



SPIS TREŚCI NUMERU 5 (516)

Z życia managera magazynu,
czyli jak składać optymalne
zamówienia?

Agnieszka Abramczyk

str. 1

Złodziej strategii

Bartłomiej Żak

str. 3

Co zobaczyła Alicja po drugiej
stronie lustra?

Marek Kordos

str. 4

Fizyka a obliczenia równoległe

Paweł Artymowicz

str. 8

Kolorowanie płaszczyzny,
prostych i okręgów

Jadwiga Czyżewska

str. 12



Zadania

str. 14

Wiszące referencje.

Czy można wyeliminować
to zagrożenie?

Andrzej Salwicki

str. 15

Migawka informatyczna:

Po co komu dowód?

(albo słów kilka
o sztucznej inteligencji)

str. 18

Informatyczny kącik
olimpijski (105):

Odtwarzanie drzewa

Kamil Dębowski

str. 19



Bizony, tury, żubry
Magdalena Fikus

str. 20

Aktualności

str. 21

Klub 44

str. 22

Prosto z nieba: Brakujące
ogniwo

str. 24

Niebo w maju

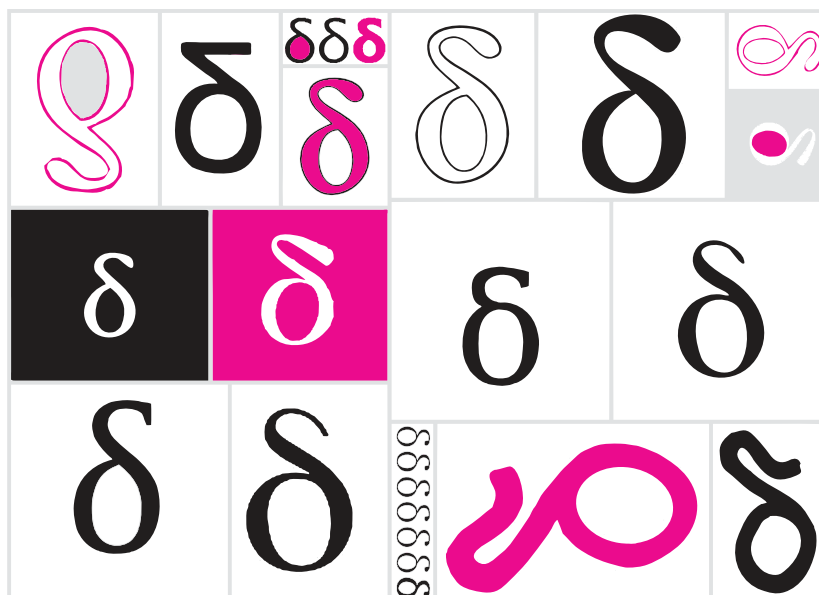
str. 24



Niby nic
Joanna Jaszewska

str. 25

W następnym numerze:



Miesięcznik *Delta* – *matematyka, fizyka, astronomia, informatyka* jest wydawany przez Uniwersytet Warszawski przy współpracy towarzystw naukowych: Polskiego Towarzystwa Matematycznego, Polskiego Towarzystwa Fizycznego, Polskiego Towarzystwa Astronomicznego i Polskiego Towarzystwa Informatycznego.

Komitet Redakcyjny: dr Waldemar Berej, dr Piotr Chrzastowski-Wachtel, dr Krzysztof Ciesielski – wiceprzewodniczący, prof. dr hab. Bożena Czerny, dr Andrzej Dąbrowski, prof. dr hab. Marek Demiański, prof. dr hab. Krzysztof Diks, dr Tomasz Greczyło, prof. dr hab. Paweł Idziak, dr hab. Agnieszka Janiuk, dr hab. Marcin Kiraga, prof. dr hab. Andrzej Majhofer, prof. dr hab. Zbigniew Marciniak, dr hab. Zygmunt Mazur, dr Adam Michalec, prof. dr hab. Michał Nawrocki – przewodniczący, dr Zdzisław Pogoda, dr Paweł Preś, prof. dr hab. Wojciech Rytter, prof. dr hab. Paweł Strzelecki.

Redaguje kolegium w składzie: Wiktor Bartol, Michał Bejger, Szymon Charzyński – z-ca red. nac., Wojciech Czerwiński, Tomasz Kazana, Piotr Kaźmierczak, Krystyna Kordos – sekr. red., Marek Kordos – red. nac., Kamila Łyczek, Katarzyna Małek, Łukasz Rajkowski, Anna Rudnik, Krzysztof Rudnik, Piotr Zalewski.

Adres do korespondencji:

Institut Matematyki UW, Redakcja *Delt*y, ul. Banacha 2, pokój 4020,
02-097 Warszawa, e-mail: delta@mimuw.edu.pl tel. 22-55-44-402.

Okładki i ilustracje: Anna Ludwicka Graphic Design & Serigrafia;
rysunki techniczne: Stanisław Walczak.

Skład systemem \TeX wykonała Redakcja.

str. 19 Wydrukowano w Drukarni Greg, ul. Górczewska 216 p. 101, 01-460 Warszawa.

PRENUMERATA

Garmond Press: www.garmondpress.pl

str. 21

Kolporter: www.kolporter.com.pl

str. 22

RUCH S.A.: www.ruch.com.pl, infolinia 804-200-600

Prenumerata realizowana przez RUCH S.A.:

str. 24

Cena prenumeraty w 2017 roku wynosi 4 zł za egzemplarz.

Zamówienia na prenumeratę w wersji papierowej można składać bezpośrednio na stronie www.prenumerata.ruch.com.pl

str. 24

Ewentualne pytania prosimy kierować na adres e-mail: prenumerata@ruch.com.pl

str. 25

lub kontaktując się z Centrum Obsługi Klienta RUCH

pod numerem: 801 800 803 lub 22 693 70 00 – czynne w dni robocze w godzinach 7⁰⁰–17⁰⁰.

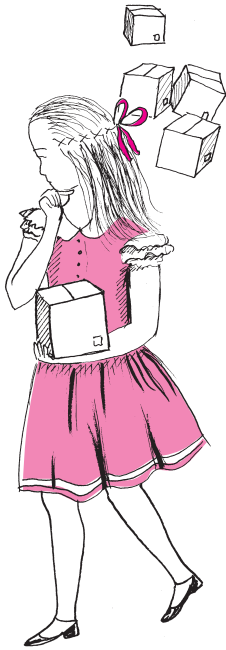
Koszt połączenia wg taryfy operatora.

Numery archiwalne (od 1987 r.) można nabyć w Redakcji osobiście lub listownie.

Strona internetowa (w tym artykuły archiwalne, linki itd.): deltami.edu.pl

Można nas też znaleźć na [facebook.com/Delta.czasopismo](https://www.facebook.com/Delta.czasopismo)

Wydawca: Uniwersytet Warszawski



Z życia managera magazynu, czyli jak składać optymalne zamówienia?

