: zbiór liczb naturalnych to system $(\mathbb{N}, S, 0)$, gdzie 0 jest wyróżnionym elementem zbioru $\mathbb{N}$, a $S\colon \mathbb{N} \longrightarrow \mathbb{N}$ jest wyróżnioną injekcją taką, że jej obraz $\operatorname{im} S$ nie zawiera elementu 0 oraz każdy podzbiór zbioru $\mathbb{N}$ zawierający 0 i zamknięty na stosowanie funkcji S jest całym $\mathbb{N}$.

W $\mathbb{N}$ wprowadza się działania $+\colon \mathbb{N}\times\mathbb{N} \longrightarrow \mathbb{N}$, $\cdot\colon \mathbb{N}\times\mathbb{N} \longrightarrow \mathbb{N}$ wzorami:
$(D1)$ $\forall_{n\in\mathbb{N}}\ n + 0 = n$,
$(D2)$ $\forall_{n,m\in\mathbb{N}}\ n + S(m) = S(n + m)$,
$(M1)$ $\forall_{n\in\mathbb{N}}\ n \cdot 0 = 0$,
$(M2)$ $\forall_{n,m\in\mathbb{N}}\ n \cdot S(m) = n \cdot m + n$.
Ponadto wprowadza się relację $n \leqslant m \iff \exists_{k\in\mathbb{N}}\ n + k = m$. Zbiór liczb całkowitych $\mathbb{Z}$ tworzy się ze zbioru liczb naturalnych następująco

$$\mathbb{Z} = \mathbb{N}\times\mathbb{N}/\sim, \quad \text{gdzie } (k,l) \sim (k',l') \iff k + l' = k' + l.$$

Klasę równoważności $[(k,l)]$ utożsamiamy z tradycyjnie rozumianą liczbą całkowitą $k - l$. Zwykle zakłada się, że $\mathbb{N} \subset \mathbb{Z}$ utożsamiając liczbę naturalną n z klasą równoważności $[(n,0)]$. Podobnie $\mathbb{Q}$ konstruuje się z $\mathbb{Z}$, $\mathbb{R}$ konstruuje się z $\mathbb{Q}$ (używając ciągów Cauchy'ego) i $\mathbb{C}$ konstruuje się z $\mathbb{R}$.

3.2. Zasada indukcji matematycznej

Twierdzenie 3.2.1 (pierwsza wersja zasady indukcji matematycznej)**.** *Niech $m \in \mathbb{Z}$ oraz niech dany będzie ciąg zdań logicznych $(p(n))_{n\geqslant m, n\in\mathbb{Z}}$. Jeśli*
(i) $p(m)$ jest prawdziwe,
(ii) dla dowolnego $k > m$ zachodzi implikacja $p(k-1) \Longrightarrow p(k)$,
to wszystkie zdania $p(n)$ $(n \geqslant m)$ są prawdziwe.

[1] matematyk włoski Giuseppe Peano (1858 – 1932)

Twierdzenie 3.2.2 (druga wersja zasady indukcji matematycznej)**.** *Niech $m \in \mathbb{Z}$ oraz niech dany będzie ciąg zdań logicznych $(p(n))_{n \geqslant m}$. Jeśli*
(i) $p(m)$ jest prawdziwe,
(ii) dla dowolnego $k > m$ zachodzi implikacja $p(m) \wedge \cdots \wedge p(k-1) \Longrightarrow p(k)$,
to wszystkie zdania $p(n)$ $(n \geqslant m)$ są prawdziwe.

Twierdzenie 3.2.3 (trzecia wersja zasady indukcji matematycznej)**.** *Niech $m \in \mathbb{Z}$ oraz niech dany będzie ciąg zdań logicznych $(p(n))_{n \geqslant m}$. Załóżmy, że $l \in \mathbb{N}$ oraz*
(i) zdania $p(m), \ldots, p(m+l)$ są prawdziwe,
(ii) dla dowolnego $k > m + l$ zachodzi implikacja $p(m) \wedge \cdots \wedge p(k-1) \Longrightarrow p(k)$,
to wszystkie zdania $p(n)$ $(n \geqslant m)$ są prawdziwe.

Twierdzenie 3.2.4. *Każda z powyższych wersji Zasady Indukcji Matematycznej jest równoważna następującej Zasadzie Minimum zwanej też Zasadą Dobrego Uporządkowania:*
(Z.M.) Każdy niepusty podzbiór zbioru $\mathbb{N}$ ma element najmniejszy.

Dowód.
(3.2.3) $\Longrightarrow$ (3.2.2) Wystarczy przyjąć $l = 0$.
(3.2.2) $\Longrightarrow$ (3.2.1) Jeśli zachodzi $p(k-1) \Longrightarrow p(k)$, to tym bardziej zachodzi $p(m) \wedge \cdots \wedge p(k-1) \Longrightarrow p(k)$. Druga wersja Z.I.M. pociąga więc pierwszą wersję.
(3.2.1) $\Longrightarrow$ (*Z.M.*) Rozważmy niepusty podzbiór A w $\mathbb{N}$. Przyjmijmy $m = 0$ oraz

$$p(n) = \text{„zbiór } A \text{ jest rozłączny ze zbiorem } \{0, 1, \ldots, n\}\text{”}.$$

Zauważmy, że albo 0 jest elementem najmniejszym w A albo $p(0)$ jest prawdziwe. Gdyby zasada minimum nie zachodziła, to $p(0)$ byłoby prawdziwe oraz dla dowolnego $k > 0$ zachodziłaby implikacja

$$p(k-1) \iff A \cap \{0, \ldots, k-1\} = \varnothing \Longrightarrow A \cap \{0, \ldots, k\} = \varnothing \iff p(k)$$

(k nie jest elementem najmniejszym zbioru A). Z twierdzenia 3.2.1 wszystkie zdania $p(n)$ $(n \in \mathbb{N})$ byłyby prawdziwe. W szczególności $\forall_{n \in \mathbb{N}}\ n \notin A$, czyli A byłby pusty. Sprzeczność dowodzi Zasady Minimum.
(*Z.M.*) $\Longrightarrow$ (3.2.3) Gdyby nie wszystkie zdania $p(n)$ w twierdzeniu 3.2.3 były prawdziwe, to niepusty zbiór

$$A = \{n - m \mid p(n) \text{ jest fałszywe}\} \subset \mathbb{N}$$

miałby element najmniejszy z Zasady Minimum. Nazwijmy ten element s. Na pewno $s > l$. Zatem zdania $p(m), \ldots, p(m+s-1)$ są prawdziwe, a zdanie $p(m+s)$ jest fałszywe. Sprzeczność z założeniem *(ii)* twierdzenia 3.2.3 dla $k = m + s$ dowodzi, że twierdzenie 3.2.3 jest prawdziwe. □

3.3. Równania rekurencyjne

Często w praktyce ciągi są zadane rekurencyjnie. Pojawia się więc klasa problemów: dla danego ciągu zadanego rekurencyjnie znaleźć jego wzór jawny. Czasem dla zastosowań wystarcza znać tylko tempo wzrostu.

3.3.A. Rodzaj pierwszy: równania liniowe

Równanie rekurencyjne

$$s_n = a_1 s_{n-1} + \cdots + a_k s_{n-k} \text{ dla } n \geqslant k$$

($k \in \mathbb{N}$ jest **stopniem równania**; $a_1, \ldots, a_k \in \mathbb{R}$ lub $\mathbb{C}$ jego **współczynnikami**).

Dane początkowe to:

$$\begin{aligned} s_0 &= b_1 \\ s_1 &= b_2 \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ s_{k-1} &= b_k, \quad \text{gdzie } b_1, \ldots, b_k \in \mathbb{R} \vee \mathbb{C}. \end{aligned}$$

To równanie jest **liniowe jednorodne** (kombinacja liniowa rozwiązań jest rozwiązaniem) oraz zadane warunki początkowe pozwalają wyznaczyć ciąg jednoznacznie. Stosujemy tzw. **metodę przewidywań** znaną z teorii równań różniczkowych. Przewidujemy rozwiązanie równania w postaci

$$s_n = d^n, \quad \text{gdzie} \quad d \in \mathbb{R} \vee \mathbb{C}.$$

Wtedy

$$\begin{aligned} d^n &= a_1 d^{n-1} + a_2 d^{n-2} + \cdots + a_k d^{n-k} \qquad / : d^{n-k}, \\ d^k &= a_1 d^{k-1} + a_2 d^{k-2} + \cdots + a_k d^0, \end{aligned}$$

zatem

$$d^k - a_1 d^{k-1} - \cdots - a_{k-1} d - a_k = 0.$$

Wielomian $W(x) = x^k - a_1 x^{k-1} - \cdots - a_k$ nazywamy **wielomianem charakterystycznym** równania rekurencyjnego, a równanie $W(x) = 0$ jego **równaniem charakterystycznym**. Jeśli pierwiastki wielomianu $W(x)$ są jednokrotne, to otrzymujemy k rozwiązań szczególnych równania (RSRLJ): $s_n = d_1^n, s_n = d_2^n, \ldots, s_n = d_k^n$, gdzie $d_1, \ldots, d_k$ są wszystkimi (być może zespolonymi) pierwiastkami wielomianu $W(x)$.

Rozwiązaniem ogólnym równania jest kombinacja liniowa tych rozwiązań szczególnych (RORLJ): $s_n = C_1 d_1^n + \cdots + C_k d_k^n$, gdzie $C_1, \ldots, C_k$ są dowolnymi stałymi. Następnie stałe $C_1, \ldots, C_k$ obliczamy podstawiając rozwiązanie ogólne do danych początkowych. Mamy:

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix} = \begin{bmatrix} 1 & \cdots & 1 \\ d_1 & \cdots & d_k \\ \vdots & \ddots & \vdots \\ d_1^{k-1} & \cdots & d_k^{k-1} \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_k \end{bmatrix}$$

Ponieważ liczby $d_1, \ldots, d_k$ są różne, to wyznacznik macierzy kwadratowej (**wyznacznik Vandermonde'a**) jest różny od zera: $\det M_k = V_k(d_1, \ldots, d_k) = \prod\limits_{\substack{i,j=1 \\ i<j}}^{k} (d_j - d_i) \neq 0$. Istnieje dokładnie jeden układ stałych $C_1, \ldots, C_k$ rozwiązujący zadany problem w tym przypadku.

Przykłady 3.3.1.

$$\begin{cases} s_n = 3s_{n-1} - 2s_{n-2}, \\ s_0 = 2, \ s_1 = 3. \end{cases}$$

Równaniem charakterystycznym jest $x^2 - 3x + 2 = 0$, $\Delta = 9 - 4 \cdot 2 = 1$, $\sqrt{\Delta} = 1$, *czyli* $x_1 = \frac{3-1}{2} = 1$, $x_2 = \frac{3+1}{2} = 2$. *Rozwiązanie ogólne ma postać*

$$s_n = C_1 \cdot 1^n + C_2 \cdot 2^n, \qquad C_1, C_2 \in \mathbb{R}.$$

Ponadto

$$\begin{cases} 2 = C_1 + C_2, \\ 3 = C_1 + 2C_2, \end{cases} \qquad \textit{czyli} \qquad \begin{cases} C_1 = 1, \\ C_2 = 1. \end{cases}$$

Rozwiązaniem problemu jest ciąg $s_n = 2^n + 1$.

Przykłady 3.3.2 (liczby Fibonacciego)**.**

$$\begin{cases} s_n = s_{n-1} + s_{n-2}, \\ s_0 = 1,\ s_1 = 1. \end{cases}$$

Równaniem charakterystycznym jest $x^2 - x - 1 = 0$, $\Delta = 1 + 4 = 5$, $\sqrt{\Delta} = \sqrt{5}$, *czyli* $d_1 = x_1 = \frac{1-\sqrt{5}}{2}$, $d_2 = x_2 = \frac{1+\sqrt{5}}{2}$. *Rozwiązanie ogólne ma postać*

$$s_n = C_1\Big(\frac{1-\sqrt{5}}{2}\Big)^n + C_2\Big(\frac{1+\sqrt{5}}{2}\Big)^n, \qquad C_1, C_2 \in \mathbb{R}.$$

Ponadto

$$\begin{cases} 1 = C_1 + C_2, \\ 1 = C_1\frac{1-\sqrt{5}}{2} + C_2\frac{1+\sqrt{5}}{2}, \end{cases} \qquad \textit{czyli} \qquad \begin{cases} C_1 = \frac{-1+\sqrt{5}}{2\sqrt{5}} = \frac{-d_1}{\sqrt{5}}, \\ C_2 = \frac{1+\sqrt{5}}{2\sqrt{5}} = \frac{d_2}{\sqrt{5}}. \end{cases}$$

Rozwiązaniem jest $s_n = \frac{1}{\sqrt{5}}\big((\frac{1+\sqrt{5}}{2})^{n+1} - (\frac{1-\sqrt{5}}{2})^{n+1}\big)$.

Jeśli któryś z pierwiastków d_i wielomianu charakterystycznego jest wielokrotny (krotności $l > 1$), to rozwiązaniami szczególnymi związanymi z tym pierwiastkiem są (RSRLJ): $s_n = d_i^n, s_n = n \cdot d_i^n, \ldots, s_n = n^{l-1} \cdot d_i^n$ (l rozwiązań szczególnych). Wtedy rozwiązanie ogólne zawiera kombinacje liniowe tych rozwiązań szczególnych. Analogicznie znajduje się stałe $C_1, \ldots, C_k$.

Przykłady 3.3.3.

$$\begin{cases} s_n = 4s_{n-1} - 4s_{n-2}, \\ s_0 = 5,\ s_1 = 8. \end{cases}$$

Równanie charakterystyczne $x^2 - 4x + 4 = 0$, $x_1 = x_2 = 2$. *Rozwiązanie ogólne ma postać*

$$s_n = C_1 \cdot 2^n + C_2 \cdot n \cdot 2^n, \qquad C_1, C_2 \in \mathbb{R}.$$

Ponadto warunki początkowe dają

$$\begin{cases} 5 = C_1 + 0, \\ 8 = C_1 \cdot 2 + C_2 \cdot 2, \end{cases} \qquad \textit{czyli} \qquad \begin{cases} C_1 = 5, \\ C_2 = -1. \end{cases}$$

Rozwiązaniem jest ciąg $s_n = 5 \cdot 2^n - n \cdot 2^n$.

Równanie liniowe niejednorodne o stałych współczynnikach rozwiązujemy korzystając z teorii równań liniowych. Jeśli równanie ma postać (RLN):

$$s_n = a_1 s_{n-1} + \cdots + a_k s_{n-k} + f(n), \qquad a_i \in \mathbb{R} \vee \mathbb{C},$$

gdzie f jest znaną funkcją, to

$$\text{RORLN} = \text{RORLJ} + \text{RSRLN}.$$

Rozwiązanie szczególne RLN znajdujemy metodą przewidywań np. według poniższej tabeli, zastepując konkretne stałe $c, c_0, ..., c_m$ stałymi nieoznaczonymi $C, C_0, ..., C_m$.

$f(n)$	RSRLN
$c_0 + c_1 n + \cdots + c_m n^m$	$C_0 + C_1 n + \cdots + C_m n^m$
$c \cdot d^n$	$C \cdot d^n$

Uwaga 3.3.4. *Tutaj do rozwiązywania równań rekurencyjnych stosowaliśmy* ***metodę wielomianu charakterystycznego****. Inną metodą jest* ***metoda funkcji tworzących****, o której można przeczytać w [B] (uwaga na błędy).*

3.3.B. Rodzaj drugi: niektóre równania sprowadzalne do liniowych

Równanie rekurencyjne zadajemy jako

$$s_{2n} = 2 \cdot s_n + f(n),$$

gdzie $f\colon \mathbb{N} \setminus \{0\} \longrightarrow \mathbb{R}$ jest znaną funkcją. Równanie to wiąże ze sobą tylko te wyrazy, których iloraz numerów jest potęgą liczby 2. Znajomość skończenie wielu początkowych wyrazów nie pozwala poznać całego ciągu. Chemy znać tylko tempo wzrostu szukanego ciągu zakładając, że jest on od pewnego momentu monotoniczny. Znajdziemy tylko wzór jawny na s_n, gdzie $n = 2^m$ dla pewnego $m \in \mathbb{N}$ (to wystarcza, bo $\lim\limits_{m\to\infty} 2^n = +\infty$). Mamy

$$s_{2^{m+1}} = 2s_{2^m} + f(2^m).$$

Lemat 3.3.5. *Zachodzi wzór* $s_{2^m} = 2^m(s_1 + \frac{1}{2}\sum\limits_{i=0}^{m-1} \frac{f(2^i)}{2^i})$, $m \in \mathbb{N}$.

Dowód. Dla $m = 0$ jest $s_1 = 1\cdot(s_1+0) = s_1$. Dla $m = 1$ jest $s_2 = 2\cdot(s_1+\frac{1}{2}\frac{f(1)}{1}) = 2s_1+f(1)$. Przyjmijmy teraz, że wzór zachodzi dla $m-1$. Wtedy

$$\begin{aligned} s_{2^m} &= 2 \cdot s_{2^{m-1}} + f(2^{m-1}) = 2 \cdot 2^{m-1}\Big(s_1 + \frac{1}{2}\sum_{i=0}^{m-2} \frac{f(2^i)}{2^i}\Big) + f(2^{m-1}) \\ &= 2^m\Big(s_1 + \frac{1}{2}\sum_{i=0}^{m-2} \frac{f(2^i)}{2^i}\Big) + 2^m \cdot \frac{1}{2} \cdot \frac{f(2^{m-1})}{2^{m-1}} \\ &= 2^m\Big(s_1 + \frac{1}{2}\sum_{i=0}^{m-1} \frac{f(2^i)}{2^i}\Big). \end{aligned}$$

□

Lemat 3.3.6. *Jeśli* $f(n) = An + B$, *to* $s_{2^m} = 2^m \cdot s_1 + 2^m \cdot m \cdot \frac{A}{2} + (2^m - 1)B$.