Agnieszka ABRAMCZYK*

Spróbuj, drogi Czytelniku, wyobrazić sobie, że zarządzasz wielkim magazynem z milionami różnych towarów na półkach. W Twoim magazynie znajdują się części zamiennie do koparek, samochodów i motocykli, śrubki, silniki, narzędzia, odzież specjalistyczna, smary i oleje napędowe, oraz inne motoryzacyjne cuda. Jeżeli wczułeś się już w swoją rolę, pozwól, że zadam Ci pytanie. W jaki sposób zarządzasz towarami? Skąd wiesz, kiedy złożyć zamówienie? Czy masz ustalony grafik i zamawiasz na początku każdego miesiąca, mimo tego, że niektóre części piętrzą się aż po sufit? A może zamawiasz dopiero wtedy, gdy zabraknie Ci towaru na półce? A co, jeżeli dostawa trwa dwa miesiące, a przed drzwiami magazynu rozentuzjasmowany tłum harleyowców domaga się nowego modelu skórzanych kurtek? Być może, zamawiasz, gdy ilość towaru jest odpowiednio mała? Ale co to znaczy odpowiednio mała? Czy 5 śrubek to wystarczająco mało? A 5 silników do supernowoczesnego, rzadkiego modelu koparki?

Poszukiwanie odpowiedzi na podobne pytania to codzienność w Synchronie – moim miejscu pracy. Nasz główny produkt jest aplikacją, dzięki której sieć magazynów może zaoszczędzić pieniądze poprzez m.in. minimalizację ilości przechowywanych towarów, optymalizację procesu składania zamówień i efektywne zarządzanie łańcuchem dostaw. W niniejszym artykule opowiem o małym wycinku jednej z funkcjonalności, to znaczy o strategii zamawiania (R, Q) . Zasada działania tej strategii jest bardzo prosta – składasz zamówienie o wielkości Q , kiedy zauważysz, że ilość towaru spadnie do poziomu R .

Ile zamawiać?

Teoretyczne opracowania na temat strategii (R, Q) nie określają jasno, ile zamawiać. Ważne jest natomiast, aby wielkość zamówienia nie zmieniała się istotnie w czasie. Przy wyborze odpowiedniego Q możemy kierować się minimalizacją istotnych kosztów, np. kosztów zamawiania (transport, rozładunek) i przechowywania (koszty związane z zamrożeniem kapitału, składowaniem, potencjalnym zniszczeniem towaru). W praktyce precyzyjne oszacowanie poszczególnych kosztów jest skomplikowane, dlatego też w ramach alternatywy możemy zastosować nieco prostszą formułę, $Q = D/N$, gdzie D to oczekiwany roczny popyt, a N to planowana liczba zamówień w roku.

Kiedy zamawiać?

Przyjmijmy na początek, że popyt i czas dostawy są deterministyczne. Przy znanym popycie możemy dokładnie przewidzieć, kiedy sprzedamy wszystko to, co mamy obecnie w magazynie. Ponieważ czas dostawy jest ustalony, możemy złożyć zamówienie odpowiednio wcześniej – na tyle wcześniej, aby dostawa wypadła w dniu, gdy klient kupuje ostatnią sztukę towaru. Formalizując nieco te rozmyślenia,

R = wielkość popytu w czasie oczekiwania na dostawę.

W rzeczywistości popyt nie jest deterministyczny, dlatego też porzucamy to założenie. Dla ustalenia uwagi przyjmujemy, że mamy do czynienia ze zmienną losową o rozkładzie normalnym $X \sim N(\mu, \sigma^2)$, gdzie μ to oczekiwany popyt w czasie oczekiwania na dostawę, a σ to niepewność naszej prognozy. Podobnie jak poprzednio, moglibyśmy złożyć zamówienie w momencie, gdy ilość towaru w magazynie jest równa μ . Zauważmy jednak, że takie postępowanie daje nam tylko 50% szans, że wszyscy klienci zostaną obsłużeni (rozkład normalny jest symetryczny względem średniej). Ta niepewność prognozy sprawia, że powinniśmy mieć dodatkową pulę towaru na „wszelki wypadek”, czyli, innymi słowy, złożyć zamówienie nieco wcześniej. Czy dużo wcześniej? To zależy od poziomu obsługi, który chcemy zapewnić naszym klientom. Pozwólcie, że na moment zboczę nieco z tematu i wtrącę kilka słów o miarach poziomu obsługi, których najczęściej używa się w praktyce.



Rozwiązanie zadania M 1528.

Odpowiedź: Na 2^{15} sposobów.

Wyróżnijmy 15 pól tablicy, które tworzą skrajny dolny wiersz i skrajną lewą kolumnę. Zauważmy, że dowolne kolorowanie tych pól dwoma kolorami jednoznacznie określa pokolorowanie całej tablicy zgodnie z warunkami zadania — kolory pozostałych pól określamy w kolejności zadanej przez numery na rysunku.

7	8	9	10	11	12	13
6	7	8	9	10	11	12
5	6	7	8	9	10	11
4	5	6	7	8	9	10
3	4	5	6	7	8	9
2	3	4	5	6	7	8
1	2	3	4	5	6	7

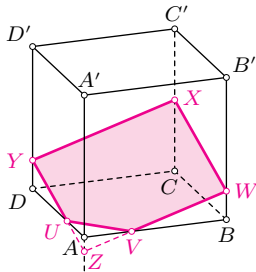
Z drugiej strony każdy sposób pomalowania wszystkich pól tablicy jednoznacznie zadaje kolory pól wyróżnionych. To oznacza, że kolorowania tablicy o zadanych własnościach są we wzajemnie jednoznacznej odpowiedniości z dowolnymi pokolorowaniami wyróżnionych pól, a takich pokolorowań jest 2^{15} .

Uwaga. Podany przykład zbioru 15 pól jednoznacznie kodującego kolorowanie całej tablicy nie jest jedyny — można zadać pytanie, jak wyglądają zestawy 15 pól o takiej własności i ile ich jest. Czytelnika Lubiącego Utożsamiać Tablice z Grafami Dwudzielnymi (a zbioru pól tablicy z odpowiednimi podgrafami) zachęcamy do udowodnienia, że omawiane zbioru pól to drzewa rozpinające takich grafów i jest ich dokładnie 2^{14} .



Rozwiązanie zadania M 1529.

Oznaczmy pewną parę przeciwległych ścian danego sześcianu przez $ABCD$ oraz $A'B'C'D'$ w taki sposób, aby odcinki AA' , BB' , CC' , DD' były krawędziami sześcianu oraz ściana $A'B'C'D'$ nie zawierała żadnego z boków uzyskanego w przekroju pięciokąta, a krawędź AA' — żadnego z jego wierzchołków. Niech ponadto U , V , W , X , Y , Z będą punktami przecięcia płaszczyzny przekroju odpowiednio z prostymi AD , AB , BB' , CC' , DD' , AA' .



Zauważmy, że czworokąt $WXYZ$ jest równoległobokiem, co wynika z równoległości przeciwległych ścian sześcianu. Okrąg wpisany w pięciokąt $UVWXY$ jest wpisany w ten równoległobok, skąd wniosek, że $WXYZ$ jest rombem. Wykażemy, że $AA'C'C$ jest płaszczyzną symetrii tego rombu; ponieważ jest to także płaszczyzna symetrii wyjściowego sześcianu, więc wyniknie stąd, że punkty U i V są względem niej symetryczne, co zakończy rozwiązanie.

Skoro $WXYZ$ jest rombem, to ma prostopadłe przekątne, a zatem punkty W oraz Y są zawarte w płaszczyźnie symetrycznej odcinka XZ . Płaszczyzna ta nie pokrywa się z płaszczyzną $BB'D'D$ (punkt Z nie należy do odcinka AA'), więc przecina się z nią wzdłuż prostej prostopadłej do płaszczyzny $AA'C'C$. Wobec tego punkty W i Y są symetryczne względem płaszczyzny $AA'C'C$, a to właśnie należało udowodnić.

- CSL (ang. *Cycle Service Level*) – prawdopodobieństwo zaspokojenia całego popytu w okresie pomiędzy dostawami wyrażone w procentach,
- IFR (ang. *Item Fill Rate*) – procent zaspokojonego popytu,
- OFR (ang. *Order Fill Rate*) – procent zamówień zrealizowanych w całości.

Aby doświadczyć różnicy pomiędzy powyższymi miarami, rozważmy scenariusz, w którym zamówienie o wielkości 99 dostarczane jest regularnie pierwszego dnia miesiąca. Popyt jest stały i wynosi 100 sztuk miesięcznie. Oczywiście jest, że $IFR = 99\%$ i $CSL = 0\%$ (bo w każdym okresie zabraknie jednej sztuki). OFR zależy od charakterystyki popytu. Jeżeli jeden klient zamówił 100 sztuk, to $OFR = 0\%$. Jeżeli dwóch klientów złożyło po jednym zamówieniu na 50 sztuk, to $OFR = 50\%$, ponieważ tylko jedno z dwóch zamówień zostało w całości zrealizowane.

Wróćmy do ustalenia satysfakcjonującej wartości R . Załóżmy, że chcemy zapewnić poziom obsługi CSL równy $n\%$. Szukamy zatem takiego R , że popyt w czasie oczekiwania na dostawę nie przekroczy tej wartości z prawdopodobieństwem $n/100$, czyli

$$P(X \leq R) = n/100.$$

Znając rozkład zmiennej losowej X , łatwo znajdujemy odpowiednią wartość R . Trochę więcej gimnastyki wymaga wyznaczenie R , które zapewni z góry zadany poziom IFR . Zgodnie z definicją,

$$IFR = \left(1 - \frac{\text{oczekiwany niezaspokojony popyt}}{\text{popyt}}\right) \cdot 100\%.$$

Jednak w praktyce stosuje się dosyć dobre przybliżenie, z którym łatwiej pracować

$$IFR = \left(1 - \frac{\text{oczekiwany niezaspokojony popyt}}{\text{oczekiwany popyt}}\right) \cdot 100\%,$$

my w dalszej części będziemy stosować ten właśnie wzór.

Oczekiwany popyt w okresie pomiędzy zamówieniami jest równy Q , a wartość oczekiwana niezaspokojonego popytu zależy od R i jest modelowana przez funkcję straty

$$G_{\mu,\sigma}(R) = \int_R^{\infty} (x - R) f_{\mu,\sigma}(x) dx,$$

gdzie $f_{\mu,\sigma}$ oznacza gęstość rozkładu normalnego o średniej μ i odchyleniu standardowym σ . Przy odrobinie cierpliwości można wykazać, że funkcja $G_{\mu,\sigma}$ jest ściśle malejąca, istnieje zatem funkcja do niej odwrotna, dzięki czemu R może być jednoznacznie wyznaczone. Odpowiednia wartość R uzyskiwana jest numerycznie. Dużym ułatwieniem przy obliczeniach jest standaryzacja rozkładu, która prowadzi do zależności

$$G_{\mu,\sigma}(R) = \sigma G_{0,1}\left(\frac{R - \mu}{\sigma}\right)$$

oraz formuła $G_{0,1}(R) = f_{0,1}(R) - R(1 - \Phi(R))$, którą można otrzymać z wyjściowego wzoru ($\Phi(R)$ to dystrybuenta rozkładu $N(0,1)$).

Przykład

Założmy, że popyt w czasie oczekiwania na dostawę modelowany jest rozkładem normalnym o średniej $\mu = 120$ i odchyleniu standardowym $\sigma = 40$. Czas oczekiwania na dostawę wynosi 1 tydzień, a zamówienie składane jest 24 razy w roku. Naszym zadaniem jest wyznaczenie strategii (R, Q) , która zapewni $CSL = 80\%$. Jaki jest oczekiwany poziom IFR dla takiej strategii?

Obliczmy najpierw wielkość zamówienia:

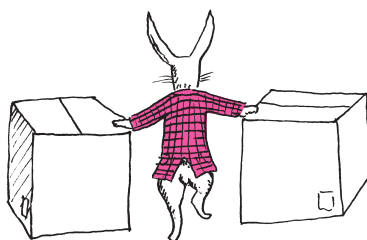
$$Q = D/N = (120/7 \cdot 365)/24 \approx 261.$$

Zgodnie z definicją CSL szukamy kwantyla rzędu 0,8 dla zmiennej losowej X o rozkładzie $N(120, 40^2)$, tzn. takiej wartości R , że $P(X \leq R) = 0,8$. Korzystając z tablic lub dowolnego pakietu statystycznego, otrzymujemy

$$R = 153,6648 \approx 154.$$

Szukana strategia to $(154, 261)$. Teraz sprawdźmy, jaki jest oczekiwany IFR dla wyliczonej strategii

$$G_{120,40}(154) = 40G_{0,1}(0,85) = 40(f_{0,1}(0,85) - 0,85(1 - \Phi(0,85))) \approx 4,4.$$



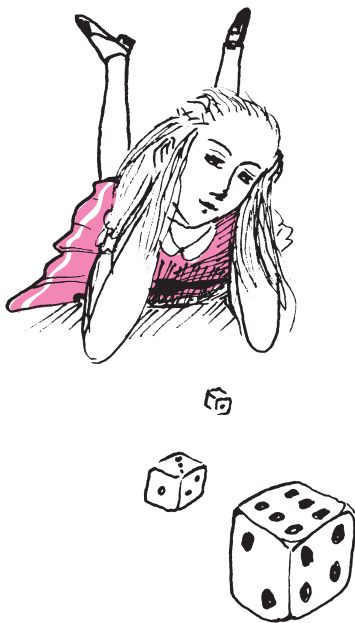
A zatem,

$$IFR = 1 - 4,4/120 \approx 93,33\%.$$

W niniejszym artykule przyjęliśmy kilka założeń ułatwiających zrozumienie badanego procesu i ułatwiających same obliczenia, np. popyt nie zawsze może być modelowany rozkładem normalnym. W przypadku towarów, które sprzedają się rzadko i w nieregularny sposób, konieczne jest użycie rozkładów dyskretnych, np. rozkładu Poissona lub ujemnego dwumianowego. To wiąże się nie tylko z zamianą rozkładu w obliczeniach, ale także ze zmianą sposobu, w jaki podchodzimy do problemu, np. przy rozkładach dyskretnych nie możemy oczekiwać, że istnieje moment, w którym ilość towaru w magazynie wynosi dokładnie R . Najczęstszą konsekwencją takiego założenia jest zbyt późne zamawianie względem modelu i nieosiągnięcie zadanych poziomów obsługi klienta.