Dowód. $f(2^i) = A \cdot 2^i + B$, $i \in \mathbb{N}$.

$$\begin{aligned} s_{2^m} &= 2^m\Big(s_1 + \frac{1}{2}\sum_{i=0}^{m-1} \frac{2^i A + B}{2^i}\Big) = 2^m s_1 + 2^m \frac{A}{2} m + 2^m \frac{1}{2}\Big(B + \frac{B}{2} + \cdots + \frac{B}{2^{m-1}}\Big) \\ &= 2^m s_1 + 2^m m \frac{A}{2} + (2^m - 1)B. \end{aligned}$$

□

Wniosek 3.3.7. *Jeśli s_n jest ostatecznie monotonicznym rozwiązaniem równania rekurencyjnego drugiego rodzaju, to $s_n = O(n \cdot \log_2 n)$, gdy f jest liniowa, oraz $s_n = O(n)$, gdy f jest stała.*

Dowód. Z lematu 3.3.6 dla $n = 2^m$ mamy

$$s_n = ns_1 + n\frac{A}{2}\log_2 n + (n-1)B.$$

Jeśli $A \neq 0$, to $s_n = O(n) + O(n\log_2 n) + O(n) = O(n\log_2 n)$. Jeśli $A = 0$, to $s_n = O(n) + O(n) = O(n)$. Ogólniej, dla $2^m < n < 2^{m+1}$ dostatecznie dużego jest:
$A \neq 0 \Longrightarrow s_{2^m} \leqslant s_n \leqslant s_{2^{m+1}} = O(2 \cdot 2^m \cdot (m+1)) = O(2^m \cdot m) = O(n\log_2 n)$.
$A = 0 \Longrightarrow s_{2^m} \leqslant s_n \leqslant s_{2^{m+1}} = O(2 \cdot 2^m) = O(2^m) = O(n)$. □

3.4. Zliczanie elementów zbiorów skończonych

Lemat 3.4.1. *Niech S i T będą zbiorami skończonymi. Wtedy:*
(a) $|S \cup T| = |S| + |T| - |S \cap T|$,
(b) $|S \times T| = |S| \cdot |T|$,
(c) $|\{f\colon S \longrightarrow T\}| = |T|^{|S|}$,
(d) $|\{f\colon S \longrightarrow T \mid f \text{ jest injekcją}\}| = \begin{cases} \frac{|T|!}{(|T|-|S|)!}, & |S| \leqslant |T| \\ 0, & |S| > |T| \end{cases}$,
(e) $|\{f\colon S \longrightarrow T \mid f \text{ jest bijekcją}\}| = \begin{cases} |T|!, & |S| = |T| \\ 0, & |S| \neq |T| \end{cases}$,
(f) $|\mathcal{P}(S)| = 2^{|S|}$,
(g) $|\{P \subset S \mid |P| = k\}| = \binom{|S|}{k} = \frac{|S|!}{k!(|S|-k)!}$ *dla* $0 \leqslant k \leqslant |S|$.

Dowód. (a) Oczywiście $|S \cup T| = |S| + |T|$, o ile $S \cap T = \varnothing$. Ogólnie: $|S \cup T| = |S \setminus T| + |S \cap T| + |T \setminus S| + |T \cap S| - |T \cap S| = |S| + |T| - |S \cap T|$.
(b) Możemy założyć, że $S = \{1, ..., |S|\}$ oraz $T = \{1, ..., |T|\}$. Wystarczy zauważyć, że istnieje bijekcja $B : \{1, ..., |S|\} \times \{1, ..., |T|\} \to \{1, ..., |S| \cdot |T|\}$ zadana wzorem $B(k, l) = (k-1) \cdot |T| + l$.
(c) Niech $S = \{s_1, ..., s_n\}$, gdzie $n = |S|$. Zauważmy, że istnieje bijekcja pomiędzy rodziną funkcji z S do T, a rodziną ciągów długości n elementów z T postaci

$$\{f : S \to T\} \ni g \mapsto (g(s_1), ..., g(s_n)) \in T \times ... \times T = T^n.$$

Stąd, wobec (b), jest $|\{f : S \to T\}| = |T^n| = |T|^n = |T|^{|S|}$.
(d) Niech $n = |S| \leqslant |T|$ i $S = \{s_1, ..., s_n\}$. Element $f(s_1)$ można wybrać na $|T|$ sposobów, element $f(s_2)$ na $|T|-1$ sposobów, ..., $f(s_n)$ na $|T|-|S|+1$ sposobów. Zatem $|\{f\colon S \longrightarrow T \mid f \text{ jest injekcją}\}| = |T| \cdot (|T|-1) \cdot ... \cdot (|T|-|S|+1) = \frac{|T|!}{(|T|-|S|)!}$.
Jeżeli $|S| > |T|$, to nie ma żadnej injekcji z S do T.
(e) Jeśli $|S| = |T|$, to jest to szczególny przypadek punktu (d), bo każda injekcja jest wtedy bijekcją (obraz injekcji jest całym T). Zatem $|\{f\colon S \longrightarrow T \mid f \text{ jest bijekcją}\}| = |T|!$.
Jeśli $|S| \neq |T|$, to nie ma żadnej bijekcji z S do T.

(f) Istnieje bijekcja $\mathcal{P}(S) \ni A \mapsto (a_1, ..., a_n) \in \{f : S \to \{0,1\}\}$ zadana wzorem

$$a_i = \begin{cases} 0, & \text{gdy } s_i \notin A, \\ 1, & \text{gdy } s_i \in A, \end{cases} \text{ gdzie } i = 1, ..., n \text{ i } S = \{s_1, ..., s_n\}.$$

Zatem (f) wynika z (c) dla $T = \{0,1\}$.

(g) Wybór podzbioru o k elementach z S polega na wyborze injekcji ze zbioru $\{1, ..., k\}$ do S i zaniedbaniu kolejności pojawiania się elementów S jako wartości funkcji. Zatem, wobec (d) i (e), zachodzi $|\{P \subset S \mid |P| = k\}| = \frac{1}{k!} \cdot \frac{|S|!}{(|S|-k)!} = \binom{|S|}{k}$ dla $0 \leqslant k \leqslant |S|$.

□

Twierdzenie 3.4.2 (Zasada włączania i wyłączania)**.** *Niech dane będą: $k \in \mathbb{N}^*$ oraz zbiory skończone $A_1, \ldots, A_k$. Wtedy*

$$|\bigcup_{i=1}^{k} A_i| = \sum_{i=1}^{k} |A_i| - \sum_{\substack{i_1,i_2=1 \\ i_1<i_2}}^{k} |A_{i_1} \cap A_{i_2}| + \sum_{\substack{i_1,i_2,i_3=1 \\ i_1<i_2<i_3}}^{k} |A_{i_1} \cap A_{i_2} \cap A_{i_3}| + \cdots + (-1)^{k-1} |A_1 \cap \cdots \cap A_k|.$$

Dowód. Indukcja na ilość zbiorów k

Dla $k = 1$ jest $|A_1| = |A_1|$.

Dla $k = 2$ jest $|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|$ (wobec (a) w lemacie 3.4.1).

Krok indukcyjny: zakładamy, że wzór jest prawdziwy dla $k-1$. Zachodzi równość zbiorów: $A_k \cap (\bigcup_{i=1}^{k-1} A_i) = \bigcup_{i=1}^{k-1} (A_k \cap A_i)$. Korzystajac z założenia indukcyjnego oraz przypadku $k = 2$ mamy

$$|\bigcup_{i=1}^{k} A_i| = |(\bigcup_{i=1}^{k-1} A_i) \cup A_k| = \sum_{i=1}^{k-1} |A_i| - \sum_{\substack{i_1,i_2=1 \\ i_1<i_2}}^{k-1} |A_{i_1} \cap A_{i_2}| + ... + (-1)^{k-2} |A_1 \cap ... \cap A_{k-1}| +$$

$$+ |A_k| - \left(\sum_{i=1}^{k-1} |A_i \cap A_k| - \sum_{\substack{i_1,i_2=1 \\ i_1<i_2}}^{k-1} |A_{i_1} \cap A_{i_2} \cap A_k| + ... + (-1)^{k-2} |A_1 \cap ... \cap A_{k-1} \cap A_k| \right) =$$

$$\sum_{i=1}^{k} |A_i| - \sum_{\substack{i_1,i_2=1 \\ i_1<i_2}}^{k} |A_{i_1} \cap A_{i_2}| + \sum_{\substack{i_1,i_2,i_3=1 \\ i_1<i_2<i_3}}^{k} |A_{i_1} \cap A_{i_2} \cap A_{i_3}| + \cdots + (-1)^{k-1} |A_1 \cap \cdots \cap A_k|.$$

□

Twierdzenie 3.4.3. *Dla $n \in \mathbb{N}^*$ oraz $r \in \mathbb{N}$, $r < n$ zachodzą wzory:*

(a) $\binom{n}{r} = \binom{n}{n-r}$,

(b) $\sum_{i=0}^{n} \binom{n}{i} = 2^n$,

(c) $\binom{n}{r} + \binom{n}{r+1} = \binom{n+1}{r+1}$,

(d) $\forall_{a,b \in \mathbb{R}} \ (a+b)^n = \sum_{l=0}^{n} \binom{n}{l} a^l b^{n-l}$.

Dowód. (a) Wynika z definicji symboli Newtona.

(b) Wynika z (f) i (g) w lemacie 3.4.1.

(c) Każdy $(r+1)$-elementowy pozdbiór w zbiorze $(n+1)$-elementowym jest dokładnie jednego z dwóch rodzajów:
1) zawiera ostatni element zbioru; takich podzbiorów jest tyle samo, co r-elementowych podzbiorów zbioru n-elementowego,
2) nie zawiera ostatniego elementu zbioru; takie podzbiory to dokładnie $(r+1)$-elementowe podzbiory zbioru n-elementowego (nie zawierającego ostatniego elementu).
Wystarcza teraz zastosować (g) z lematu 3.4.1.

(d) Indukcja na n.
Dla $n=1$ jest $(a+b)^1 = a+b = \binom{1}{1}a^1b^0 + \binom{1}{0}a^0b^1$.
Krok indukcyjny: załóżmy, że wzór zachodzi dla pewnego $k \geqslant 1$. Wtedy dla $k+1$ mamy:

$$(a+b)^{k+1} = (a+b)^k(a+b) = \left(\sum_{l=0}^{k}\binom{k}{l}a^lb^{k-l}\right)\cdot(a+b) = \sum_{l=0}^{k}\binom{k}{l}a^{l+1}b^{k-l} +$$
$$\sum_{l=0}^{k}\binom{k}{l}a^lb^{k+1-l} = \sum_{m=1}^{k+1}\binom{k}{m-1}a^mb^{k+1-m} + \sum_{m=0}^{k}\binom{k}{m}a^mb^{k+1-m} =$$
$$\binom{k}{k}a^{k+1}b^0 + \sum_{m=1}^{k}\left(\binom{k}{m-1}+\binom{k}{m}\right)a^mb^{k+1-m} + \binom{k}{0}a^0b^{k+1} =$$
$$\binom{k+1}{k+1}a^{k+1}b^0 + \sum_{m=1}^{k}\binom{k+1}{m}a^mb^{k+1-m} + \binom{k+1}{0}a^0b^{k+1} =$$
$$\sum_{m=0}^{k+1}\binom{k+1}{m}a^mb^{k+1-m}.$$

□

Twierdzenie 3.4.4. *Niech $n,k \in \mathbb{N}^*$. Jest $\binom{n+k-1}{k-1}$ sposobów rozmieszczenia n identycznych przedmiotów w k różnych pudełkach. Inaczej: Jest $\binom{n+k-1}{k-1}$ sposobów przedstawienia liczby n jako sumy uporządkowanej k składników naturalnych.*

Dowód. Numerujemy pudełka liczbami $1,\ldots,k$. Pomiędzy sąsiednie pudełka wsadzamy przegródki (jest ich $k-1$). Rozmieszczenie n jednakowych przedmiotów w k pudełkach jest równoważne wyborowi $k-1$ przegródek ze zbioru $(n+k-1)$-elementowego przedmiotów i przegródek. □

Uporządkowanym podziałem zbioru A na k podzbiorów nazywamy taką uporządkowaną k-tkę $(A_1,\ldots,A_k)$ podzbiorów zbioru A, że $A = \bigcup_{i=1}^{k} A_i$ oraz $A_i \cap A_j = \varnothing$ dla $i \neq j$.

Twierdzenie 3.4.5. *Niech dane będzie przedstawienie $n = n_1 + \cdots + n_k$ liczby naturalnej n jako sumy składników naturalnych. Wtedy liczba podziałów uporządkowanych $(A_1,...,A_k)$ n-elementowego zbioru A spełniających warunki $|A_1| = n_1,\ldots,|A_k| = n_k$ wynosi $\frac{n!}{n_1!\cdots n_k!}$.*

Dowód. Wybór podziału uporządkowanego $(A_1,\ldots,A_k)$ polega na wyborze podzbioru A_1 spośród n elementów, następnie podzbioru A_2 spośród pozostałych $n-n_1$ elementów, itd.,

wreszcie wyboru A_k spośród pozostałych $n_k = n - n_1 - \cdots - n_{k-1}$ elementów. Wybory te są niezależne, więc ich liczba wynosi

$$\binom{n}{n_1}\binom{n-n_1}{n_2}\cdots\binom{n_k}{n_k} = \frac{n!(n-n_1)!\cdots n_k!}{n_1!(n-n_1)!n_2!(n-n_1-n_2)!\cdots n_k!n_k!0!} = \frac{n!}{n_1!\cdots n_k!}. \quad \square$$

Twierdzenie 3.4.6 (Zasada Szufladkowa Dirichleta)**.**

(*a*) *Jeśli skończony zbiór S jest podzielony na k podzbiorów, to któryś z podzbiorów ma co najmniej $\frac{|S|}{k}$ elementów.*

(*b*) *Niech S, T będą zbiorami skończonymi i niech $f\colon S \longrightarrow T$ będzie surjekcją. Załóżmy, że $|S| > r \cdot |T|$ dla pewnego $r \in \mathbb{R}$. Wtedy istnieje taki element $t \in T$, że $|f^{-1}(t)| > r$.*

Dowód.

(*a*) Nie wprost. Jeśli każdy podzbiór ma mniej niż $\frac{|S|}{k}$ elementów, to w S jest mniej niż $\frac{|S|}{k} \cdot k = |S|$ elementów, sprzeczność.

(*b*) Zauważmy, że $T \neq \varnothing$. Włókna (przeciwobrazy) elementów zbioru T tworzą podział S na $|T|$ podzbiorów. Któryś z tych podzbiorów ma co najmniej $\frac{|S|}{|T|}$ elementów wobec punktu (*a*). Zatem ma on więcej niż r elementów. $\square$

Twierdzenie 3.4.7 (Uogólniona Zasada Szufladkowa Dirichleta)**.** *Niech $k \in \mathbb{N}^*$ i niech $A_1, \ldots, A_k \subset S$ będą zbiorami skończonymi. Załóżmy, że każdy element S należy do co najmniej l zbiorów spośród $A_1, \ldots, A_k$. Wtedy średnia arytmetyczna mocy zbiorów A_i ($i = 1, \ldots, k$) wynosi co najmniej $\frac{l}{k}|S|$.*

Dowód. Niech

$$I = \{(s,i) \in S \times \{1,\ldots,k\} \mid s \in A_i\} \text{ (zbiór incydencji).}$$

Wtedy $|I| = \sum\limits_{i=1}^{k} |\{s \in S \mid s \in A_i\}| = \sum\limits_{i=1}^{k} |A_i|$ oraz $|I| = \sum\limits_{s\in S} |\{i \in \{1,\ldots,k\} \mid s \in A_i\}| \geqslant |S| \cdot l$. Mamy $l \cdot |S| \leqslant \sum\limits_{i=1}^{k} |A_i|$, czyli $\frac{l}{k}|S| \leqslant \frac{1}{k}\sum\limits_{i=1}^{k} |A_i|$. $\square$

3.5. Teoria podzielności w $\mathbb{Z}$

Niech $a, b, c \in \mathbb{Z}$. Mówimy, że a **dzieli** b (piszemy $a \mid b$) jeśli $a \cdot k = b$ dla pewnego $k \in \mathbb{Z}$. Wtedy liczbę a nazywamy **dzielnikiem** liczby b, natomiast b nazywamy **wielokrotnością** liczby a.

Jeśli $b \mid a - c$, to mówimy, że a **przystaje do** c **modulo** b (zapisujemy $a \equiv c \pmod{b}$). Relacja podzielności na $\mathbb{N}^*$ jest zgodna z relacją nierówności $\leqslant$, tzn. $a \mid b \Longrightarrow a \leqslant b$.

Lemat 3.5.1. *Dla dowolnych $a, b, c, d \in \mathbb{Z}$ zachodzą własności:*

(1) $(a \mid b \wedge a \mid c) \Longrightarrow (a \mid b \pm c \wedge a \mid b \cdot d)$,

(2) $(a \mid b \wedge b \mid a) \Longrightarrow (a = b \vee a = -b)$,

(3) $(a \mid b \wedge b \neq 0) \Longrightarrow |a| \leqslant |b|$,

(4) $(a \mid b \wedge b \mid c) \Longrightarrow a \mid c$.