W zespole Data Science naszym głównym zadaniem jest modyfikowanie bazowych modeli poprzez ulepszanie i rozwijanie pożądaných funkcjonalności. Dla modelu optymalnego zamawiania są to m.in. uwzględnienie zmienności czasu dostawy, modelowanie opóźnień w centralnych magazynach, rozpatrywanie wielu różnych dostawców, agregacja oraz redystrybucja zamówień i wiele innych. Kluczem jest wymyślenie (lub znalezienie w literaturze) takiego modelu, który balansuje teoretyczne wyrafinowanie z praktycznymi możliwościami implementacyjnymi i dostępnymi danymi. Czasami warto wybrać mniej dokładny model, który pozwala na prostsze i szybsze obliczenia. Warto uświadomić sobie tutaj rozmiar danych, z którymi mamy do czynienia – są to dziesiątki milionów zależnych od siebie części z kilkuletnią historią.

Złodziej strategii



Jedną z rzeczy, które trudno wytłumaczyć niematematykom, to dowody niekonstrukttywne. W takim dowodzie autorzy dochodzą do wniosku, iż pewien obiekt matematyczny istnieje, często wiedząc o nim bardzo mało. Dzieje się tak dlatego, że stwierdzamy istnienie takiego obiektu, nie próbując go skonstruować, tylko powołując się na inne fakty. Jednym z najprostszych przykładów jest dowód przez „kradzież strategii”, który pokażę na przykładzie prostej gry.

Grą tą będą *dzielniki*, w które gra się następująco: na początku na tablicy mamy wypisane pewne liczby. Dwaj gracze na przemian wykonują ruchy. Ruch polega na wybraniu nieskreślonej liczby z tablicy i skreśleniu jej wraz ze wszystkimi jej dzielnikami. Ten, kto nie może się ruszyć, przegrywa. Jak widać, gra jest bez elementów losowych i rozstrzyga się w czasie skończonym. Dzięki temu wiemy, że któryś z graczy ma strategię wygrywającą. Teraz czas na właściwe pytanie: rozważmy grę w dzielniki na zbiorze liczb od 1 do 100. Który z graczy ma wówczas strategię wygrywającą?

Jak należy w tę grę grać, żeby wygrać? Nie wiemy... i nie musimy wiedzieć! Przy pytaniu o wygraną interesuje nas „kto” wygra, a nie „jak”. I odpowiemy na to pytanie dzięki następującemu spostrzeżeniu: liczba 1 pełni w tej grze osobliwą rolę – na pewno zostanie skreślona w pierwszym ruchu. Zanim z tego skorzystamy, rozpatrzmy grę w dzielniki na zbiorze liczb od 2 do 100. Który gracz wygrywa w drugiej grze?

Załóżmy wpieryw, że strategię wygrywającą ma wtedy gracz rozpoczynający. W pierwszym ruchu wybiera liczbę x . W takim razie w pierwszej grze również wygrywa gracz rozpoczynający, w pierwszym ruchu skreślając x . Potem wykonuje ruchy jak w grze w dzielniki od 2 do 100, bo to już ta sama gra, w końcu jedynek nie można wybrać.

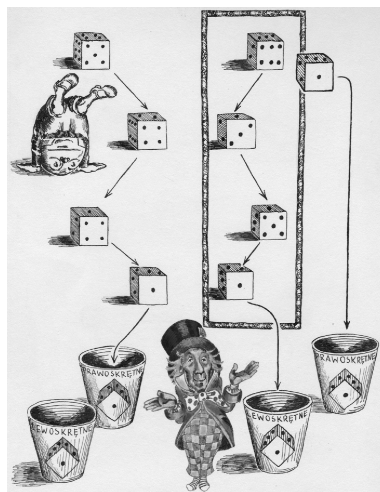
Załóżmy teraz, że gracz drugi ma strategię wygrywającą w grze „od dwójki”. Wtedy grając w grę „od jedynek”, w pierwszej turze skreślamy 1 i od tego momentu zachowujemy się jak gracz drugi w grze „od dwójki”... kradnąc jego strategię i tym samym wygrywając.

Oznacza to, że w grze „od jedynek” gracz rozpoczynający może wygrywać zawsze. Ale jak? Nie wiemy i tym sposobem, niestety, się nie dowiemy...

*Student, Wydział Matematyki,
Informatyki i Mechaniki, Uniwersytet
Warszawski

Co zobaczyła Alicja po drugiej stronie lustra?

Marek KORDOS



Chciałbym opowiedzieć o najtrudniejszym pojęciu matematyki. Najtrudniejszym, choć intuicyjnie prostym i używanym powszechnie również przez niematematyków. Chodzi o orientację.

Zacznijmy od jednego z najbardziej znanych matematyków – to Charles L. Dodgson (1832–1898), student, a później profesor Christ Church College w Oxfordzie. Oczywiście, spieszę uspokoić tych, którzy nie odnajdują w swojej pamięci takiego matematyka: Charles Dodgson nie jest znany jako matematyk i, na dodatek, nie jest znany pod swoim nazwiskiem. To Lewis Carroll, którego dwie książeczki – *Alice's Adventures in Wonderland* (1865) i *Through the Looking Glass* (1875) – należą do (przynajmniej europejskiego) kanonu lektur dziecięcych. Pierwsza z nich ma ustaloną polską wersję tytułu: *Alicja w krainie czarów*, druga zaś znana jest w Polsce pod różnymi tytułami, wśród których jest i *Co zobaczyła Alicja po drugiej stronie lustra?*

No właśnie – co zobaczyła? Nasuwająca się odpowiedź to: zobaczyła siebie. Każdy jednak, kto np. próbował z pomocą lustra coś sobie obciąć, powiedzmy wąsy albo grzywkę, wie, że napotkamy przy takiej czynności pewne trudności i nasza sprawność sterowania ruchami widzianych w lustrze nożyc jest znacznie naruszona. W ogóle wydaje się, że trzymane w prawej ręce nożyce w lustrze trzymamy w ręce lewej.

Pozwala to postawić hipotezę, że lustro zamienia prawą i lewą stronę. Każdy zresztą chętnie poprze nasz pogląd, uważając przy tym, że wie to od dziecka. Jeśli jednak mamy umysł przenikliwy i zwyczaj zadawania sobie i innym pytania *dłaczego?*, natychmiast zaczniemy poszukiwać odpowiedzi, dlaczego akurat prawą i lewą, a nie, powiedzmy dół i górę.

Ufni w potęgę poznawczą doświadczeń, możemy wykonać eksperyment: z nagle, aby kompletnie zaskoczyć lustro, obróćmy je o 90°. Lustro dalej będzie zamieniało prawą i lewą, choć teraz będą one wypadały w tych akurat miejscach lustra, w których poprzednio znajdowały się niezamieniane dół i góra.

I tak pytanie staje się poważne, jak każde pytanie, na które nie umiemy dać zadowalającej odpowiedzi.