Dowód. Ćwiczenie. $\square$

Twierdzenie 3.5.2 (o dzieleniu z resztą)**.** *Jeśli $a \in \mathbb{N}^*$, to dla dowolnego $b \in \mathbb{Z}$ istnieje dokładnie jedna para $(q, r) \in \mathbb{Z} \times \mathbb{Z}$ spełniająca warunek $b = a \cdot q + r$ i $0 \leqslant r < a$.*

Dowód. **Istnienie:** Niech

$$\mathcal{Z} = \{k \in \mathbb{Z} \mid b - a \cdot k \geqslant 0\}.$$

Zbiór $\mathcal{Z}$ jest ograniczony od góry przez $|b|$, więc ma on element największy q (zasada minimum). Niech $r = b - aq$. Wtedy $r \geqslant 0$, bo $q \in \mathcal{Z}$. Gdyby $r \geqslant a$, to $b - aq - a \geqslant 0$ oraz $q + 1 \in \mathcal{Z}$. Sprzeczność dowodzi, że $r < a$.

Jednoznaczność: Jeśli $b = aq_1 + r_1 = aq_2 + r_2$, to $a \mid r_1 - r_2$ ale $|r_1 - r_2| \leqslant a - 1 < a$. Wobec lematu 3.5.1 (3) jest $r_1 - r_2 = 0$, czyli $r_1 = r_2$. Stąd $aq_1 = b - r_1 = b - r_2 = aq_2$ i $q_1 = q_2$. □

Przykłady 3.5.3 (kod korygujący ISBN-10)**.** *W bibliotekarstwie stosuje się kod* ISBN *(International Standard Book Number) identyfikujący książki wydane we wszystkich krajach stosujących ten kod. W celu zapobiegania pomyłkom ostatnia (dziesiąta) „cyfra" kodu jest cyfrą kontrolną i jest wyznaczona wzorem*

$$c_{10} \equiv c_1 \cdot 1 + c_2 \cdot 2 + c_3 \cdot 3 + \cdots + c_8 \cdot 8 + c_9 \cdot 9 \pmod{11}.$$

Jeśli ta reszta wynosi 10*, to używamy „cyfry" X. W ten sposób zapobiegamy większości pomyłek przy przepisywaniu kodów* ISBN-10*. (Ostatnio używa się już nowej wersji kodu korygującego* ISBN-13*.)*

Dzielnikiem właściwym liczby $a \in \mathbb{N}^*$ nazywamy taki dzielnik $d \in \mathbb{N}$, że $d \neq 1$ i $d \neq a$ (czyli $1 < d < a$). Liczbę $p \in \mathbb{N} \setminus \{0, 1\}$ nazywamy **pierwszą**, gdy nie ma ona dzielników właściwych. Liczby naturalne większe od 1, które nie są pierwsze nazywamy **złożonymi**.

Lemat 3.5.4. *Niech $a \in \mathbb{N} \setminus \{0, 1\}$. Jeśli a jest liczbą złożoną, to najmniejszy dzielnik właściwy d liczby a jest liczbą pierwszą i spełnia nierówność $1 < d \leqslant \sqrt{a}$.*

Dowód. Gdyby dzielnik d nie był liczbą pierwszą, to miałby dzielnik właściwy d_1, gdzie $1 < d_1 < d$. Wtedy $d_1 \mid a$ i d_1 jest mniejszym od d dzielnikiem właściwym liczby a, sprzeczność. Weźmy teraz $k \in \mathbb{Z}$ takie, że $d \cdot k = a$. Wtedy $d \leqslant k < a$, zatem $d^2 \leqslant d \cdot k = a$. □

Wniosek 3.5.5. *Jeśli liczba naturalna $n > 1$ nie ma dzielników właściwych co najwyżej równych $\sqrt{n}$, to jest pierwsza.*

Twierdzenie 3.5.6 (Euklidesa)**.** *Zbiór liczb pierwszych jest nieskończony.*

Dowód. Załóżmy, że $p_1, \ldots, p_n$ są wszystkimi liczbami pierwszymi. Wtedy liczba $N = p_1 \cdots p_n + 1$ nie dzieli się przez żadną liczbę pierwszą, $N \equiv 1 \pmod{p_i}$ dla każdego $i = 1, \ldots, n$. Liczba N jest większa od każdej liczby pierwszej, więc nie jest pierwsza. Nie może też być złożona, bo miałaby wtedy dzielnik właściwy pierwszy. Sprzeczność, zatem liczb pierwszych jest nieskończenie wiele. □

Algorytm 3.5.7 (Sito Eratostenesa)**.**

Utwórz tablicę liczb od 2 *do N.*
Powtarzaj:
wypisz nastepną nie skreśloną liczbę i skreśl wszystkie jej wielokrotności,
aż wszystkie liczby będą skreślone.

Przykładowo, dla $N = 30$, mamy (liczby wypisane przed skreśleniem są pogrubione):

	2	**3**	4	**5**	6	**7**	8	9	10
11	12	**13**	14	15	16	**17**	18	**19**	20
21	22	**23**	24	25	26	27	28	**29**	30

Zatem liczby $2, 3, 5, 7, 11, 13, 17, 19, 23, 29$ *są pierwsze.*

Lemat 3.5.8. *Każda liczba całkowita różna od zera ma skończenie wiele dzielników.*

Dowód. Jeśli $a \neq 0$ i $d \mid a$, to $-|a| \leqslant d \leqslant |a|$ z lematu 3.5.1. □

Największym wspólnym dzielnikiem liczb $a, b \in \mathbb{Z}$ nie równych jednocześnie zeru nazywamy liczbę

$$\operatorname{nwd}(a,b) = \max\nolimits_{\leqslant}\{c \in \mathbb{Z} \mid c \mid a \wedge c \mid b\} = \max\nolimits_{\leqslant}\{c \in \mathbb{N}^* \mid c \mid a \wedge c \mid b\}.$$

Twierdzenie 3.5.9. *Jeśli $a, b \in \mathbb{Z}$ nie są jednocześnie równe zeru, to istnieją takie $x_0, y_0 \in \mathbb{Z}$, że $ax_0 + by_0 = \operatorname{nwd}(a,b)$.*

Dowód. Niech $\Omega = \{ax + by : x, y \in \mathbb{Z},\ ax + by > 0\}$. Oznaczmy $d = \min \Omega$ (element najmniejszy). Wtedy $d = ax_0 + by_0$ dla pewnych $x_0, y_0 \in \mathbb{Z}$. Skoro $\operatorname{nwd}(a,b) \mid a$ i $\operatorname{nwd}(a,b) \mid b$, to $\operatorname{nwd}(a,b) \mid ax_0 + by_0 = d$ i $\operatorname{nwd}(a,b) \leqslant d$. Mamy jednak $d \mid a$, bo inaczej byłoby $a = kd + r$, $0 < r < d$, $r = a - kd = a - k(ax_0 + by_0) = a(1 - kx_0) + b(-ky_0) \in \Omega$ i $d \leqslant r$, sprzeczność. Podobnie dowodzi się, że $d|b$. Ostatecznie $d = \operatorname{nwd}(a,b)$. □

Wniosek 3.5.10. *Liczba* $\operatorname{nwd}(a,b)$ *jest elementem największym w sensie relacji $\mid$ w zbiorze dodatnich wspólnych dzielników liczb a, b, tzn. $\forall_{c \in \mathbb{Z}} : c \mid a \wedge c \mid b \Longrightarrow c \mid \operatorname{nwd}(a,b)$.*

Dowód. Istnieją $k, l \in \mathbb{Z}$ takie, że $a = ck$, $b = cl$. Istnieją $x_0, y_0 \in \mathbb{Z}$ takie, że $\operatorname{nwd}(a,b) = ax_0 + by_0 = ckx_0 + cly_0 = c(kx_0 + ly_0)$. □

Twierdzenie 3.5.11. *Jeśli $a = b \cdot q + r$, $0 \leqslant r < b$, to* $\operatorname{nwd}(a,b) = \operatorname{nwd}(b,r)$.

Dowód. $d \mid a = bq + r, d \mid b \Longrightarrow d \mid b, d \mid r = a - bq$. Podobnie $d \mid b, d \mid r \Longrightarrow d \mid b, d \mid a = bq + r$. Te same liczby są wspólnymi dzielnikami par (a,b) i (b,r). Ta sama liczba jest największą z nich. □

Wniosek 3.5.12 (algorytm Euklidesa)**.** *Załóżmy, że $a, b \in \mathbb{Z}$ nie są równocześnie zerem. Dopóki $a \neq 0$ powtarzaj*

$$\begin{cases} r := b \pmod a, \\ b := a, \\ a := r. \end{cases}$$

Zwróć b.

Jeśli $\operatorname{nwd}(a,b) = 1$, to mówimy, że liczby a i b są **względnie pierwsze**.

Twierdzenie 3.5.13 (Zasadnicze twierdzenie arytmetyki)**.** *Jeśli $(a,b) \in \mathbb{Z}^2 \setminus \{(0,0)\}$, $c \in \mathbb{Z}$,* $\operatorname{nwd}(a,b) = 1$*, $a \mid bc$, to $a \mid c$.*

Dowód. Istnieją $k \in \mathbb{Z}$ oraz $x_0, y_0 \in \mathbb{Z}$ takie, że $ak = bc$ oraz $1 = \operatorname{nwd}(a,b) = ax_0 + by_0$, Zatem $c = c \cdot 1 = c(ax_0 + by_0) = cax_0 + aky_0 = a(cx_0 + ky_0)$. □

Twierdzenie 3.5.14. *Jeśli liczba pierwsza p dzieli iloczyn $a_1 \cdot \ldots \cdot a_n$, to p dzieli któryś czynnik a_i.*

Dowód. Indukcja na n. Dla $n = 1$ oczywiste. Dla $n = 2$ niech $p|a \cdot b$. Jeśli p nie dzieli a, to $\operatorname{nwd}(p,a) = 1$ i z tw. 3.5.13 wtedy p dzieli b.

Krok indukcyjny: załóżmy tezę dla $n-1$ czynników. Jeżeli p nie dzieli a_1, to, z przypadku $n = 2$, mamy $p|a_2 \cdot \ldots \cdot a_n$. Teza wynika z założenia kroku indukcyjnego. □

Twierdzenie 3.5.15 (faktorialność pierścienia $\mathbb{Z}$)**.** *Każdą liczbę $n \in \mathbb{N}^*$ można zapisać jednoznacznie (z dokładnością do kolejności czynników) jako iloczyn liczb pierwszych.*

Dowód. **Istnienie rozkładu:** Indukcja na n. Liczba $n = 1$ jest iloczynem pustej rodziny liczb pierwszych, a $n = 2$ jest iloczynem jednej liczby pierwszej.

Krok indukcyjny: Załóżmy prawdziwość tezy dla $n' < n (n \geqslant 3)$. Jesli n jest pierwsza, to teza jest oczywista. Jeśli n jest złożona, to $n = a \cdot b$, gdzie a i b są dzielnikami właściwymi n. Stosując do a i b założenie indukcyjne mamy $n = a \cdot b = (p_1 \cdot ... \cdot p_m) \cdot (q_1 \cdot ... \cdot q_{m'})$, gdzie p_i, q_j są pierwsze.

Jednoznaczność rozkładu: Niech $n = p_1 \cdot ... \cdot p_k = q_1 \cdot ... \cdot q_l, k, l \in \mathbb{N}$. Indukcja na n. Dla $n = 1$, $n = 2$ teza jest oczywista.

Krok indukcyjny: załóżmy, że mamy jednoznaczność dla każdego $n' < n (n \geqslant 3)$. Z twierdzenia 3.5.14 liczba p_1 dzieli którąś liczbę q_j, zatem $p_1 = q_j$. Wystarcza teraz zastosować założenie indukcyjne do $n' = p_2 \cdot ... \cdot p_n = q_1 \cdot ... \cdot q_{j-1} \cdot q_{j+1} \cdot ... \cdot q_l$. Każda z liczb p_i jest którąś z liczb q_j i na odwrót. Rozkłady te mogą się różnić tylko kolejnością czynników. □

Lemat 3.5.16. *Jeśli dla $i = 1, \ldots, n$ mamy* $\operatorname{nwd}(a_i, c) = 1$*, to* $\operatorname{nwd}(a_1 \cdots a_n, c) = 1$.

Dowód. Indukcja względem n. Dla $n = 1$ nie ma czego dowodzić. Niech $n = 2$. Skoro $a_1x_1 + cy_1 = 1$ i $a_2x_2 + cy_2 = 1$ dla pewnych $x_1, x_2, y_1, y_2 \in \mathbb{Z}$, to

$$1 - cy_2 = a_2x_2 = a_2x_2(a_1x_1 + cy_1)$$

$$1 = a_1a_2x_1x_2 + a_2x_2cy_1 + cy_2$$

$$1 = (a_1a_2)x_1x_2 + c(y_2 + a_2x_2y_1),$$

czyli $\operatorname{nwd}(a_1a_2, c) = 1$.

Krok indukcyjny: załóżmy, że lemat jest prawdziwy dla $n - 1 (n \geqslant 3)$. Wtedy $\operatorname{nwd}(a_1 \cdot ... \cdot a_{n-1}, c) = 1$ oraz $\operatorname{nwd}(a_n, c) = 1$. Z kroku $n = 2$ wynika, że $\operatorname{nwd}(a_1 \cdot ... \cdot a_n, c) = 1$. □

Lemat 3.5.17. *Jeśli liczby $a_1, \ldots, a_n$ są parami względnie pierwsze i każda jest dzielnikiem liczby c, to $a_1 \cdots a_n$ jest dzielnikiem c.*

Dowód. W rozkładzie c na liczby pierwsze pojawią się (bez powtórzeń) czynniki pierwsze liczb $a_1, \ldots, a_n$. Zatem ich iloczyn jest częścią iloczynu tworzącego liczbę c. □

Twierdzenie 3.5.18 (chińskie o resztach)**.** *Załóżmy, że liczby $a_1, \ldots, a_n \in \mathbb{N}^*$ są parami względnie pierwsze. Wtedy:*

(1) *dla dowolnych $b_1, \ldots, b_n \in \mathbb{Z}$ układ równań*

$$\begin{cases} x \equiv b_1 \pmod{a_1} \\ \cdots\cdots\cdots\cdots \\ x \equiv b_n \pmod{a_n} \end{cases}$$

ma rozwiązanie,

(2) *jeśli x oraz x' są rozwiązaniami układu równań, to $x \equiv x' \pmod{a_1 \cdots a_n}$.*

Dowód.

(1) Indukcja na n. Dla $n = 1$ wystarczy wziąć $x = x_1 = b_1$. Jeśli $n > 1$ i (1) jest prawdziwe dla $n - 1$, to z lematu 3.5.16 mamy $\operatorname{nwd}(a_1 \cdots a_{n-1}, a_n) = 1$. Z twierdzenia 3.5.9 jest $1 = a_1 \cdots a_{n-1} \cdot x_0 + a_n \cdot y_0$ dla pewnych $x_0, y_0 \in \mathbb{Z}$. Określmy $x = x_n = x_{n-1} + (b_n - x_{n-1}) \cdot a_1 \cdots a_{n-1} \cdot x_0$, gdzie x_{n-1} jest rozwiązaniem układu pierwszych $n - 1$ równań. Wtedy dla $i = 1, \ldots, n - 1$ jest $x_n \equiv x_{n-1} \pmod{a_i} \equiv b_i \pmod{a_i}$. Ponadto dla $i = n$ mamy $x_n \equiv x_{n-1} + (b_n - x_{n-1}) \cdot 1 \pmod{a_n} \equiv b_n \pmod{a_n}$. Zatem $x = x_n$ jest rozwiązaniem układu n równań.

(2) Jeśli x i x' są rozwiązaniami układu równań, to $x - x' \equiv 0 \pmod{a_i}$ dla $i = 1, \ldots, n$. Zatem $\forall_{i=1,\ldots,n} : a_i \mid x - x'$. Z lematu 3.5.17 mamy wtedy $a_1 \cdots a_n \mid x - x'$, czyli $x \equiv x' \pmod{a_1 \cdots a_n}$. □

Przykłady 3.5.19. *Należy rozwiązać układ równań*

$$\begin{cases} x \equiv 3 \pmod 3 \\ x \equiv 0 \pmod 4 \\ x \equiv -1 \pmod 5 \\ x \equiv 2 \pmod 7. \end{cases}$$

Dla tego układu mamy: $x_1 = 3$*;*
$x_2 = 3 + (0-3) \cdot 3 \cdot (-1)$*, bo* $3 \cdot (-1) + 4 \cdot 1 = 1$*, czyli* $x_2 = 12$*;*
$x_3 = 12 + (-1-12) \cdot 3 \cdot 4 \cdot 3$*, bo* $3 \cdot 4 \cdot 3 + 5 \cdot (-7) = 1$*, czyli* $x_3 = 12 - 13 \cdot 36 = -456$*;*
$x_4 = -456 + (2-(-456)) \cdot 3 \cdot 4 \cdot 5 \cdot 2$*, bo* $3 \cdot 4 \cdot 5 \cdot 2 + 7 \cdot (-17) = 1$
$x_4 = -456 + 458 \cdot 120 = 54504$*. Ponadto* $3 \cdot 4 \cdot 5 \cdot 7 = 420$*.*
Rozwiązanie ogólne: $x = 54504 + k \cdot 420, k \in \mathbb{Z}$*.*
Prościej: $x = 324 + k \cdot 420, k \in \mathbb{Z}$*.*

Dla $m \in \mathbb{N}^*$ określamy

$$\varphi(m) = |\{k \in \mathbb{N} : k \leqslant m, \operatorname{nwd}(k, m) = 1\}|.$$

W szczególności $\varphi(1) = 1$ i $\varphi(p) = p-1$ dla $p \in \mathcal{P}$. Funkcję $m \longmapsto \varphi(m)$ nazywamy **funkcją Eulera**. Zauważmy, że $\varphi(m)$ jest liczbą elementów odwracalnych w pierścieniu $(\mathbb{Z}_m, +, \cdot)$, czyli mocą (multiplikatywnej) grupy $U(\mathbb{Z}_m)$ elementów odwracalnych tego pierścienia.

Funkcją arytmetyczną nazywamy dowolny ciąg liczbowy $f\colon \mathbb{N}^* \longrightarrow \mathbb{C}$. Mówimy, że funkcja arytmetyczna f jest **multiplikatywna**, jeśli:
(i) f nie jest ciągiem zerowym,
(ii) jeśli $\operatorname{nwd}(m, n) = 1$, to $f(m) \cdot f(n) = f(m \cdot n)$.

Twierdzenie 3.5.20. *Funkcja Eulera* φ *jest multiplikatywną funkcją arytmetycną.*

Dowód. Dla $p \in \mathcal{P}$ mamy $\varphi(p) = p - 1 \neq 0$. Zauważmy, że dla $m \geqslant 2$ mamy $\varphi(m) = |U(\mathbb{Z}_m)|$. Jeśli $\operatorname{nwd}(m, n) = 1$, to odwzorowanie $h\colon \mathbb{Z}_{m \cdot n} \longrightarrow \mathbb{Z}_m \times \mathbb{Z}_n$ określone wzorem $h([a]_{m \cdot n}) = ([a]_m, [a]_n)$ jest izomorfizmem pierścieni (patrz twierdzenie chińskie o resztach). Zatem $\varphi(m \cdot n) = |U(\mathbb{Z}_{m \cdot n})| = |U(\mathbb{Z}_m) \times U(\mathbb{Z}_n)| = |U(\mathbb{Z}_m)| \cdot |U(\mathbb{Z}_n)| = \varphi(m) \cdot \varphi(n)$. □

Twierdzenie 3.5.21. *Jeśli* $n = p_1^{\alpha_1} \cdots p_n^{\alpha_n}$ *jest rozkładem liczby naturalnej* $n > 1$ *na czynniki pierwsze, to* $\varphi(n) = p_1^{\alpha_1 - 1} \cdots p_n^{\alpha_n - 1}(p_1 - 1) \cdots (p_n - 1)$.

Dowód. Skoro φ jest multiplikatywna, to wystarczy obliczyć $\varphi(p^\alpha)$, $p \in \mathcal{P}$, $\alpha \in \mathbb{N}^*$. Dla $\alpha = 1$ jest $\varphi(p) = p - 1$. Dla $\alpha > 1$ liczby względnie pierwsze z p^α są to dokładnie liczby nie dzielące się przez p. Zatem $\varphi(p^\alpha) = p^\alpha - p^{\alpha-1} = p^{\alpha-1}(p-1)$. □

Twierdzenie 3.5.22 (Eulera)**.** *Dla* $a \in \mathbb{Z}$, $m > 1$ *takich, że* $\operatorname{nwd}(a, m) = 1$ *zachodzi* $a^{\varphi(m)} \equiv 1 \pmod m$.

Dowód. Niech $r_1, \ldots, r_{\varphi(m)}$ będzie rodziną wszystkich reszt modulo m w $\mathbb{Z}_m$, które są względnie pierwsze z m (tworzą one grupę multiplikatywną $U(\mathbb{Z}_m)$). Skoro $\operatorname{nwd}(a, m) = 1$, czyli $[a]_m \in U(\mathbb{Z}_m)$, to elementy $[ar_1]_m, \ldots, [ar_{\varphi(m)}]_m$ są również względnie pierwsze z m

oraz wszystkie różne – stanowią permutację elementów $r_1, \dots, r_{\varphi(m)}$. Iloczyn tych elementów jest postaci

$$[a^{\varphi(m)} r_1 \cdots r_{\varphi(m)}]_m = [r_1 \cdots r_{\varphi(m)}]_m.$$

Skracając w grupie $U(\mathbb{Z}_m)$ kolejno przez $r_1, \dots, r_{\varphi(m)}$ otrzymujemy $[a^{\varphi(m)}]_m = [1]_m$, czyli $a^{\varphi(m)} \equiv 1 \pmod m$. □

Twierdzenie 3.5.23 (małe twierdzenie Fermata)**.** *Jeśli $p \in \mathcal{P}$ i $a \in \mathbb{Z}$, to $a^p \equiv a \pmod p$. Ponadto jeśli $p \nmid a$, to $a^{p-1} \equiv 1 \pmod p$.*

Dowód. Jeśli $p \mid a$, to $a \equiv 0 \pmod p$. Zatem $a^p \equiv 0^p \equiv 0 \equiv a \pmod p$. Jeśli $p \nmid a$, to stosujemy twierdzenie Eulera dla $m = p$. □

Twierdzenie 3.5.24 (wielkie twierdzenie Fermata, wykazane w 1996 r.)**.** *Dla $n \geqslant 3$ nie istnieją $x, y, z \in \mathbb{Z} \setminus \{0\}$ takie, że $x^n + y^n = z^n$.*

Oznaczmy $\pi(n) = |\{p \in \mathcal{P} : p \leqslant n\}|$ dla $n \in \mathbb{N}$.

Twierdzenie 3.5.25 (Czebyszewa)**.** *Istnieją stałe dodatnie $A, B \in \mathbb{R}$ takie, że dla $n \geqslant 2$ jest $B\dfrac{n}{\ln n} \leqslant \pi(n) \leqslant C\dfrac{n}{\ln n}$. Inaczej: $\pi(n) = O(\frac{n}{\ln n})$.*

Przykłady 3.5.26 (kryptosystem RSA)**.** *Twórcami kryptosystemu są Rivest, Shamir i Adleman, którzy stworzyli go w roku 1978.*

Chcemy zaszyfrować tekst dowodlnej długości w alfabecie skończonym Σ. Dzielimy tekst na bloki l-literowe (l powinno być duże !), aby przesłać kolejno zaszyfrowane bloki.

W kryptosystemie RSA wybieramy najpierw dwie duże liczby pierwsze p i q tak, aby $n = p \cdot q > |\Sigma|^l$. Wtedy jest dokładnie $\varphi(n) = (p-1) \cdot (q-1)$ liczb naturalnych mniejszych od n i względnie pierwszych z n.

W pierścieniu $(\mathbb{Z}_n, +, \cdot)$ reszt modulo n mamy grupę multiplikatywną elementów odwracalnych

$$(U(\mathbb{Z}_n), \cdot) = \{[m]_n : \exists\, [s]_n \in \mathbb{Z}_n\, [m]_n \cdot [s]_n = [1]_n\} = \{[m]_n : \operatorname{nwd}(m, n) = 1\}.$$

W grupie abelowej $(U(\mathbb{Z}_n), \cdot)$ jest $\varphi(n)$ elementów.

*Wybieramy liczbę e względnie pierwszą z $\varphi(n)$. Liczby n i e ogłaszamy jako nasz **klucz publiczny**. Natomiast naszym **kluczen prywatnym** jest para (n, d), gdzie $d \in \mathbb{N}$ jest takie, że $d \cdot e \equiv 1 \pmod{\varphi(n)}$.*

Teraz każdy blok l-literowy alfabetu Σ traktujemy jako liczbę naturalną mniejszą od n.

Użycie kryptosystemu:

1) Proste: odbiór zaszyfrowanej wiadomości.

Dowolna osoba znajaca nasz klucz publiczny oraz założenia systemu (zatem alfabet Σ i długość bloku l) moze wysłać do nas kryptogram K zadany wzorem

$$K = J^e \pmod n.$$

Osoba znająca klucz prywatny odczytuje tekst jawny ze wzoru

$$J = K^d \pmod n,$$

bo $K^d \equiv J^{ed} \equiv J^1 \pmod n$. (Jeśli $\operatorname{nwd}(J, n) = 1$, to $J^{\varphi(n)} \equiv J^0 \equiv 1 \pmod n$. Jeśli $\operatorname{nwd}(J, n) = p$, to $J^q \equiv J \pmod n$, a jeśli $\operatorname{nwd}(J, n) = q$, to $J^p \equiv J \pmod n$. Pozostaje przydadek trywialny $J \equiv 0 \pmod n$, również dający pożądaną równość modularną.) Wartość

kryptosystemu polega na tym, że jedyną sensowną drogą do odnalezienia d jest rozłożenie n na czynniki pierwsze, a to wymaga uruchomienia algorytmu o złożoności czasowej typu $O(|\Sigma|^{l/2})$, czyli $O(\sqrt{n})$ lub, w przypadku najlepszych znanych algorytmów rozkładu, algorytmu o złożoności $O(e^{C\sqrt[3]{l}\sqrt[3]{(\ln l)^2}})$.

Samo szyfrowanie i deszyfrowanie wykorzystuje algorytmy o złożoności czasowej $O(l^3)$. Jeśli l jest dostatecznie duże, to zadanie rozłożenia liczby n na czynniki pierwsze jest obliczeniowo niewykonalne (o ile nie mamy szybkiego „komputera kwantowego").

Dygresja: *Szacuje się, że od początku Wszechswiata minęło znacznie mniej niż 10^{20} sekund.*

2) Proste: podpis elektroniczny.

Przesyłając komuś tekst zaszyfrowany naszym kluczem prywatnym, czyli przesyłając (podpisaną) wiadomość

$$P = J^d \pmod n$$

możemy potwierdzić autentyczność tekstu. Wtedy dowolna osoba znająca założenia systemu oraz nasz klucz publiczny może zweryfikować tekst, odczytując

$$J = P^e \pmod n.$$

3) Użycie kombinowane.

Jeśli wymienimy z przyjacielem klucze publiczne (poznamy parę (n', e') przyjaciela), to mozemy przesyłać wiadomości podpisane i zaszyfrowane. Jeśli $n < n'$, to wysyłamy wiadomość-kryptogram

$$K = (J^d \pmod n)^{e'} \pmod{n'}.$$

Natomiast przyjaciel może nam przesłać kryptogram

$$K = (J^e \pmod n)^{d'} \pmod{n'}.$$

(Jeśli $n > n'$, to zamieniamy sie rolami z przyjacielem.) Takie użycie gwarantuje tajność i wiarygodność przekazu.

Uwagi:

1) Liczby p, q, e powinny być jak najbardziej przypadkowe, by nie dało sie ich zgadnąć.

2) W szczególności liczby p i q nie powinny być zbyt blisko siebie. Inaczej wróg mógłby wykorzystać równość

$$n = (\frac{p+q}{2})^2 - (\frac{p-q}{2})^2$$

i odgadnąć p i q.

Przykłady 3.5.27. *Przyjmijmy dla ułatwienia, że $l = 1$ oraz $|\Sigma| = 32$, na przykład:*

$$\Sigma = \{A, B, ..., Y, Z, ., !, ?, , , :, \ \}.$$

Litery alfabetu będziemy utożsamiać z liczbami od 1 do 32. Niech $(n, e) = (55, 7)$. Przechwyconą wiadomością jest ciąg liczb $20, 25, 24$. Złamiemy szyfr! Rozkładamy $n = 5 \cdot 11$, zatem $\varphi(n) = 4 \cdot 10 = 40$. Ponadto $d \cdot 7 \equiv 1 \pmod{40}$, czyli $d = 23$.

$$J = K^{23} \pmod{55}$$

$23 = 16 + 4 + 2 + 1$

$K^{23} = K^{16} \cdot K^4 \cdot K^2 \cdot K^1$
$20^2 = 400 \equiv 15 \pmod{55}$
$20^4 \equiv 15^2 = 225 \equiv 5 \pmod{55}$
$20^8 \equiv 5^2 = 25 \pmod{55}$
$20^{16} \equiv 25^2 = 625 \equiv 20 \pmod{55}$
$20^{23} \equiv 20 \cdot 5 \cdot 15 \cdot 20 \equiv 25 \pmod{55}$. *Pierwsza litera: Y.*
$25^2 \equiv 20 \pmod{55}$
$25^4 \equiv 15 \pmod{55}$
$25^8 \equiv 5 \pmod{55}$
$25^{16} \equiv 25 \pmod{55}$
$25^{23} \equiv 25 \cdot 15 \cdot 20 \cdot 25 \equiv 5 \pmod{55}$. *Druga litera: E.*
$24^2 = 576 \equiv 26 \pmod{55}$
$24^4 \equiv 26^2 = 676 \equiv 16 \pmod{55}$
$24^8 \equiv 16^2 = 256 \equiv 36 \pmod{55}$
$24^{16} \equiv 36^2 = 1296 \equiv 196 \equiv -24 \pmod{55}$
$24^{23} \equiv (-24) \cdot 16 \cdot 26 \cdot 24 \equiv -36 \equiv 19 \pmod{55}$. *Trzecia litera: S.*
Sprawdzenie:
$25^7 \equiv 25 \cdot 20 \cdot 15 = 7500 \equiv 2000 \equiv 20 \pmod{55}$
$5^7 \equiv 5 \cdot 25 \cdot 20 \equiv 2500 \equiv 300 \equiv 25 \pmod{55}$
$19^7 \equiv 19 \cdot 31 \cdot 26 \equiv 24 \pmod{55}$

Rozdział 4

Grafy

4.1. Podstawowe definicje - digrafy

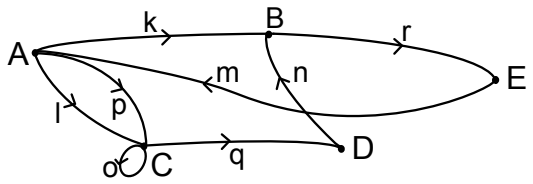
Grafem skierowanym (digrafem) nazywamy trójkę $D = (V(D), A(D), \psi_D)$, gdzie $V(D)$ jest zbiorem niepustym zwanym zbiorem **wierzchołków**, $A(D)$ jest dowolnym zbiorem zwanym zbiorem **łuków** oraz $\psi_D : A(D) \to V(D) \times V(D)$ jest tak zwaną **funkcją incydencji** digrafu. Jeśli $a \in A(D)$ i $\psi_D(a) = (v_1, v_2)$, to mówimy, że v_1 jest **początkiem**, a v_2 jest **końcem** łuku a.

Rysunkiem digrafu D nazywamy rysunek, w którym wierzchołki oznaczone są punktami, a łuki są strzałkami biegnącymi od początku łuku do końca łuku.

Przykłady 4.1.1. *Niech* $V(D) = \{A, B, C, D, E\}$, $A(D) = \{k, l, m, n, o, p, q, r\}$, *oraz* ψ_D *niech będzie dana tabelką:*

a	$\psi_d(a)$
k	*(A, B)*
l	*(A, C)*
m	*(E, A)*
n	*(D, B)*
o	*(C, C)*
p	*(A, C)*
q	*(C, D)*
r	*(B, E)*

Wtedy rysunkiem digrafu jest



Jeśli początek łuku jest równy końcowi łuku, to łuk nazywamy **pętlą**. Może się zdarzyć, że wiele łuków ma ten sam początek i ten sam koniec; wtedy mówimy o **łukach wielokrotnych** digrafu.

Digraf prosty to digraf bez pętli i bez łuków wielokrotnych. Mówimy, że wierzchołek v_1 jest **sąsiedni do** wierzchołka v_2 gdy istnieje łuk o początku v_1 i końcu v_2. Relacja bycia „sąsiednim do" nie musi być symetryczna.

Jeśli D nie ma łuków wielokrotnych, to możemy go utożsamiać z relacją sąsiedztwa na zbiorze $V(D)$, czyli z podzbiorem $A(D)$ iloczynu kartezjańskiego $V(D) \times V(D)$ (wtedy łuk $a \in A(D)$ utożsamiamy z $\psi_D(a) \in V(D) \times V(D)$).

[Wielu autorów tak definiuje digraf, nazywając digrafy z łukami wielokrotnymi **multigrafami skierowanymi**.]

Trasą skierowaną (*ang.:* directed walk) w digrafie nazywamy skończony ciąg łuków taki, że koniec poprzedniego łuku jest początkiem następnego łuku.
Długością trasy skierowanej jest długość tego ciągu.
Ścieżką skierowaną (*ang.:* directed trail) jest trasa skierowana o wszystkich łukach różnych.
Drogą skierowaną (*ang.:* directed path) jest trasa skierowana, której wszystkie początki są różne i wszystkie końce są różne. (Droga skierowana jest zawsze ścieżką skierowaną.)
Trasa skierowana jest **zamknięta** jeśli początek pierwszego łuku (początek całej trasy skierowanej) jest końcem jej ostatniego łuku (końcem całej trasy skierowanej). **Cykl skierowany** to skierowana droga zamknięta dodatniej długości. *(Uwaga: Niektórzy autorzy zakładają o drogach skierownych, że są niezamknięte i definiują cykl skierowany jako skierowaną ścieżkę zamkniętą dodatniej długości, w której wszystkie wierzchołki są różne za wyjątkiem równości początku i końca całego cyklu.)*

Przykłady 4.1.2. *W digrafie z Przykładu 4.1.1 łuk o jest pętlą. Jest tu łuk wielokrotny (para:l, p). Trasa skierowana pqnrml jest ścieżką skierowaną długości 6. Trasa skierowana pqnrm jest cyklem skierowanym długości 5. Wierzchołek C nie jest sąsiedni do B, ale jest sąsiedni do siebie.*

Ilość łuków **wychodzących** z danego $v \in V(D)$ (tzn. o początku v nazywamy **stopniem wyjściowym** v i oznaczamy $outdeg(v)$. Ilość łuków **wchodzących** do wierzchołka $v \in V(D)$ (tzn. o końcu v)nazywamy **stopniem wejściowym** v i oznaczamy $indeg(v)$. Jeśli $indeg(v) = 0$, to v nazywamy **źródłem**. Jeśli $outdeg(v) = 0$, to v nazywamy **ujściem**. Digraf D jest **skończony** jeśli oba zbiory $V(D)$ i $A(D)$ są skończone.