Właściwa odpowiedź jest zaskakująca: lustro nie zamienia prawej i lewej. Wydaje się to bzdurą, choć jest prawdą. Zrozumienie tej prawdy jest rzeczą trudną, bo też nie leży ona w optyce, ale głęboko w strukturze używanych przez nas pojęć, właściwie wręcz w filozofii.

Chodzi o to, że używane przez nas pojęcia pochodzą „z różnych szuflad”. Każdy przecież uzna, że *prostopadłość* i *postuluszność* to pojęcia o różnym pochodzeniu – wywodzą się one z zupełnie odmiennych kręgów doświadczeń. Okazuje się jednak, że różne pochodzenie mogą mieć pojęcia dotyczące tego samego obszaru działania. Dla ilustracji przytoczę trzy przykłady. Pierwszy będzie dotyczył etyki, drugi geografii, a trzeci geometrii.

Przykład I. W okresie największego rozkwitu Starożytnych Aten działała grupa filozofów nazywających się sofistami. Podstawowym kierunkiem ich dociekań było badanie reguł prowadzenia dyskusji. W szczególności interesowało ich, jak można wygrać debatę niezależnie od tego, jakie racje się głosi i jakie racje ma do przedstawienia kontrdyskutant. Swoje przemyślenia potrafili też zastosować praktycznie, udzielając (cóż tu ukrywać – odpłatnie) porad politykom, na co w demokratycznym systemie Aten było duże zapotrzebowanie. Sofistą – należącym do końcowego okresu ich działalności – był Sokrates. Jego końcowa konkluzja była taka, że gdy za przesłanki naszego wywodu przyjmiemy jedno stwierdzenie z zakresu prawa naturalnego i drugie stwierdzenie z zakresu prawa moralnego, wówczas będziemy mogli uzasadnić zupełnie dowolną tezę. Na przykład, jeśli w debacie będziemy powoływać się z jednej strony na wolny rynek



Rozwiązanie zadania F 927.
Korzystając z praw Kirchhoffa dla obwodów elektrycznych, znajdujemy, że wartość oporu R powinniśmy obliczać według wzorów:
dla metody 1

$$R = \frac{U_1 R_V}{R_V I_1 - U_1} = \frac{R_V R_1}{R_V - R_1},$$

a dla metody 2

$$R = \frac{U_2}{I_2} - R_A = R_2 - R_A.$$

Odwracając obie zależności, otrzymujemy:

$$R_1 = \frac{R R_V}{R + R_V}$$

oraz

$$R_2 = R + R_A,$$

a więc wzory uproszczone prowadzą do wyniku zaniżonego w przypadku metody 1 i zawyżonego w przypadku metody 2. Porównajmy wartości bezwzględne błędów obu metod i zbadajmy, kiedy metoda 1 jest dokładniejsza:

$$R_2 - R = R_A > R - R_1$$

– ta nierówność jest spełniona, gdy

$$R < \left(R_A + \sqrt{4R_V R_A + R_A^2} \right) / 2 \approx \sqrt{R_V R_A}.$$

Wypisując równość przybliżoną, skorzystaliśmy z faktu, że $R_A \ll R_V$.



i konkurencję, a z drugiej na solidarność międzyludzką i miłość bliźniego, uda się nam bez trudu uzasadnić wszystko, co tylko się komu zamarzy. Można to zresztą zaobserwować w dyskusjach naszych dzisiejszych polityków. Nie byłiby oni jednak zadowoleni z powyższego stwierdzenia. Niezadowoleni byli też politycy współczesnych Sokratesowi Aten, więc oskarżyli go o bezbożność i demoralizację młodzieży. Skazali go następnie na śmierć, każąc mu wypić kielich cykuty, co też uczynił.

Przykład II. Dla określenia swojego położenia na kuli ziemskiej żeglarz musi stwierdzić, na jakiej szerokości i długości geograficznej się znajduje. Szerokość geograficzna w pogodną noc daje się łatwo określić – jest to bowiem pojęcie przyrodnicze: wystarczy stwierdzić, jaki kąt z poziomem tworzy kierunek ku Gwieździe Polarnej. Ten kąt to właśnie szerokość geograficzna na naszej półkuli. Nie ma natomiast doświadczenia przyrodniczego pozwalającego określić długość geograficzną. Już sam południk zerowy jest ustalany umownie – do pierwszej wojny światowej były zresztą trzy takie umowy: brytyjska z południkiem zerowym przechodzącym przez londyńskie przedmieście Greenwich, francuska z zerowym południkiem przez Paryż (mam nawet globus z tak zaznaczonym południkiem zerowym) i hiszpańska – przez Kadyks. Dziś powszechnie przyjmuje się tę pierwszą umowę. Do określenia długości geograficznej (nawet przez GPS!) potrzebna jest znajomość różnicy czasu astronomicznego z Greenwich (np. kiedy jest południe) – ta różnica czasu w godzinach pomnożona przez 15 daje długość geograficzną w stopniach (wschodnią, jeśli nasz zegar jest szybszy od londyńskiego, a zachodnią, gdy jest wolniejszy). W czasach, gdy nie było dostatecznie dokładnych zegarów, kłopoty z ustalaniem długości geograficznej były bardzo poważne (posługiwano się tzw. tablicami astronomicznych efemeryd). Stąd w XVII wieku admiralicje posiadających największe floty Anglii i Holandii ufundowały wielką nagrodę za skonstruowanie morskiego chronometru, co nie udało się mimo starań ani Galileuszowi, ani Huygensowi. Dopiero John Harrison taki chronometr skonstruował i ten przy niewielkich zmianach dotrwał do pojawienia się radia. Inna rzecz, że Harrison musiał przez blisko 20 lat procesować się ze skąpymi admirałami, zanim nagrodę uzyskał.

Przykład III. Jeśli chcemy opisać rozwartość jakiegoś kąta, wystarczy podać, jaką część (właściwą lub nie) kąta pełnego on stanowi. Rozwartość kąta pełnego możemy opisać na różne sposoby: a to, że jest to 360° , a to, że 2π radianów, a to, że 400 gradusów (jak chciała Rewolucja Francuska), a to, że 6 000 tysięcznych (jak chcą artylerzyści) i to wyznaczy rozwartość naszego kąta. Tak czy owak możemy poinformować o rozwartości kąta, nie odwołując się do żadnego materialnego wzorca. Ten fakt określamy, mówiąc, że kąt ma miarę naturalną. Natomiast odcinek nie ma miary naturalnej i, aby podać jego długość, obok podania jakiejś liczby musimy odwołać się do jakiegoś materialnego wzorca. Za czasów mojej młodości była to szyna z dwoma nacięciami przechowywana w Sévres pod Paryżem.

Wróćmy do Alicji. Spróbujmy – pamiętając, że mamy różne możliwości – przyjrzeć się pochodzeniu rozmaitych charakterystycznych kierunków jej ciała (wstydlivi mogą, oczywiście, przyglądać się sobie).

Góra i dół są pojęciami przyrodniczymi. Zarówno fizyka (ciężenie), jak wynikająca stąd anatomia (nogi u dołu, głowa u góry) wyraźnie różnicę góry i dołu wyznaczają.

Przód i tył – też są przyrodnicze, bo każdy z nas coś innego ma z przodu (np. nos) niż z tyłu.