Macierzą sąsiedztwa digrafu skończonego D nazywamy macierz $A_D = [a_{ij}](i, j = 1, \ldots, n)$, gdzie $V(D) = \{v_1, \ldots, v_n\}$ oraz a_{ij} jest liczbą łuków o początku v_i i końcu v_j. Jeśli D nie ma łuków wielokrotnych, to A_D jest macierzą relacji sąsiedztwa wierzchołków digrafu.

Jeśli dany wierzchołek jest początkiem lub końcem określonego łuku, to mówimy, że wierzchołek i łuk są **incydentne**. Wierzchołek pętli jest z nią dwukrotnie incydentny.
Macierzą incydencji skończonego digrafu D jest macierz $M_D = [m_{ik}]$ gdzie

$$m_{ik} = \begin{cases} 0, & \text{gdy } v_i \text{ nie jest incydentny z } a_k, \\ 1, & \text{gdy } v_i \text{ jest jednokrotnie incydentny z } a_k, \\ 2, & \text{gdy } v_i \text{ jest dwukrotnie incydentny z } a_k. \end{cases}$$

$V(D) = \{v_1, \ldots, v_n\}$, $A(D) = \{a_1, \ldots, a_m\}$, $i = 1, \ldots, n$; $k = 1, \ldots, m$.

Macierz incydencji digrafu prostego jest macierzą relacji incydencji w $V(D) \times A(D)$. Suma wyrazów w każdej kolumnie M_D wynosi 2.

Przykłady 4.1.3. *W digrafie z przykładu 4.1.1 mamy $outdeg(C) = 2$, $indeg(C) = 3$. Macierzą sąsiedztwa jest*

$$A_D = \begin{bmatrix} 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \text{ a macierzą incydencji jest } M_D = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 2 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Nie ma źródła ani ujścia.

Digraf D jest **silnie spójny**, gdy dla każdej pary wierzchołków $(v_1, v_2) \in V(D) \times V(D)$ istnieje droga skierowana (wystarcza istnienie trasy skierowanej) o początku v_1 i końcu v_2 (mówimy wtedy, że v_2 jest **osiągalny** z v_1).
Sumą rozłączną digrafów D_1 i D_2 nazywamy digraf $D_1 + D_2$, którego zbiorem wierzchołków jest suma rozłączna $V(D_1) \sqcup V(D_2)$, zbiorem łuków jest suma rozłączna $A(D_1) \sqcup A(D_2)$ oraz funkcją incydencji jest suma rozłączna $\psi_{D_1+D_2} = \psi_{D_1} \sqcup \psi_{D_2}$. Digraf D jest **spójny** jeśli nie jest sumą rozłączną dwóch digrafów.

Każdy digraf można rozłożyć na sumę rozłączną pewnego zbioru digrafów spójnych zwanych **składowymi spójnymi** digrafu.

Poddigrafem digrafu D jest taki digraf D', że $V(D') \subset V(D)$, $A(D') \subset A(D)$ i $\psi_{D'} \subset \psi_D$ (tzn. $\psi_{D'} = \psi_D|_{A_{D'}}$).

Digraf pusty to digraf bez łuków (jest wiele digrafów pustych).

Digrafy D_1 i D_2 są **izomorficzne** jeśli istnieją bijekcje $\varphi_V : V(D_1) \to V(D_2)$, $\varphi_A : A(D_1) \to A(D_2)$ takie, że jeśli $\psi_{D_1}(a) = (v_1, v_2)$, to $\psi_{D_2}(\varphi_A(a)) = (\varphi_V(v_1), \varphi_V(v_2))$.

Przykłady 4.1.4. *W przykładzie 4.1.1 digraf D jest silnie spójny. Każdy poddigraf nie zawierający wierzchołka E nie zawiera też łuków: r, m. Gdyby odjąć łuki r, m, to digraf rozpadłby się na dwie składowe spójne: digraf o wierzchołkach A, B, C, D oraz digraf pusty o wierzchołku E.*

Dowolne dwa digrafy puste o tej samej liczbie wierzchołków są izomorficzne. (Ogólniej: digrafy izomorficzne różnią się tylko nazwami wierzchołków i łuków.)

4.2. Podstawowe definicje - grafy nieskierowane

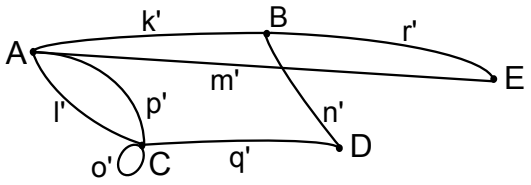
Grafem nieskierowanym (krócej: **grafem**) nazywamy trójkę $G = (V(G), E(G), \psi_G)$ gdzie $V(G)$ jest zbiorem niepustym zwanym zbiorem **wierzchołków** G, $E(G)$ jest dowolnym zbiorem zwanym zbiorem **krawędzi** G, a ψ_G jest **funkcją incydencji** grafu G, która każdej krawędzi $e \in E(G)$ przyporządkowuje parę nieuporządkowaną (niekoniecznie różnych) wierzchołków G zwanych **końcami** krawędzi e.
Rysunkiem grafu G nazywamy rysunek, w którym wierzchołki są oznaczone jako punkty, a krawędzie jako krzywe (lub odcinki) łączące końce krawędzi.

Szkieletem digrafu D nazywamy graf G, który powstaje z zastąpienia łuków krawędziami.

Przykłady 4.2.1. *Szkieletem digrafu z Przykładu 4.1.1 jest graf $G = (V(G), E(G), \psi_G)$ gdzie $V(G) = \{A, B, C, D, E\}$, $E(G) = \{k', l', m', n', o', p', q', r'\}$ oraz funkcja incydencji dana jest tabelką:*

a	$\psi_d(a)$
k'	$\{A, B\}$
l'	$\{A, C\}$
m'	$\{A, E\}$
n'	$\{D, B\}$
o'	$\{C, C\}$
p'	$\{A, C\}$
q'	$\{C, D\}$
r'	$\{B, E\}$

Rysunkiem G jest



Pętlą (*ang.:* loop) nazywamy krawędź, której końce są równe. Pozostałe krawędzie nazywamy **łączami** (*ang.:* link). Jeśli wiele krawędzi ma te same końce, to mówimy o **krawędziach wielokrotnych**. Graf bez pętli i bez krawędzi wielokrotnych nazywamy **grafem prostym**.
[Niektórzy autorzy grafy mogące mieć krawędzie wielokrotne i pętle nazywają **multigrafami**.] Wierzchołek i krawędź są **incydentne** jeśli wierzchołek jest końcem krawędzi. Przy tym pętla i jej wierzchołek są dwukrotnie incydentne. Ilość incydencji danego wierzchołka v z krawędziami grafu nazywamy **stopniem** wierzchołka v i oznaczamy $\deg(v)$. Wierzchołki nazywamy **sąsiednimi** jeśli istnieje krawędź, która je łączy.

Macierzą sąsiedztwa grafu G nazywamy macierz $A_G = [a_{ij}]$, gdzie a_{ij} jest liczbą krawędzi łączących wierzchołek v_i $(i = 1, \ldots, |V(G)|)$ z wierzchołkiem v_j. Macierz sąsiedztwa dowolnego grafu jest symetryczna.

Macierzą incydencji grafu G nazywamy macierz $M_G = [m_{ik}]$ gdzie m_{ik} jest krotnością incydencji pomiędzy wierzchołkiem v_i a krawędzią e_k $(i = 1, \ldots, |V(G)|,\ k = 1, \ldots, |E(G)|)$. Suma wyrazów w każdej kolumnie macierzy M_G wynosi 2. Dla grafu prostego macierz sąsiedztwa jest macierzą relacji sąsiedztwa wierzchołków grafu, a macierz incydencji jest macierzą relacji incydencji wierzchołków i krawędzi G.

Przykłady 4.2.2. *W grafie z przykładu 4.2.1 wierzchołki B i C nie są sąsiednie, a C jest sąsiedni ze sobą. Mamy* $\deg(C) = 5$. *Krawędź o' jest pętlą, a krawędzie l', p' tworzą krawędź wielokrotną. Macierzą sąsiedztwa jest*

$$A_G = \begin{bmatrix} 0 & 1 & 2 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 2 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}, \text{ a macierzą incydencji jest } M_G = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 2 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Lemat 4.2.3 (o uściskach dłoni)**.** *W każdym digrafie suma stopni wyjściowych wszystkich wierzchołków jest równa sumie stopni wejściowych. W każdym grafie suma stopni wszystkich wierzchołków jest parzysta.*

Dowód. Każdy łuk ma 1 początek i 1 koniec, zatem

$$\sum_{v \in V(D)} indeg(v) = |A(D)| = \sum_{v \in V(D)} outdeg(v).$$

Każda krawędź ma dwa końce (licząc z krotnościami), zatem

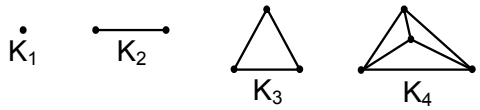
$$\sum_{v \in V(G)} \deg(v) = 2|E(G)|.$$

□

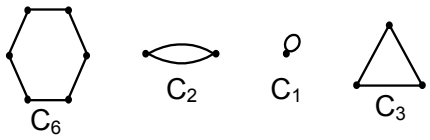
Dla grafów mamy pojęcia: trasy, ścieżki, drogi i cyklu analogiczne do pojęć z teorii grafów skierowanych. (Czasem trasy są zapisywane inaczej, na przykład jako ciągi wierzchołków i krawedzi albo jako ciągi tylko wierzchołków, ale zawsze długość trasy to długość ciągu krawędzi.)

Przykłady 4.2.4. *1. Graf pusty o n wierzchołkach oznaczamy N_n.*

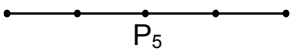
2. Graf prosty o n wierzchołkach i maksymalnej liczbie krawędzi nazywamy ***grafem pełnym*** *lub* ***kliką*** *o n wierzchołkach i oznaczamy K_n. Taki graf ma $\binom{n}{2}$ krawędzi.*



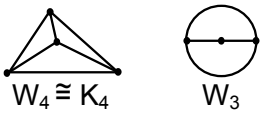
3. Graf spójny o n wierzchołkach, w którym każdy wierzchołek jest stopnia 2 nazywamy ***grafem cyklicznym****. Często będziemy go utożsamiać z* ***n-cyklem*** *(czyli z cyklem długości n) i oznaczać go przez C_n.* ***Trójkąt*** *to 3-cykl.*



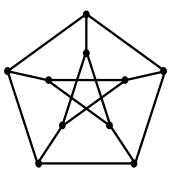
4. Graf powstający z n-cyklu przez usunięcie jednej krawędzi nazywamy ***grafem liniowym*** *o n wierzchołkach i oznaczamy P_n.*



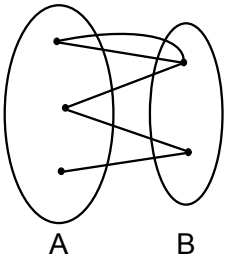
5. ***Kołem*** *o n wierzchołkach nazywamy graf powstający z C_{n-1} przez dodanie nowego wierzchołka („środka koła") i połączenia go krawędzią z każdym dotychczasowym wierzchołkiem. Koło o n wierzchołkach oznaczymy W_n.*



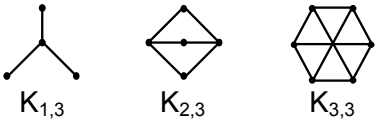
6. ***Grafem r-regularnym*** *nazywamy taki graf, w którym każdy wierzchołek ma stopień r. Na przykład klika K_n jest (n-1)-regularna, grafy puste są 0-regularne, grafy cykliczne są 2-regularne. Grafy 3-regularne nazywamy* ***kubicznymi****. Na przykład poniższy* ***graf Petersena*** *jest kubiczny:*



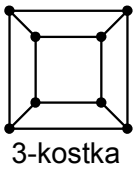
7. ***Grafem dwudzielnym*** *nazywamy taki graf, w którym zbiór wierzchołków jest podzielony na 2 podzbiory i krawędzie mogą łączyć tylko wierzchołki z różnych podzbiorów*



Graf dwudzielny pełny *to graf dwudzielny prosty mający maksymalnie wiele krawędzi. Oznaczamy go $K_{n,m}$ gdy podzbiory zbioru wierzchołków mają odpowiednio moce n, m. Graf $K_{n,m}$ ma $n \cdot m$ krawędzi.*



8. ***K-kostką*** *jest graf prosty, którego wierzchołkami są wszystkie ciągi zerojedynkowe długości k, a krawędzie łączą te wierzchołki, które różnią się na dokładnie jednym miejscu. K-kostka ma 2^k wierzchołków oraz $k \cdot 2^{k-1}$ krawędzi, jest grafem k-regularnym i dwudzielnym (wierzchołki mają parzystą lub nieparzystą liczbę jedynek).*



Mówimy, że grafy G_1 i G_2 są tożsame jako **grafy nieoznakowane** jeśli istnieje izomorfizm między nimi. Graf nieoznakowany można więc utożsamiać z „klasą równoważności" w sensie „relacji" izomorficzności grafów.

Mówimy, że grafy G_1 i G_2 są tożsame jako **grafy oznakowane** jeśli wierzchołki obu grafów mają nazwy z tego samego zbioru **etykiet** (*ang.* labels) oraz istnieje izomorfizm G_1 na G_2 taki, że dowolny wierzchołek G_1 jest przekształcony na wierzchołek o tej samej nazwie (etykiecie) w G_2. *Będziemy używać tylko takich grafów oznakowanych, w których wszystkie wierzchołki mają różne etykiety!* Po ustaleniu bijekcji pomiędzy zbiorem n etykiet, a liczbami ze zbioru $\{1, ..., n\}$, skończony prosty graf oznakowany można więc utożsamiać z macierzą sąsiedztwa grafu.

4.3. Drzewa i lasy

Lasem nazywamy graf acykliczny, czyli graf bez cykli. **Drzewem** jest las spójny. Każdy las jest grafem prostym, bo nie ma 1-cykli i 2-cykli jako podgrafów. **Liściem** w lesie nazywamy wierzchołek stopnia 1.

Lemat 4.3.1. *Jeśli G jest skończony i spójny oraz każdy wierzchołek ma stopień co najmniej 2, to G ma cykl. W szczególności każde niepuste drzewo ma liść.*

Dowód. Wybieramy dowolny wierzchołek i prowadzimy ścieżki wychodzące z tego wierzchołka. Najdłuższa taka ścieżka musi mieć powtarzający się wierzchołek (inaczej kończyłaby

się na liściu). Najkrótszy odcinek tej ścieżki zawierający powtarzający się wierzchołek jest cyklem. Druga część tezy wynika z pierwszej przez kontrapozycję. □

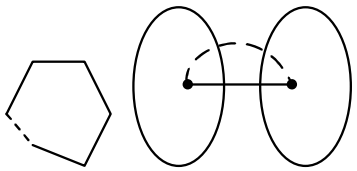
Dowód. Jeśli wszystkie wierzchołki sa stopnia co najmniej 2, to z lematu o uściskach dłoni mamy $|E(G)| \geqslant |V(G)|$, więc G nie może być lasem. Druga część tezy wynika z pierwszej przez kontrapozycję. □

Lemat 4.3.2. *Jeśli graf G ma n wierzchołków i k składowych spójnych, to $|E(G)| \geqslant n-k$.*

Dowód. Jeśli $m = 0$, to graf jest pusty. Wtedy $n = k$ i teza jest spełniona. Załóżmy, że $m > 0$ i lemat jest prawdziwy dla $m-1$. Wyjęcie jednej krawędzi z G zwiększa liczbę składowych co najwyżej o jeden (może co najwyżej rozbić jedną składową na dwie). Z założenia indukcyjnego $m-1 \geqslant n-(k+1)$ jeśli zwiększyła się liczba składowych lub $m-1 \geqslant n-k$ w przeciwnym razie. Zawsze mamy $m \geqslant n-k$. □

Mostem w G jest taka krawędź, której wyjęcie powiększa liczbę składowych grafu.

Uwaga 4.3.3. *1. Żaden cykl nie ma mostów, bo graf liniowy jest spójny.*
2. Most nie jest częścią żadnego cyklu.
3. Krawędź nie będąca częścią żadnego cyklu jest mostem (rozspaja składową jej końców na dwie).



Twierdzenie 4.3.4. *Dla grafu G o n wierzchołkach następujące warunki są równoważne:*
1. G jest drzewem,
2. G jest acykliczny i $|E(G)| = n-1$,
3. G jest spójny i ma $n-1$ krawędzi,
4. G jest spójny i każda jego krawędź jest mostem,
5. każde dwa wierzchołki G są połączone jedyną drogą,
6. G jest acykliczny i dodanie jednej dowolnej krawędzi tworzy jeden cykl.

Dowód. Indukcja na n. Dla $n = 1$ twierdzenie jest oczywiste. Niech $n \geqslant 2$ i załóżmy, że twierdzenie jest prawdziwe dla $n' < n$.

1) $\Rightarrow$ 2) G jest acykliczny, więc każda krawędź jest mostem. Po odjęciu dowolnej krawędzi G rozspaja się, czyli $G-\{e\} = G_1 + G_2$, przy tym G_1 i G_2 są składowymi spójnymi powstałego grafu. Z założenia indukcyjnego dla G_1 i G_2 ($|V(G_1)| < n$, $|V(G_2)| < n$) mamy $|V(G_1)| = |E(G_1)| + 1$ bo G_1 jest spójny,
$|V(G_2)| = |E(G_2)| + 1$ bo G_2 jest spójny,

$$n = |E(G-\{e\})| + 2 = |E(G)| + 1.$$

2) $\Rightarrow$ 3) Każda składowa spójna G_i grafu G jest drzewem, zatem $|E(G_i)| = |V(G_i)| - 1$ dla każdej składowej spójnej G_i grafu G. Zatem $|E(G)| = |V(G)| - k$, gdzie $k =$ ilość składowych spójnych. G ma jedną składową spójną, czyli jest spójny.

3) $\Rightarrow$ 4) G jest spójny i wyjęcie dowolnej krawędzi rozspaja G, bo $\forall e \in E(G)$ $E(G-\{e\}) = n-2$ więc $G-\{e\}$ ma co najmniej 2 składowe spójne.

4) $\Rightarrow$ 5) Każda para wierzchołków da się połączyć drogą. Gdyby takich dróg było więcej, to różniące się kawałki utworzyłyby przynajmniej jeden cykl. Nie każda krawędź byłaby mostem.
5) $\Rightarrow$ 6) Gdyby G miał cykl, to wierzchołki cyklu dałyby się połączyć różnymi drogami. Dodanie dowolnej krawędzi jej wierzchołki dają się połączyć starą drogą lub nową krawędzią, więc powstał jeden cykl.
6) $\Rightarrow$ 1) G jest acykliczny. Gdyby był niespójny, to dodanie krawędzi łączącej wierzchołki różnych składowych (mostu) nie tworzyłoby cyklu. To jest niemożliwe, więc G jest spójny.

□

Wniosek 4.3.5. *Jeśli G jest lasem o k drzewach, to $|E(G)| = |V(G)| - k$.*

Dowód. Stosujemy tw. 4.3.4 pkt 3) do każdej składowej spójnej G. □

Podgraf T grafu spójnego G zawierający wszystkie jego wierzchołki i będący drzewem nazywamy **drzewem spinającym** grafu G (lub **drzewem rozpinającym** graf G).
Dla dowolnego grafu G **las rozpinający** G jest lasem składającym się z drzew rozpinających każdą składową spójną G.

Twierdzenie 4.3.6. *Każdy graf skończony ma las rozpinający.*

Dowód. Usuwamy z grafu nie-mosty o ile to możliwe. Gdy stanie się to niemożliwe, to każda krawędź będzie mostem, a więc graf stanie się lasem rozpinającym wyjściowy graf, bo żaden wierzchołek nie został usunięty. □

Twierdzenie 4.3.7. *(a) (Cayley, 1889) Istnieje n^{n-2} drzew o zbiorze wierzchołków $\{1, 2, \ldots, n\}$ (tzn. drzew oznakowanych o n wierzchołkach etykietowanych kolejnymi liczbami naturalnymi).*
(b) Klika K_n ma n^{n-2} różnych drzew spinających.

Dowód. (a) Dla $n = 1$ i $n = 2$ twierdzenie jest oczywiste. Niech $n \geqslant 3$. Utworzymy bijekcję pomiędzy zbiorem wszystkich drzew oznakowanych o wierzchołkach $1, 2, \ldots, n$, a zbiorem wszystkich ciągów długości $n - 2$ o wyrazach ze zbioru $\{1, 2, \ldots, n\}$.

Dla danego drzewa oznakowanego T tworzymy ciąg $(a_1, \ldots, a_{n-2})$ w następujący sposób: Wybieramy najmniejszy numer wierzchołka b_1 stopnia 1 (liścia).
Jako a_1 bierzemy numer wierzchołka sąsiadującego z b_1, a b_1 wraz z krawędzią z nim incydentną usuwamy z T. Robimy to samo z drzewem $T - \{b_1\}$: wybieramy najmniejszy wierzchołek b_2 stopnia 1 oraz wpisujemy do ciągu jako a_2 ten wierzchołek, który z b_2 sąsiaduje ... Po wybraniu b_{n-2} i a_{n-2} i wyrzuceniu b_{n-2} otrzymujemy drzewo o dwóch wierzchołkach i jednej krawędzi. Takie drzewo jest jedyne.

Odwrotnie: dla dowolnego ciągu $(a_1, \ldots, a_{n-2})$ wyrazów z $\{1, \ldots, n\}$ tworzymy odpowiadające drzewo oznakowane następująco: Liczby nie występujące w tym ciągu oznaczają wierzchołki stopnia 1 (ogólniej: wierzchołek stopnia k występuje w ciągu $k-1$ razy). Niech b_1 będzie najmniejszą liczbą ze zbioru wierzchołków $V = \{1, 2, \ldots, n\}$ nie występującą w ciągu. Łączymy b_1 z a_1 oraz usuwamy b_1 z V i a_1 z ciągu. (O ile było to ostatnie wystąpienie a_1 w ciągu, to od teraz a_1 jest wśród wierzchołków nie występujących w ciągu). Powtarzamy to postępowanie aż wybierzemy b_{n-2} i a_{n-2}. Na koniec dodajemy krawędź pomiędzy dwoma wierzchołkami pozostałymi w V.

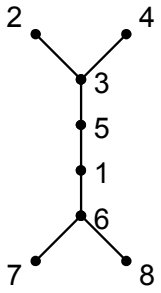
Można zauważyć, że procesy te są wzajemnie odwrotne, zatem wyznaczają wzajemnie odwrotne bijekcje.

(b) Klika zawiera wszystkie możliwe krawędzie grafu prostego, zatem w klice K_n drzewami spinającymi są wszystkie drzewa oznakowane o wierzchołkach $\{1,\dots,n\}$. □

Przykłady 4.3.8. *Dla ciągu* $(3,3,5,1,6,6)$ *mamy*

$b_1 = 2$	*krawędź*	$\{2,3\}$
$b_2 = 4$		$\{4,3\}$
$b_3 = 3$		$\{3,5\}$
$b_4 = 5$		$\{5,1\}$
$b_5 = 1$		$\{1,6\}$
$b_6 = 7$		$\{7,6\}$

pozostała krawędź $\{6,8\}$.



Dla powstałego drzewa odpowiadającym ciągiem jest $(3,3,5,1,6,6)$.

Grafem z wagami nazywamy graf, w którym każdej krawędzi $e \in E(G)$ przypisano pewną liczbę $w(e) \in \mathbb{R}$, zwaną **wagą krawędzi** e. Wtedy dla dowolnego podgrafu H grafu G lub dla dowolnej trasy w G **wagą podgrafu** H nazywamy liczbę $\sum_{e_i \in E(H)} w(e_i)$, a **wagą trasy** jest suma wag wszystkich jej krawędzi (Wagi są liczone z krotnościami dla tras nie będących ścieżkami). *Będziemy zakładać, że wagi krawędzi są liczbami dodatnimi!*

Problem najkrótszych połączeń

Zagadnienie praktyczne: Stworzyć sieć kolejową pomiędzy n miastami ($n \geqslant 2$) tak, by podróż od dowolnego miasta do dowolnego miasta była możliwa, ale by całkowita długość torów (zatem koszt budowy lub eksploatacji) była najmniejsza.

Problem z teorii grafów: Dla danego skończonego spójnego grafu z wagami G znaleźć taki podgraf łączący wszystkie wierzchołki, którego waga jest najmniejsza.

Uwaga: Rozwiązanie jest drzewem spinającym G i możemy założyć, że G jest prosty. Zatem problem można sformułować: Znaleźć **optymalne** (= o najmniejszej wadze) **drzewo spinające** G.

Algorytm 4.3.9 (Kruskala[1], 1956)**.** *Ponumeruj krawędzie* G *wg rosnących wag. Podstaw* $E = \varnothing$.
Dla j *od* 1 *do* $|E(G)|$ *wykonuj:*
jeśli zbiór $E \cup \{e_j\}$ *wyznacza graf acykliczny, to podstaw* $E := E \cup \{e_j\}$.
Zwróć E.

O ile użyjemy właściwych typów danych, to algorytm ten ma złożoność $O(m \log m)$, gdzie $m = |E(G)|$.

Twierdzenie 4.3.10. *Algorytm Kruskala wybiera krawędzie optymalnego drzewa spinającego.*

Dowód. Powstały graf T utworzony z wierzchołków G i krawędzi wybranych przez algorytm jest drzewem spinającym G, bo jest acykliczny i dodanie dowolnej krawędzi utworzyłoby cykl. Jeśli S jest innym drzewem spinającym G, to niech e_k będzie najwcześniejszą krawędzią należącą do $E(T) \setminus E(S)$. Wtedy $S + \{e_k\}$ ma cykl C, który zawiera krawędź $e_l \in E(S) \setminus E(T)$. Oczywiście $e_k \neq e_l$. Załóżmy, że e_l jest najwcześniejszą z takich krawędzi. Gdyby $w(e_l) < w(e_k)$, to oznaczałoby, że e_l została odrzucona w procesie tworzenia T,

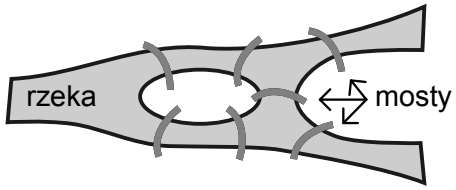
[1] matematyk amerykański Joseph Kruskal (1928 –)

bo tworzyła cykl. Musiałaby zatem tworzyć cykl w drzewie S. Sprzeczność dowodzi, że $w(e_l) \geqslant w(e_k)$. Dla drzewa spinającego $S' = S + \{e_k\} - \{e_l\}$ zachodzi $w(S') \leqslant w(S)$ oraz S' ma o przynajmniej jedną więcej wspólną krawędź z T z „początkowego odcinka wspólnych krawędzi". Powtarzając to rozumowanie otrzymamy ciąg drzew spinających G o nierosnących wagach, który kończy się drzewem T. Zatem $w(T) \leqslant w(S)$. □

Istnieją ponadto: odwrócony algorytm Kruskala, algorytm Prima – patrz [RW].

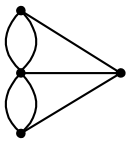
4.4. Grafy eulerowskie i półeulerowskie

Przykłady 4.4.1 (problem mostów królewieckich)**.** *Plan miasta Królewca w 18 wieku wyglądał mniej więcej tak:*



Problem*: Czy można przejść dokładnie raz przez każdy most tak, by powrócić do punktu wyjścia?*

Odpowiedź Eulera (1736): Nie można! Przyczyna: poniższy graf nie jest eulerowski.



Graf spójny skończony G jest **eulerowski**, gdy można w nim znaleźć **cykl Eulera**, czyli zamkniętą ścieżkę przechodzącą przez wszystkie krawędzie G.

Ostrzeżenie: cykl Eulera nie musi być drogą, więc nie musi być też cyklem!

Graf spójny skończony G nazywamy **półeulerowskim**, gdy można w nim znaleźć **ścieżkę Eulera**, czyli ścieżkę przechodzącą przez wszystkie krawędzie, ale nie można znaleźć cyklu Eulera.

Twierdzenie 4.4.2 (Eulera, 1736)**.** *Jeżeli G jest skończony i spójny, to:*
G jest eulerowski $\iff \forall v \in V(G)\ 2|\deg(v)$.

Dowód. ($\Rightarrow$) Cykl Eulera przechodząc przez dowolny wierzchołek powiększa każdorazowo udział krawędzi tej ścieżki w stopniu wierzchołka o 2. Ponieważ cykl Eulera zawiera każdą krawędź grafu dokładnie raz, to dowolny wierzchołek G ma stopień parzysty.

($\Leftarrow$) Indukcja na $m = |E(G)|$. Twierdzenie jest oczywiste dla $m = 0$, czyli grafu pustego. Zakładamy, że $m > 0$ i twierdzenie jest prawdziwe dla $m' < m$. Każdy wierzchołek G ma stopień co najmniej 2, zatem w G jest cykl C z lematu 4.3.1. Graf $G' = G - E(C)$ ma mniej niż m krawędzi oraz każda jego składowa spójna ma wierzchołek wspólny z C. Z założenia indukcyjnego każda ze składowych zawiera cykl Eulera. Jeśli obejdziemy cykl C dookoła tak, że napotykając jakąś składową grafu G' obchodzimy jej cykl Eulera i wracamy w to samo miejsce cyklu C, to utworzymy cykl Eulera całego G. □

Z dowodu twierdzenia wynika:

Wniosek 4.4.3. *Skończony graf spójny jest eulerowski wtedy i tylko wtedy, gdy zbiór jego krawędzi można rozłożyć na cykle. W szczególności graf eulerowski nie ma mostów.*

Twierdzenie 4.4.4. *Jeśli G jest skończony i spójny, to:*
G jest półeulerowski $\Leftrightarrow$ w G są dokładnie 2 wierzchołki stopni nieparzystych.

Dowód. ($\Rightarrow$) Początek i koniec ścieżki Eulera są jedynymi wierzchołkami nieparzystych stopni.
($\Leftarrow$) Jeśli G ma 2 wierzchołki nieparzystych stopni, to dodając krawędź e pomiędzy nimi otrzymujemy graf eulerowski $G + \{e\}$. Skoro $G + \{e\}$ ma cykl Eulera Z, to $Z - \{e\}$ jest ścieżką Eulera w G. □

Algorytm 4.4.5 (Fleury'ego, 1883)**.** *Dane: Graf eulerowski lub półeulerowski G.*
Szukane: cykl Eulera lub ścieżka Eulera w G.
Wybierz wierzchołek stopnia nieparzystego $v \in V(G)$ lub dowolny wierzchołek, jeśli takich nie ma. Podstaw $Z := (v)$, $G' := G$.
Dopóki $\deg(v) > 0$ *wykonuj*
jeśli to możliwe, to weź krawędź $e \in E(G')$ incydentną z v, która nie jest mostem, w przeciwnym razie weź dowolną krawędź incydentną z v w G';
podstaw $G' := G' - \{e\}$, $v :=$ przeciwległy koniec e; $Z := (Z, e, v)$.
Zwróć ciąg Z.

Algorytm Fleury'ego ma złożoność $O(n^2m)$, gdzie $n = |V(G)|, m = |E(G)|$.

Twierdzenie 4.4.6. *W grafie eulerowskim algorytm Fleury'ego zwraca cykl Eulera, a w grafie półeulerowskim zwraca ścieżkę Eulera.*

Dowód. Algorytm Fleury'ego kończy się, gdy dochodzimy do wierzchołka, który stał się izolowany. Jeśli G jest eulerowski, to jest to wierzchołek początkowy, a jeśli G jest półeulerowski, to jest to pozostały (poza początkowym) wierzchołek stopnia nieparzystego, gdyż każdy inny wierzchołek ma nieodpowiedni stopień. Otrzymujemy zatem odpowiednią ścieżkę zamkniętą lub niezamkniętą Z. Gdyby nie zawierała ona wszystkich krawędzi G, to powstały graf $G' = G - E(Z)$ miałby nieizolowane wierzchołki. Każdy wierzchołek powstałego G' miałby parzysty stopień, więc każda składowa G' byłaby grafem eulerowskim. Wychodząc z ostatniego wierzchołka z Z, który nie stał się izolowany w powstałym ostatecznie G', przechodziliśmy przez most, zatem inne krawędzie incydentne z tym wierzchołkiem również były mostami i nimi pozostały. Ale w ostatecznym G' nie ma mostów z Wniosku 4.4.3. Sprzeczność dowodzi, że Z zawiera wszystkie krawędzie. □

Problem chińskiego listonosza

Zagadnienie praktyczne: Listonosz ma przejść przez wszystkie ulice danego miasta i powrócić na pocztę. Znaleźć najkrótszą trasę spełniającą ten warunek.
Problem z teorii grafów: W danym skończonym spójnym grafie z wagami G znaleźć trasę zamkniętą o najmniejszej wadze, która zawiera wszystkie krawędzie.

Rozwiązanie częściowe

1. Jeśli G jest eulerowski, to każdy cykl Eulera jest rozwiązaniem problemu i można je znaleźć algorytmem Fleury'ego.
2. Jeśli G jest półeulerowski, to wystarcza znać **najkrótszą drogę** pomiędzy wierzchołkami nieparzystych stopni.

Ogólny algorytm stworzony przez Edmondsa i Johnsona (1973) jest skomplikowany.

Problem najkrótszej drogi

Zagadnienie praktyczne: Znaleźć najkrótsze połączenie pomiędzy dwoma miastami.
Problem z teorii grafów: W grafie skończonym spójnym z wagami G znaleźć trasę[2] pomiędzy zadanymi wierzchołkami o minimalnej wadze.
Uwaga: Można założyć, że G jest prosty!

Długość (= wagę) najkrótszej (najlżejszej) drogi pomiędzy dwoma wierzchołkami u, v będziemy nazywali **odległością** tych wierzchołków i oznaczali $d(u,v)$. **Odległością** wierzchołka v **od zbioru wierzchołków** S jest $\min_{u\in S} d(v,u)$.