Natomiast lewa i prawa to tylko konwencja. Do dwóch strzałek prostopadłych dobieramy trzecią – można to zrobić na dwa równoważne sposoby. Arbitralność widać w trudnościach, jakie mają dzieci (często do późnej starości) z nauczeniem się, gdzie jest prawo, a gdzie lewo – chodzi o czystą umowę; tego nie można zrozumieć, tego trzeba się nauczyć na pamięć. A żarty w rodzaju; *lewa ręka to ta, gdzie kciuk jest z prawej strony* tylko ową arbitralność podkreślają.



Rozwiązanie zadania F 928.

Zarówno wdychane, jak i wydychane powietrze znajduje się pod ciśnieniem atmosferycznym, a więc możemy przyjąć, że objętość molowa w obu przypadkach jest taka sama i dla temperatury 20°C (temperatura pokojowa) wynosi około 24 l (wszystkie gazy biorące udział w oddychaniu traktujemy jak gazy doskonale). Ułamki objętościowe odpowiadają więc ułamkom molowym składu wdychanego i wydychanego powietrza. Biorąc pod uwagę masy molowe: $\text{N}_2 = 28\text{ g}$, $\text{O}_2 = 32\text{ g}$, $\text{CO}_2 = 44\text{ g}$ i fakt, że w jednym oddechu wymieniamy $x = (0,5\text{ l}/24\text{ l}) = 0,021$ mola powietrza, otrzymujemy, że w każdym oddechu tracimy około 0,01 g. W ciągu 8 godzin snu daje to około 58 g.

Do tej umowy wrócimy za chwilę, ale teraz zakończmy sprawę z lustrem:

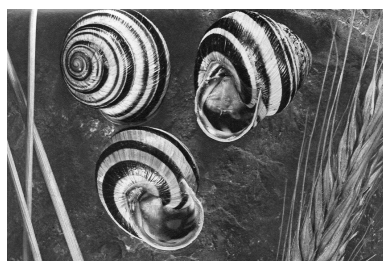
- góry i dołu zmienić nie może;
- przód i tył, oczywiście, zamienia: nie zwróciliśmy na to uwagi, ale przecież stoimy ze swoim odbiciem nos w nos; zmienia strzałkę skierowaną do lustra na skierowaną od lustra;
- prawej i lewej też nie zmienia: przecież ręka z zegarkiem jest z tej samej strony, co była; **to my zmieniamy jej nazwę!**



I tu pojawia się matematyczne pojęcie orientacji. Jest to kolejność wzajemnie prostopadłych strzałek. Wybór tej kolejności to *orientacja* przestrzeni. I tu zdumiewający fakt: zarówno w jedno-, jak w dwu-, trój- i więcej wymiarowej przestrzeni jest tyle samo orientacji – dwie. Czyli w dowolnie wymiarowej przestrzeni każdy zestaw maksymalnej liczby wzajemnie prostopadłych strzałek z ustaloną kolejnością daje się przemieścić tak, że się nałoży na jeden z ustalonych dwu takich zestawów. Utało się, że o jednej orientacji mówimy dodatnia, a o drugiej – ujemna. W przestrzeni trójwymiarowej funkcjonują też techniczne nazwy: prawoskrętna i lewoskrętna.

Matematycznie obie orientacje są równoważne. Ale w szkole – gdzie wszystko musi być jednoznacznie zadekretowane, ustalono, by nazywać je niejako geograficznie: dodatnia (prawoskrętna) orientacja to wschód-północ-góra, na płaszczyźnie wschód-północ i na prostej wschód. Ale nic by się nie zmieniło, gdybyśmy się umówili odwrotnie.

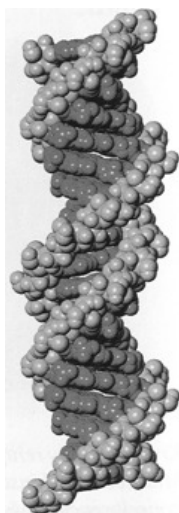
Powstaje pytanie, czy przyroda – jak matematyka – jest wobec tych równoprawnych możliwości sprawiedliwa. A odpowiedź jest – nie. Sprawa ma jednak różne oblicza.



Zacznijmy od siebie – było takie hasło wyborcze: *serce masz po lewej stronie, a po prawej masz wątrobę*. I przeważnie istotnie tak jest. Skąd biorą się takie nieregularności?

Nierównoprawność w wyborze orientacji przez przyrodę dostrzeżono dawno. Najbardziej oczywistym przykładem są muszle ślimaków – rzeczywistość przytłaczająca większość z nich odpowiada orientacji prawoskrętnej. Od pół wieku objaśnia się to tak: jakkolwiek nierównomierność wzrostu muszli – z jednej strony troszeczkę szybciej niż z drugiej – powoduje jej kształt spiralny; proces ten pogłębił się podczas upływu pokoleń, bo odchylenia się genetycznie przekazywały, a w wyniku ewolucji jedna z orientacji stała się dominująca.

Odkrycie helisy DNA (też praktycznie zawsze o tej samej orientacji) nie zmieniło poglądu na tę sprawę – po prostu jedna z orientacji wygasła.



Asymetria jest też w chemii – tam zamiast o orientacji mówi się o *chiralności*. Złożone związki organiczne mają bardzo skomplikowaną strukturę przestrzenną i mogą istnieć w wersjach o odmiennej chiralności (czyli orientacji). Okazuje się, że ma to wielkie znaczenie, gdy zostają wchłonięte przez żywy organizm – pożywne substancje w wersji o odmiennej chiralności niejednokrotnie stają się truciznami. Stąd znaczenie kontrolowania chiralności przy produkowaniu substancji odżywczych czy też zwłaszcza lekarstw. Oczywiście, nie zawsze substancja odmienna chiralnie od milej nam musi być dla nas niemiła: przykładem są zapachy cytrynowy i pomarańczowy, które są swymi lustrzanymi odbiciami.

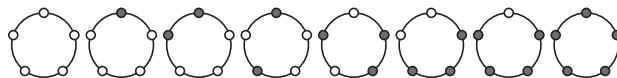
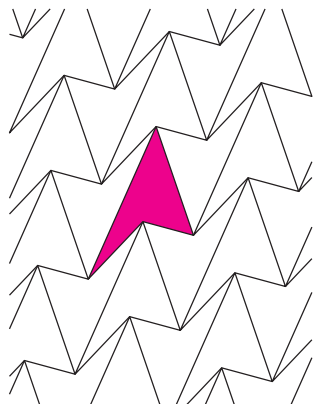
Znaczenie chiralności zostało podkreślone przyznaniem w 2001 roku Nagrody Nobla właśnie za *chirality-controlled synthesis* (chiralność i po angielsku jest chiralnością, bo to słowo pochodzi z greki i oznacza rękę – patrz żart o kciuku). Otrzymali ją Wiliam S. Knowles, Ryoji Noyori i K. Barry Sharpless.

Jeśli pod turystycznym kompasem położymy miedziany drucik w kierunku wskazywanym przez igłę magnetyczną i przez ten drucik przepuścimy prąd (wystarczy paluszek AA), to igła odchyli się. Gdy zmienimy bieguny paluszka (czyli prąd „popłynie w przeciwną stronę”), to igła odchyli się też w przeciwną stronę niż poprzednio. Zatem orientacja pary (przepływ prądu – odchylenie igły) będzie znów taka sama.