Uwaga: Niech $S \subsetneq V(G)$ i $u \in S$. Jeśli najkrótszą drogą od zadanego wierzchołka u do zbioru $V(G) \setminus S$ jest
$D = u \dots u'\, u'v$, to $u' \in S$ oraz
$\quad d(u,v) = d(u,u') + w(u'v)$, zatem
$(*)\quad d(u, V(G)\setminus S) = \min_{u'\in S, v\notin S} d(u,u') + w(u'v))$.

Algorytm 4.4.7 (Dijkstra[3] 1959, Whiting & Hillier 1960)**.** *Idea: Tworzymy ciąg wstępujący podzbiorów $S_0 \subset S_1 \subset \dots \subset V(G)$. Zaczynając od wierzchołka początkowego $S_0 = \{u\}$. Niech u_1 będzie wierzchołkiem najbliższym (jednym z najbliższych) u, tzn. $d(u,u_1) = d(u, V(G)\setminus S_0)$. Oznaczmy $P_1 = uu_1$ oraz $S_1 = \{u, u_1\}$. Jeśli dla $i = 1, \dots, k$ zdefiniowano już zbiory S_i oraz najkrótsze drogi P_i od u do u_i, to obliczamy $d(u, V(G)\setminus S_k)$ z $(*)$ i wyznaczamy wierzchołek $u_{k+1} \in V(G)\setminus S_k$ taki, że*

$$d(u,u_{k+1}) = d(u, V(G)\setminus S_k).$$

Zachodzi $d(u,u_{k+1}) = d(u,u_j) + w(u_ju_{k+1})$ dla pewnego $j \leqslant k$. Tworzymy $S_{k+1} = S_k \cup \{u_{k+1}\}$ oraz $P_{k+1} = P_j + u_ju_{k+1}$. W ten sposób znajdujemy najkrótszą drogę od u do dowolnego wierzchołka G.

Wprowadzając etykiety odległości od $u = u_0$ dla wierzchołków G możemy krócej zapisać ten algorytm:

Podstaw $l(u_0) := 0$; $l(v) = +\infty$ dla $v \neq u_0$, $S_0 := \{u_0\}$,
Dla i od 0 do $|V(G)| - 1$ wykonuj:
dla każdego $v \in V(G)\setminus S_i$ zamień $l(v)$ na $\min(l(v), l(u_i) + w(u_iv))$,
oblicz $\min_{v\in V(G)\setminus S_i}(l(v))$ i wybierz u_{i+1}, w którym minimum jest przyjęte
oraz podstaw $S_{i+1} = S_i \cup \{u_{i+1}\}$.

Po wykonaniu tego algorytmu znamy odległość $l(v)$ pomiędzy u_0 a dowolnym wierzchołkiem v z G. Ma on złożoność $O(n^2)$, gdzie $n = |V(G)|$. Najkrótszą drogę można odtworzyć badając listę odległości i tablicę wag albo w czasie wykonywania algorytmu pozostawiając wskaźniki do „wcześniejszych wierzchołków".

4.5. Grafy hamiltonowskie

Grafy hamiltonowskie swą nazwę zawdzięczają matematykowi irlandzkiemu Williamowi Hamiltonowi (1805 –1865).

[2] rozwiązanie będzie drogą
[3] holenderski informatyk Edsger Dijkstra (1930 – 2002)

Graf G nazywamy **hamiltonowskim** gdy można w nim znaleźć **cykl Hamiltona**, czyli cykl przechodzący przez każdy wierzchołek G.

Graf G nazywamy **półhamiltonowskim** gdy nie jest hamiltonowski, ale ma **drogę Hamiltona**, czyli drogę (niezamkniętą) przechodzącą przez wszystkie wierzchołki G.

Uwaga 4.5.1. *Najprostszym warunkiem koniecznym na hamiltonowskośc (a nawet półhamiltonowskość) grafu jest spójność grafu. W szczególności musi być* $|E(G)| \geqslant |V(G)| - 1$.

Twierdzenie 4.5.2 (warunek konieczny)**.** *Jeśli G jest hamiltonowski, to dla każdego niepustego i właściwego podzbioru $S \subset V(G)$ jest*

$$\omega(G - S) \leqslant |S|,$$

gdzie $\omega(G - S)$ oznacza liczbę składowych spójnych grafu $G - S$ (tj. grafu powstałego z G przez wyjęcie wierzchołków ze zbioru S i incydentnych z nimi krawędzi).

Dowód. Jeśli C jest cyklem Hamiltona w G, to

$$\omega(C - S) \leqslant |S|,$$

ale $C - S$ zawiera wszystkie wierzchołki grafu $G - S$, zatem

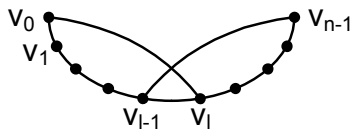
$$\omega(G - S) \leqslant \omega(C - S).$$

□

Warunki wystarczające na hamiltonowskość:

Twierdzenie 4.5.3 (Ore[4], 1960)**.** *Jeśli graf prosty G ma $n \geqslant 3$ wierzchołków oraz dla dowolnych różnych i niesąsiednich wierzchołków $v, w \in V(G)$ jest* $\deg(v) + \deg(w) \geqslant n$, *to G jest hamiltonowski.*

Dowód. Klika K_n jest grafem hamiltonowskim. Wystarcza wykazać, że dla każdego podgrafu H w K_n będącego nadgrafem G jeśli wszystkie grafy $H + \{e\}$, gdzie $e \in E(K_n) - E(H)$, są hamiltonowskie, to H również. Taki graf H ma drogę Hamiltona $v_0 v_1 \dots v_{n-1}$. Jeśli v_0 i v_{n-1} są sąsiednie, to H jest hamiltonowski. Jeśli nie są sąsiednie, to $\deg(v_0) + \deg(v_{n-1}) \geqslant n$, zatem zbiory $S = \{i|\ v_0 v_i \in E(H)\}$ i $T = \{i|\ v_{i-1} v_{n-1} \in E(H)\}$ spełniają warunki $|S| + |T| \geqslant n$ i $|S \cup T| \leqslant n - 1$, zatem $S \cap T \neq \varnothing$.
Istnieje takie l, że
v_0 jest sąsiedni z v_l
v_{l-1} jest sąsiedni z v_{n-1}.
Zatem $v_0 v_1 \dots v_{l-1} v_{n-1} v_{n-2} \dots v_l v_0$ jest cyklem Hamiltona w H.



□

Wniosek 4.5.4 (tw. Diraca, 1952)**.** *Jeśli G jest grafem prostym o $n \geqslant 3$ wierzchołkach takich, że dla każdego $v \in V(G)$ jest* $\deg(v) \geqslant \frac{n}{2}$, *to G jest hamiltonowski.*

[4] matematyk norweski Øystein Ore (1899 – 1968)

Problem komiwojażera

Zagadnienie praktyczne: Komiwojażer ma odwiedzić wszystkie miasta danej mapy i powrócić do punktu wyjścia. Znaleźć najkrótszą trasę.
Problem z teorii grafów: Znaleźć najkrótszą trasę przechodzącą przez wszystkie wierzchołki skończonego, spójnego grafu G.
Uwaga: o ile G jest hamiltonowski, to rozwiązanie będzie cyklem Hamiltona, czyli należy znaleźć **optymalny cykl Hamiltona**. Nie ma dobrego (czyli wielomianowego względem ilości wierzchołków i ilości krawędzi grafu) algorytmu rozwiązującego ten problem.

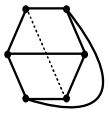
4.6. Grafy planarne

Graf G jest **planarny** jeśli istnieje rysunek G na płaszczyźnie, na którym krawędzie G nie przecinają się.

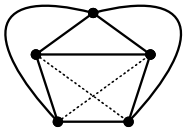
Przykłady 4.6.1. *Grafy K_1, K_2, K_3, K_4 oraz $K_{2,n}$ dla $n \in \mathbb{N}*$ są planarne.*

Twierdzenie 4.6.2. *Grafy K_5 i $K_{3,3}$ są nieplanarne.*

Dowód. Załóżmy, że $K_{3,3}$ jest planarny. Ma on 6-cykl, którego rysunek jest sześciokątem. Każda z krawędzi łączących przeciwległe wierzchołki może leżeć wewnątrz lub na zewnątrz sześciokąta. Dwie z nich muszą leżeć wewnątrz lub dwie na zewnątrz, zatem muszą się przecinać.

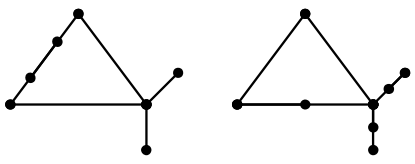


Gdyby K_5 był planarny, to miałby cykl długości 5. Rysunek grafu zawierałby pięciokąt. Z pozostałych 5-ciu krawędzi przynajmniej 3 leżą wewnątrz pięciokąta lub 3 leżą na zewnątrz pięciokąta. Zatem któreś dwie (mające oba końce różne) muszą się przecinać.



□

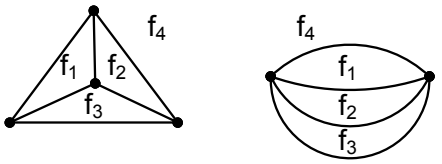
Grafy G_1 i G_2 nazywamy **homeomorficznymi** jeśli oba powstają przez dodanie wierzchołków stopnia 2 wewnątrz krawędzi pewnego grafu H. Na przykład grafy na rysunku są homeomorficzne.



Twierdzenie 4.6.3 (Kuratowski, 1930)**.** *Graf G jest planarny wtedy i tylko wtedy, gdy nie ma podgrafu homeomorficznego z K_5 lub $K_{3,3}$.*

Uwaga 4.6.4. *Podgraf grafu planarnego jest planarny. Zatem planarność oznacza posiadanie nie za dużej ilości krawędzi, odwrotnie do hamiltonowskości.*

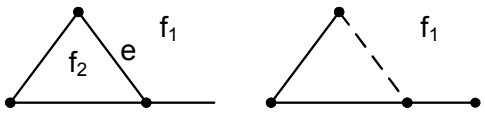
Ścianą rysunku płaskiego grafu planarnego G jest składowa spójna dopełnienia tego rysunku do płaszczyzny, czyli każdy obszar ograniczony krawędziami grafu G. Ściany grafu mogą być nieograniczone. Jeśli G jest skończony, to ma on jedyną ścianę nieograniczoną.



Twierdzenie 4.6.5 (Euler, 1750)**.** *Dla dowolnego rysunku płaskiego spójnego grafu planarnego o n wierzchołkach, m krawędziach i f ścianach zachodzi równość*

$$n - m + f = 2$$

Dowód. Indukcja względem $m = |E(G)|$. Jeśli $m = 0$, to $n = 1$ i $f = 1$, twierdzenie jest prawdziwe. Zakładamy, że twierdzenie jest prawdziwe dla $m - 1$. Jeśli G jest drzewem, to $n = m+1$ i $f = 1$, więc $n-m+f = 2$. Jeśli G nie jest drzewem, to wyjmując z G nie-most e otrzymujemy graf $G' = G-\{e\}$ spójny i planarny. Z założenia indukcyjnego dla G' mamy $n-(m+1)+f' = 2$, ale wyjęcie e z G rozerwało jeden cykl oddzielający wnętrze cyklu od przyległej ściany. Zatem $f' = f - 1$ oraz $n - m + f = 2$. □



Wniosek 4.6.6. *Dla dowolnego rysunku płaskiego grafu planarnego G o n wierzchołkach, m krawędziach i f ścianach i k składowych spójnych zachodzi równość*

$$n - m + f = k + 1$$

Dowód. Stosujemy twierdzenie 4.6.5 do każdej składowej G. Liczby wierzchołków, krawędzi i ścian sumują się za wyjątkiem ściany nieograniczonej, która występuje w każdej (z k) składowej spójnej. Stąd $n - m + f = 2k - (k - 1) = k + 1$. □

Wniosek 4.6.7. *1. Jeśli G jest spójnym, planarnym grafem prostym o $n \geqslant 3$ wierzchołkach i m krawędziach, to $m \leqslant 3n - 6$.*

2. Jeśli dodatkowo w G nie ma trójkątów, to $m \leqslant 2n - 4$.

Dowód. 1. Jeśli G jest drzewem o 3 wierzchołkach, to teza zachodzi (bo $2 \leqslant 3\cdot 3-6$). W przeciwnym przypadku każda ściana jest ograniczona przynajmniej trzema krawędziami (bo G jest prosty). Każda krawędź rozdziela co najwyżej dwie ściany. Zatem $f \leqslant \frac{2}{3}m$. Z twierdzenia Eulera mamy $2 = n - m + f \leqslant n - m + \frac{2}{3}m = n - \frac{1}{3}m$, czyli $n \geqslant 2 + \frac{1}{3}m$.

2. Jeżeli G jest drzewem o $\leqslant 4$ wierzchołkach, to nierówność jest spełniona ($3 \leqslant 2\cdot 4-4$ lub $2 \leqslant 2\cdot 3-4$). W przeciwnym przypadku jeśli w G nie ma trójkątów, to każda ściana jest ograniczona przynajmniej czterema krawędziami, zatem $f \leqslant \frac{2}{4}m$. Z twierdzenia Eulera jest $2 = n - m + f \leqslant n - m + \frac{1}{2}m = n - \frac{1}{2}m$, czyli $n \geqslant 2 + \frac{1}{2}m$.

□

Twierdzenie 4.6.8. *Każdy planarny skończony graf prosty G ma wierzchołek stopnia $\leqslant 5$.*

Dowód. Możemy założyć, że G jest spójny i ma co najmniej 3 wierzchołki. Gdyby dla każdego $v \in V(G)$ było $\deg(v) \geqslant 6$, to $6n \leqslant 2m$ (z lematu o uściskach dłoni) oraz $3n \leqslant m \leqslant 3n - 6$. Sprzeczność. □

4.7. Kolorowanie grafów

Graf bez pętli jest **k-kolorowalny** (wierzchołkowo) jeśli każdemu wierzchołkowi możemy przypisać jeden z k kolorów tak, by sąsiednie wierzchołki miały różne kolory.
Graf k-kolorowalny, który nie jest $(k-1)$-kolorowalny nazywamy **k-chromatycznym**.

Przykłady 4.7.1. *Graf trywialny K_1 jest 1-chromatyczny. Drzewo niepuste jest 2-chromatyczne, podobnie cykl o parzystej długości. Natomiast cykl nieparzystej długości $\geqslant 3$ jest 3-chromatyczny. Klika K_n jest n-chromatyczna. Grafy niepuste dwudzielne są 2-chromatyczne.*

Twierdzenie 4.7.2. *Graf prosty skończony G o wszystkich wierzchołkach stopnia $\leqslant \Delta$ jest $(\Delta+1)$-kolorowalny.*

Dowód. Indukcja na $n = |V(G)|$. Dla $n = 1$ G jest pusty, więc jest 1-kolorowalny. Załóżmy, że twierdzenie jest prawdziwe dla $n-1$. Dla dowolnego $v \in V(G)$ graf $G' = G - \{v\}$ ma $n-1$ wierzchołków oraz wszystkie wierzchołki G' są stopni $\leqslant \Delta$. Z założenia indukcyjnego G' jest $(\Delta+1)$-kolorowalny. Ale $\deg(v) \leqslant \Delta$. Wystarcza więc pomalować v na kolor różny od kolorów sąsiadów v w kolorowaniu G'. □

Twierdzenie 4.7.3 (Brooks, 1941)**.** *Jeśli G jest skończony spójny i prosty oraz nie jest kliką, to dla $\Delta \geqslant 3$ jeśli każdy z wierzchołków G jest stopnia $\leqslant \Delta$, to G jest Δ-kolorowalny.*

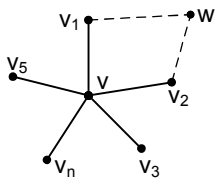
Od teraz zakładamy, że rozważane grafy są planarne.

Twierdzenie 4.7.4. *Każdy planarny skończony graf prosty jest 6-kolorowalny.*

Dowód. Indukcja na $n = |V(G)|$. Jeśli $n \leqslant 6$ to twierdzenie jest oczywiste. Załóżmy, że twierdzenie jest prawdziwe dla $n-1$. Z twierdzenia 4.6.8 graf G ma wierzchołek v stopnia $\leqslant 5$. Skoro $G-\{v\}$ jest 6-kolorowalny, to malując v ma kolor różny od kolorów jego sąsiadów otrzymujemy 6-kolorowalność grafu G. □

Twierdzenie 4.7.5. *Każdy planarny skończony graf prosty jest 5-kolorowalny.*

Dowód. Indukcja na $n = |V(G)|$. Twierdzenie jest oczywiste dla $n \leqslant 5$. Załóżmy, że $n > 5$ i twierdzenie jest prawdziwe dla $n' < n$. Z twierdzenia 4.6.8 graf G ma wierzchołek v stopnia $\leqslant 5$. Graf $G' = G - \{v\}$ jest 5-kolorowalny z założenia indukcyjnego. Jeśli $\deg(v) < 5$, to możemy nadać v kolor różny od kolorów jego sąsiadów. Pozostaje przypadek $\deg(v) = 5$. Oznaczmy sąsiadów v jako v_1, v_2, v_3, v_4, v_5. Ponieważ G nie zawiera kliki K_5, któreś dwa wierzchołki spośród v_1, v_2, v_3, v_4, v_5 nie są sąsiednie. Możemy założyć, że v_1 i v_2 nie są sąsiednie. Utożsamiając wierzchołki v, v_1, v_2 w G oraz utożsamiając krawędzie łączące v_1, v_2 i ewentualnie v ze wspólnymi sąsiadami otrzymijemy graf G'' o $n-2$ wierzchołkach, który jest 5-kolorowalny.



Graf G jest 5-kolorowalny następująco: bierzemy 5-kolorowanie G'', wierzchołki v_1, v_2 kolorujemy tak, jak wierzchołek powstały z utożsamienia v, v_1, v_2, natomiast v kolorujemy tym kolorem, który nie występuje wsród kolorów wierzchołków v_1, v_2, v_3, v_4, v_5. □

Twierdzenie 4.7.6 (Appel & Haken z pomocą komputera, 1976)**.** *Każdy skończony planarny graf prosty jest 4-kolorowalny.*

Graf G jest **k-spójny** jeśli do rozspojenia G trzeba z niego wyjąć przynajmniej k wierzchołków. Przyjmuje się przy tym, że K_1 jest 0-spójny i ogólniej, że graf pełny K_n jest $(n-1)$-spójny. Jeżeli G jest spójny i nie jest pełny, to jego **spójnością** $\kappa(G)$ nazywamy najmniejszą liczbę wierzchołków, których wyjęcie rozspaja graf. (Jeśli G jest 1-spójny i 2-spójny, ale nie 3-spójny, to $\kappa(G) = 2$).

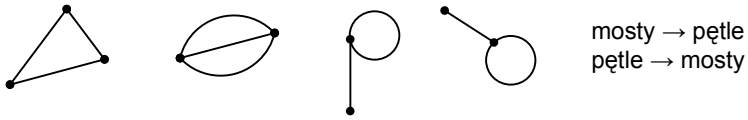
Przykłady 4.7.7. *Każdy nietrywialny graf spójny jest 1-spójny. Cykl długości co najmniej 3 jest 2-spójny. Drzewo mające nie-liść jest 1-spójne. Drzewo mające tylko dwa liście jest również 1-spójne.*

Mapa to skończony graf planarny 3-spójny. W szczególności mapa nie ma mostów i każdy wierzchołek mapy ma stopień co najmniej 3.

Dla danego rysunku płaskiego skończonego grafu planarnego G **grafem (geometrycznie) dualnym** $G^\star$ nazywamy graf, którego rysunek otrzymujemy w następujacy sposób:
1. ścianom G odpowiadają wierzchołki $G^\star$,
2. wierzchołkom G odpowiadają ściany $G^\star$,
3. wierzchołki $G^\star$ są połączone krawędzią wtedy i tylko wtedy, gdy odpowiadające ściany w G są rozdzielone krawędzią i na odwrót.

Graf $G^{\star\star}$ dualny do $G^\star$ jest **geometrycznie izomorficzny** z grafem G, czyli rysunki G i $G^{\star\star}$ nie różnią się istotnie.

Przykłady 4.7.8. *Graf drugi jest dualny do pierwszego, a graf czwarty jest dualny do trzeciego:*



Lemat 4.7.9. *Jeśli G jest skończonym grafem planarnym, to:*

$$G \text{ jest } k\text{-kolorowalny (wierzchołkowo)} \Leftrightarrow G^\star \text{ jest } k\text{-kolorowalny ścianowo.}$$

Dowód. Wierzchołki G są sąsiednie wtedy i tylko wtedy, gdy odpowiadające im ściany $G^\star$ są sąsiednie. □

Twierdzenie 4.7.10. *Każdą mapę można pokolorować czterema barwami.*

Dowód. Jeśli G jest mapą, to $G^\star$ jest planarny i prosty. Z twierdzenia 4.7.6 $G^\star$ jest 4-kolorowalny wierzchołkowo. Zatem $G \cong G^{\star\star}$ jest 4-kolorowalny ścianowo. □

4.8. Przepływy w sieciach

Sieć to digraf skończony z wagami oraz wyróżnionymi zbiorami źródeł i ujść, to znaczy obiekt postaci $N = (D, c, X, Y)$, gdzie $D = (V(D), A(D), \psi_D)$ jest digrafem, $c : A(D) \mapsto [0, +\infty)$ jest funkcją przypisującą wagi łukom zwaną **funkcją przepustowości** (waga $c(a)$ łuku $a \in A(D)$ jest nazywana jego **przepustowością** (ang.: *capacity*), zbiór $X \subset V(D)$ jest zwany **zbiorem źródeł**, zbiór $Y \subset V(D)$ jest zwany **zbiorem ujść** sieci N. Zakładamy przy tym, że $X \cap Y = \varnothing$, a zbiór $V(D) \setminus (X \cup Y)$ nazywamy **zbiorem wierzchołków pośrednich**.

Stopniem wyjściowym wierzchołka sieci $v \in V(D)$ nazywamy liczbę

$$\operatorname{outdeg}(v) = \sum_{v \text{ jest początkiem } a} c(a).$$

Stopniem wejściowym wierzchołka sieci $v \in V(D)$ nazywamy liczbę

$$\operatorname{indeg}(v) = \sum_{v \text{ jest końcem } a} c(a).$$

Lemat 4.8.1 (o uściskach dłoni)**.** *Suma stopni wejściowych wszystkich wierzchołków sieci jest równa sumie stopni wyjściowych wszystkich wierzchołków sieci.*

Dowód. Zachodzą równości

$$\sum_{v \in V(D)} \operatorname{indeg}(v) = \sum_{a \in A(D)} c(a) = \sum_{v \in V(D)} \operatorname{outdeg}(v).$$

□

Uwaga 4.8.2. *Można założyć, że w dowolnej sieci N jest tylko jedno źródło i tylko jedno ujście, a przy tym są to źródło i ujście w sensie teorii digrafów. Łatwo bowiem przekształcić dowolna sieć w sieć o tych własnościach: dodajemy jeden wierzchołekjako nowe źródło i prowadzimy łuki z niego do wszystkich starych źródeł o przepustowości nieskończonej albo wystarczająco dużej i podobnie dodajemy jeden wierzchołek jako nowe ujście i prowadzimy łuki ze starych ujść do nowego ujścia o przepustowości nieskończonej albo skończonej, ale wystarczająco dużej. Odtąd będziemy zatem przyjmować powyższe założenie.*

Przepływem w sieci N nazywamy funkcję $\phi : A(D) \mapsto [0, +\infty)$ taką, że

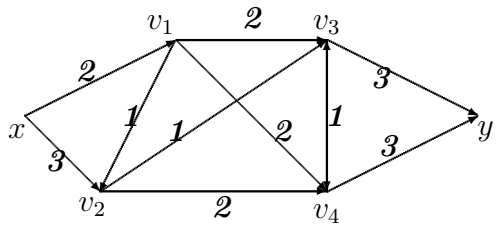
$$\forall a \in A(D)\ \phi(a) \leqslant c(a)$$

oraz

$$\forall v \in V(D) \setminus (X \cup Y) \sum_{v \text{ jest końcem } a} \phi(a) = \sum_{v \text{ jest początkiem } a} \phi(a).$$

Przykłady 4.8.3. *1) Funkcja zerowa, zwana **przepływem zerowym**, jest przepływem w każdej sieci.*

2) W sieci jak na rysunku (przy łukach zaznaczono ich przepustowości)



możliwym przepływem jest funkcja:

$\psi_D(a)$	(x, v_1)	(x, v_2)	(v_1, v_2)	(v_1, v_3)	(v_2, v_4)	(v_2, v_3)	(v_1, v_4)	(v_4, v_3)	(v_3, y)	(v_4, y)
$\phi(a)$	2	2	0	2	2	0	0	1	3	1

Łuk $a \in A(D)$ nazywamy **nasyconym** dla przepływu ϕ jeśli $\phi(a) = c(a)$ oraz **nienasyconym** jeśli $\phi(a) < c(a)$.
Wartością przepływu ϕ nazywamy liczbę

$$\mathrm{val}(\phi) = \sum_{a \text{ ma początek w } x} \phi(a) = \sum_{a \text{ ma koniec w } y} \phi(a), \text{ gdzie } X = \{x\}, Y = \{y\}.$$

Przepływ maksymalny w sieci N to przepływ w N o maksymalnej wartości.

Przykłady 4.8.4. *W sieci i dla przepływu z przykładu 4.8.3 punkt 2) łuki $(x, v_1), (v_1, v_3), (v_2, v_4), (v_3, y), (v_4, v_3)$ są nasycone, a łuki $(x, v_2), (v_1, v_2), (v_1, v_4), (v_2, v_3), (v_4, y)$ są nienasycone. Wartością przepływu jest 4. Nie jest to przepływ maksymalny, z łatwością można znaleźć przepływ o wartości 5.*

Przekrojem P sieci N nazywamy taki zbiór $P \subset A(D)$, że każda trasa od źródła do ujścia w N musi zawierać łuk z P.
Przepustowością przekroju P jest liczba $c(P) = \sum_{a \in P} c(a)$.
Przekrój minimalny sieci N to przekrój o najmniejszej przepustowości.

Przykłady 4.8.5. *W sieci z przykładu 4.8.3 punkt 2) zbiór $P_1 = \{(x, v_1), (x, v_2)\}$ jest przekrojem o przepustowości 5. Zbiór $P_2 = \{(v_3, y), (v_4, y)\}$ jest przekrojem o przepustowości 6. Ponieważ istnieje przepływ o przepustowości 5 w tej sieci, to każdy przekrój tej sieci ma przepustowość co najmniej 5. Zatem przekrój P_1 jest minimalny.*

Uwaga 4.8.6. *Łatwo można zauważyć, że wartość dowolnego przepływu (również maksymalnego) jest zawsze nie większa od przepustowości dowolnego przekroju (również minimalnego).*

Twierdzenie 4.8.7 (o maksymalnym przepływie i minimalnym przekroju, Ford i Fulkerson, 1956)**.** *W dowolnej sieci wartość maksymalnego przepływu jest równa przepustowosci przekroju minimalnego.*

Drogą nieskierowaną (dwukierunkową) w digrafie D bedziemy nazywać taki ciąg wierzchołków i łuków, którego odpowiednikiem w szkielecie digrafu jest droga. Droga nieskierowana ma początek i koniec, ale może mieć łuki zarówno skierowane do przodu (**łuki postępowe**) jak i łuki sierowane do tyłu (**łuki wsteczne**). Dla dowolnej drogi nieskierowanej d w sieci N z zadanym przepływem ϕ określamy „możliwą inkrementację przepływu ϕ wzdłuz d" jako liczbę $i(d) = \min_{a \in A(D)} i(a)$, gdzie

$$i(a) = \begin{cases} c(a) - \phi(a), & \text{jeśli } a \text{ jest łukiem postępowym w } d, \\ \phi(a), & \text{jeśli } a \text{ jest łukiem wstecznym w } d. \end{cases}$$

Droga nieskierowana d jest: **nasycona** jeśli $i(d) = 0$ oraz **nienasycona** jeśli $i(d) > 0$. **Droga powiększajaca przepływ** to droga nieskierowana nienasycona od źródła x do ujścia y w sieci N. Jeżeli dla danego przepływu ϕ w sieci N da sie znaleźć drogę d powiększającą przepływ, to przepływ ϕ nie jest maksymalny, bo można wskazać przepływ $\hat{\phi}$ zadany wzorem:

$$\hat{\phi} = \begin{cases} \phi(a) + i(d), & \text{jeśli } a \text{ jest łukiem postępowym w } d, \\ \phi(a) - i(a), & \text{jeśli } a \text{ jest łukiem wstecznym w } d, \\ \phi(a), & \text{jeśli } a \text{ nie jest łukiem drogi } d. \end{cases}$$

i ten nowy przepływ $\hat{\phi}$ (**modyfikacja ϕ na bazie d**) ma wartość $\mathrm{val}(\hat{\phi}) = \mathrm{val}(\phi) + i(d)$, większą od $\mathrm{val}(\phi)$.

Lemat 4.8.8. *W sieci N przepływ ϕ jest maksymalny wtedy i tylko wtedy, gdy w sieci N nie ma drogi powiększajacej przepływ ϕ.*

Dowód. Jeśli w N istnieje droga powiększająca przepływ ϕ, to ϕ nie jest maksymalny.

Załóżmy, że nie ma drogi powiększającej przepływ ϕ. Niech S będzie zbiorem tych wierzchołków, do których da się dotrzeć ze źródła x drogą nienasyconą. Oczywiście $x \in S$ i $y \notin S$. Zbiór K tych łuków, które wychodzą z wierzchołka w S i wchodzą do wierzchołka spoza S, jest przekrojem sieci N.

Niech a bedzie łukiem z D wychodzącym z wierzchołka $u \in S$ i wchodzącym do wierzchołka $v \notin S$. Istnieje droga nienasycona od x do u, ale nie istnieje droga nienasycona od x do v. Zatem łuk a jest nasycony.

Niech teraz a bedzie łukiem z D wychodzącym z $u \notin S$ i wchodzącym do $v \in S$. Istnieje droga nienasycona od x do v, ale nie istnieje droga nienasycona od x do u. Zatem łuk a musi byc zerowy (to znaczy $\phi(a) = 0$).

Otrzymujemy równość

$$\text{val}(\phi) = \quad \text{całkowity wypływ z wierzchołków zbioru } S \quad = c(K).$$

Oznacza to, że ϕ jest przepływem maksymalnym oraz K jest przekrojem minimalnym. □

Dowód. (Twierdzenia 4.8.7.) Analizę sytuacji możemy zacząć od dowolnego przepływu, na przykład od przepływu zerowego. O ile znajdziemy drogę powiększającą przepływ, to dokonujemy modyfikacji przepływu na bazie tej drogi nieskierowanej. Zauważmy, że po takiej modyfikacji ta droga nieskierowana już nigdy nie będzie powiększać przepływu. W digrafie skończonym jest tylko skończenie wiele dróg nieskierowanych. Po skonczenie wielu krokach nie będzie już drogi powiększającej przepływ. Wobec lematu 4.8.8, istnieje maksymalny przepływ o wartosci równej przepustowości pewnego (minimalnego) przekroju. □

Jak znaleźć minimalny przekrój i maksymalny przepływ?

Zaczynamy od dowolnego przepływu (na przykład od przepływu zerowego). Szukamy drogi powiększającej przepływ. Jeśli taką znajdziemy, to modyfikujemy przepływ na bazie tej drogi i wracamy do punktu wyjścia algorytmu.

Jeśli nie da się znaleźć drogi powiększającej przepływ, to nasz przepływ jest maksymalny i jesteśmy w stanie wskazać przekrój minimalny rozumując tak, jak w dowodzie lematu 4.8.8.

Jak szukamy drogi powiększającej przepływ?

„Uprawiamy" („hodujemy") drzewo nienasycone T tych wierzchołków, do których istnieje droga nienasycona ze źródła x.

Drzewo nienasycone w sieci N dla przepływu ϕ to drzewo będące podgrafem szkieletu digrafu D, które zawiera źródło x oraz dla dowolnego wierzchołka v poza źródłem zawierające (jedyną w T) drogę nieskierowaną od x do v nienasyconą dla przepływu ϕ w sieci N.

Na początku drzewo nienasycone T składa się tylko ze źródła x, któremu nadajemy etykietę $l(x) = \infty$.

Powtarzamy następujące kroki: Każdy zaetykietowany już wierzchołek u **skanujemy**, czyli przeszukujemy wierzchołki sąsiednie do u w szkielecie digrafu D oraz:

1. jeśli w D jest łuk nienasycony a o początku u i nieetykietowanym końcu, to dołączamy łuk a i jego koniec v do drzewa T oraz etykietujemy koniec łuku liczbą $l(v) = \min(l(u), c(a) - \phi(a))$;
2. jesli w D jest łuk pozytywny a (to znaczy $\phi(a) > 0$) o końcu u i nieetykietowanym początku v, to dołączamy łuk a i wierzchołek v do drzewa T oraz etykietujemy początek łuku liczbą $l(v) = \min(l(u), \phi(a))$.

Albo dołączymy wierzchołek y do drzewa T (znajdując drogę powiekszającą przepływ), albo dojdziemy do sytuacji, w której wszystkie zaetykietowane wierzchołki D są już zeskanowane i nie można zaetykietować żadnego nowego wierzchołka (wtedy wierzchołki drzewa T tworzą zbiór S z dowodu lematu 4.8.8).

Powyższy algorytm nazywamy **metodą etykietowania**.

Bibliografia

[B] Cz. Bagiński, *„Funkcje tworzące”*, `http://aragorn.pb.bialystok.pl/~baginski/Pdf/F_tw.pdf`.

[C] J. Cichoń, *„Wykłady ze wstępu do matematyki”*, Dolnośląskie Wydawnictwo Edukacyjne (2003).

[G] J. Gancarzewicz, *„Arytmetyka”*, Wydawnictwo UJ (2000).

[GKP] R. Graham, D. Knuth, O. Patashnik, *„Matematyka konkretna”*, Wyd. Naukowe PWN (wiele wydań).

[LM] W. Lipski, W. Marek, *„Analiza kombinatoryczna”* (BM 59), PWN (1986).

[MO] W. Marek, J. Onyszkiewicz, *„Elementy logiki i teorii mnogości w zadaniach”*, Wyd. Naukowe PWN (wiele wydań).

[N] W. Narkiewicz, *„Teoria liczb”*, Wyd. Naukowe PWN (2003).

[R] H. Rasiowa, *„Wstęp do matematyki współczesnej”* (BM 30), PWN (wiele wydań).

[RW] K. Ross, C. Wright, *„Matematyka dyskretna”*, Wyd. Naukowe PWN (wiele wydań).

[W] R. Wilson, *„Wprowadzenie do teorii grafów”*, Wyd. Naukowe PWN (wiele wydań).

Wiele informacji o matematyce dyskretnej można znaleźć w internecie!

Spis treści