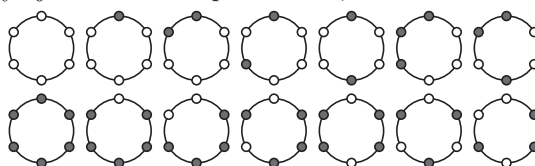


O sytuacjach, w których przy zmianach zachowywana jest orientacja, mówimy, że „niosą” orientację. Nie wątpię, że Czytelnik Sprytny wskaże bez trudu wiele sytuacji, które niosą orientację i wiele takich, które nie niosą.

W geometrii łatwo wskazać zarówno obiekty, które niosą na sobie orientację (np. kostka do gry), jak i te, które nie niosą (np. kula). Nie każdy od razu domyśli się, że malując dwoma kolorami „zegar pięciogodzinny”, nie można uzyskać obrazka niosącego orientację,



a „sześciogodzinny” już można tak pomalować,



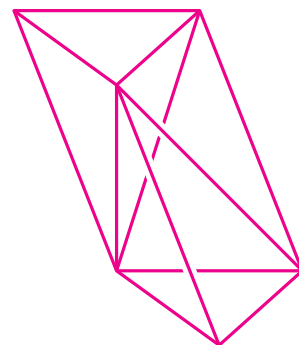
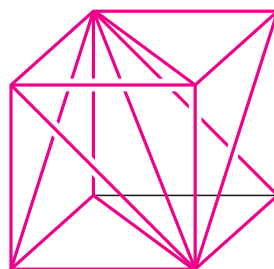
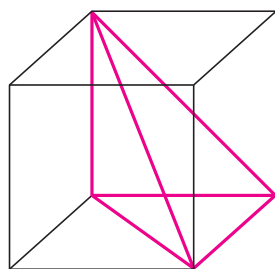
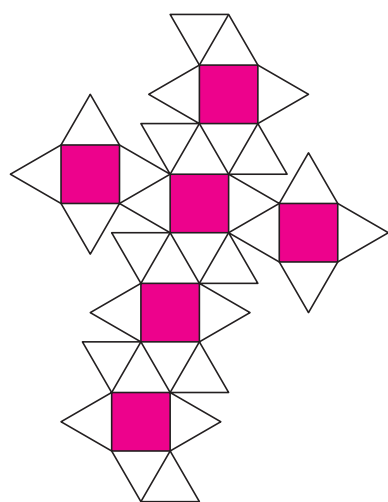
by dwa spośród tych „zegarów” (które?) niosły na sobie orientację.

Orientację niesie na sobie każdy czworokąt różnoboczny, o czym przekonuje rysunek na marginesie. Czytelnik Domyślny zauważy, że patrząc na niego, można odkryć przepis na ułożenie parkietu z dowolnych jednakowych czworokątnych płytek (jaki to przepis?).



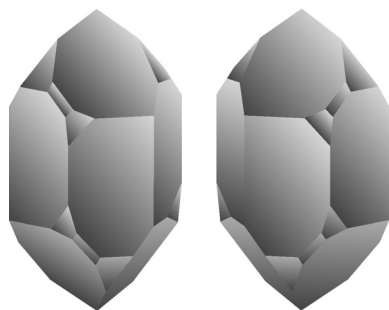
Lustrzany czworokąt już nie pasuje do parkietu.

Podobne, a może jeszcze bardziej złożone mogą być sytuacje w przestrzeni. Czworoscian widoczny poniżej na rysunku z lewej to (jak każdy Czytelnik Dokładny obliczy) jedna szósta sześcianu, w którym został zbudowany. Nie można jednak z jego sześciu egzemplarzy złożyć tego sześcianu. Jak pokazuje rysunek środkowy, trzy z nich muszą być jego lustrzanymi odbiciami.



Mimo to same takie czworosciany o jednakowej orientacji mogą szczelnie wypełnić przestrzeń, o czym przekonuje rysunek prawy: takimi pochylonymi graniastosłupami każdy potrafi „wyparkietować” całą trójwymiarową przestrzeń.

Oczywiście, sześcian nie niesie na sobie orientacji, podobnie jak żaden z pozostałych czterech wielościanów foremnych, czyli platońskich. Natomiast wśród wielościanów półforemnych, czyli archimedesowych (to takie, których ściany są wielokątami foremnymi niekoniecznie jednakowymi, ale naroże w każdym wierzchołku jest takie samo) już takie są. Na marginesie narysowana jest siatka wielościanu archimedesowego, w którego każdym wierzchołku zbiegają się cztery trójkąty równoboczne i kwadrat (to ile jest tych trójkątów?). Jeśli skleimy tę siatkę tak, aby widoczne były pokolorowane ściany kwadratowe, i drugi raz, tak by kolory schowały się w środku, otrzymamy dwa różne, choć lustrzane wielościany – każdy z nich niesie więc orientację.



Naturalne kryształy SiO₂ o odmiennej orientacji

Natura sama też potrafi „skleić” wielościan niosący orientację na dwa sposoby. Obok mamy zdjęcie dwóch kryształów kwarcu, będących swymi odbiciami lustrzanymi. Tym, jakie kryształy niosą na sobie orientację i jakie w ogóle są możliwe, a także tym, czy Natura wykorzystuje wszystkie matematycznie dopuszczalne możliwości, zajmuje się leżąca na pograniczu matematyki i geologii krytalografia. Ale o niej innym razem.

Fizyka a obliczenia równoległe

Paweł ARTYMOWICZ*

Komputery, największy wynalazek nowożytności, mają coraz więcej zastosowań. Jedne pracują w telefonach komórkowych i urządzeniach przenośnych. Innym – superkomputerom – zlecamy np. symulację historii wszechświata. Może kiedyś nauczymy je myśleć podobnie do tego jak sami myślimy. We wstępnym artykule z serii o współczesnych kierunkach w technikach obliczeniowych (zwłaszcza superkomputerowych) pokażemy, jak fizyka tranzystora umożliwiła technologii mikroprocesorowej podwoić prędkość komputerów więcej niż dwadzieścia kolejnych razy (tj. o czynnik $> 2^{20} \approx$ milion razy), umożliwiając szybki internet, smartfony i współczesną naukę obliczeniową, i dlatego kontynuacja dotychczasowego wykładniczego rozwoju techniki komputerowej od trzynastu lat wymaga od programistów zasadniczo nowego podejścia: programowania współbieżnego procesorów wielordzeniowych.

Komputery są niezbędne do coraz szerszej listy zastosowań, od użytkowych (komunikatory osobiste, media i handel internetowy) i rozrywkowych (zaawansowane gry) do naukowych i inżynierskich, czyli od wymiany i analizy wartko rosnącego strumienia informacji, do symulacji rzeczywistości i prób prawdziwie inteligentnego przetwarzania informacji. Gdy wybiegniemy nieco w przyszłość, ale sądząc z rosnącego tempa rozwoju inteligencji maszynowej, w szczególności uczenia maszynowego – w całkiem niedaleką już przyszłość, nasze życie i światowy rynek pracy będą zmieniać się szybko dzięki tej jakościowo nowej fali komputeryzacji w stopniu nie mniejszym od tego, jak mechanizacja produkcji zmieniła społeczeństwa w epoce rewolucji przemysłowej. Niekoniecznie musi to oznaczać dominację maszyn i bezrobocie ludzi. Na przykład to, że pojazdy mechaniczne wyeliminowały konie i woźniców, nie oznaczało bynajmniej spadku zatrudnienia w transporcie. Pojawił się zawód kierowcy. Niedługo znajdzie się on na liście zawodów zagrożonych. Kierowców zastąpią komputery z wielkimi bazami danych, łącznością satelitarną, widzeniem i adaptacyjnym uczeniem maszynowym. W przewidywalnej przyszłości naukowcy czy lekarze i – niestety – również niektórzy politycy pójdą w ślady zapalaczy lamp gazowych (popularna w XIX wieku profesja).

Odpowiednikiem maszyny parowej, motoru rewolucji przemysłowej, jest dziś procesor komputera. Obliczenia równoległe jednego zadania na procesorach wielordzeniowych (tj. na wielu kalkulatorach naraz) mają zaś historyczny odpowiednik w metodzie produkcji masowej rozwiniętej w fabryce samochodów Forda w 1913 roku – zwiększając znacznie tempo obliczeń. Analogia jest o tyle niepełna, że każdy pracownik na taśmie produkcyjnej Forda mógł być (z założenia) mało wykształcony, w odróżnieniu od rzemieślnika tworzącego cały produkt samemu. Dzisiaj natomiast, jak zobaczymy, zachodzi konieczność przestawienia się programisty z myślenia i programowania sekwencyjnego na tworzenie programów współbieżnych albo równoległych (jak jednoczesna produkcja samochodów na równoległych taśmach produkcyjnych). A to wymaga nie mniejszych, lecz większych umiejętności niż programowanie zwykłego komputera von Neumanna, który pobiera i wykonuje jedną po drugiej instrukcje przy użyciu jednego kalkulatora (rdzenia obliczeniowego).

Postęp geometryczny

W latach 60. ubiegłego wieku współzałożyciel firmy Intel, Gordon Moore, sformułował empiryczną regułę geometrycznego postępu szybkości komputerów w ich kolejnych, regularnie co dwa lata (lub nawet nieco częściej) powstających pokoleniach procesorów komputerowych. Ścisłej, prawo Moore'a mówi, że liczba tranzystorów na jednostkę powierzchni, tj. gęstość



Rozwiązanie zadania M 1530.
Niech $\ell(P)$ będzie długością łuku (mierzoną zgodnie z ruchem wskazówek zegara) łączącego P_0 z P , tzn. dla każdego $i = 0, 1, 2, \dots, 2^n - 1$

$$\ell(P_i) = \frac{i(i+1)}{2} \pmod{2^n}.$$

Z założeń zadania wynika, że ℓ jest bijekcją zbioru wierzchołków 2^n -kąta i $\mathbb{Z} \cap [0, 2^n - 1]$.

Udowodnimy, że dla $n \geq 3$ zachodzi równość

$$\ell(\mathcal{E}) = \ell(\mathcal{F}) + 2^{n-2} \pmod{2^n},$$

czyli że zbiór \mathcal{E} jest obrazem \mathcal{F} przy obrocie o 90° wokół środka danego okręgu zgodnie z ruchem wskazówek zegara. Konkretnie, wykażemy, że dla każdego $P_i \in \mathcal{F}$

$$\ell(P_{f(i)}) \equiv \ell(P_i) + 2^{n-2} \pmod{2^n},$$

gdzie funkcja $f: \{i: P_i \in \mathcal{F}\} \rightarrow \{i: P_i \in \mathcal{E}\}$ określona jest następująco

$$f(i) = \begin{cases} 2^{n-1} + i & \text{dla } 2 \mid i, \\ 2^{n-1} - i - 1 & \text{dla } 2 \nmid i. \end{cases}$$

W obliczu definicji funkcji ℓ wystarczy więc sprawdzić, że dla każdego $i < 2^{n-1}$ liczba

$$g(i) := f(i)(f(i)+1) - i(i+1) - 2^{n-1}$$

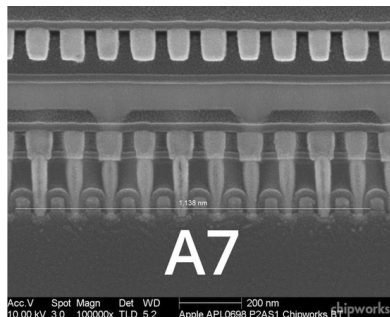
jest podzielna przez 2^{n-1} . Rzeczywiście, bezpośrednio z definicji funkcji f otrzymujemy, że jeżeli $2 \mid i$, to

$$g(i) = 2^{2n-2} + i^2 \equiv 0 \pmod{2^{n-1}},$$

a jeżeli $2 \nmid i$, to

$$g(i) = 2^{2n-2} - (i+1)2^n \equiv 0 \pmod{2^{n-1}},$$

gdyż 2^{2n-2} dzieli się przez 2^{n-1} dla $n \geq 3$. Pozostaje bezpośrednio sprawdzić, że dla $n = 2$ rozważane zbiory także są przystające (odpowiednia izometria znów jest obrotem o 90° , ale w przeciwną stronę).



Rys. 1. Fragment procesora Apple A7, gdzie dziesięć tranzystorów ułożonych jest na odcinku 1138 nm. Procesor ten, znajdujący się w telefonach iPhone 5S, oparty jest na połączeniach o grubości 28 nm, tj. rzędu 70 odległości między atomami krzemu. (W latach 2010–2014 szerokość ścieżki procesorów Intel Corp. zmalała z 32 i 28 nm do 22 nm, a następnie do 14 nm). Szybko rosnący koszt produkcji tych niezawodnych, mikroskopijnych układów musi się producentom zwrócić. To zmusza ich do nieco rzadszej niż dawniej zmiany technologii produkcji. Pesymiści ogłosili wręcz, że pokolenie procesorów w technologii 10 nm będzie już ostatnim opłacalnym, czyli końcem technologii krzemowej. To niekoniecznie prawda. W odwodzie są też nowe substraty, np. grafen, prawdziwie 2-wymiarowa, heksagonalna sieć atomów węgla, która może kiedyś zastąpić krzem. Istnieje poważny, atomowo-kwantowy limit miniaturyzacji, związany z odległościami między atomami krzemu równymi 0,43 nm. Standardowy tranzystor nie będzie działał klasycznie po miniaturyzacji do kilku nanometrów i schemat funkcjonowania komputera musi być wtedy wymyślony od podstaw.

upakowania tranzystorów w płaskim układzie scalonym podwaja się co około dwa lata. Prawo jest szeroko znane, gdyż faktycznie tak zmieniała się ta liczba przez niesłychanie wiele pokoleń technologii procesorowych: ponad 20 (czyli przez ostatnie 40 lat) i ponieważ taki wykładniczy przyrost liczby bramek logicznych i komórek pamięci lokalnej w procesorze leży u podstaw błyskotliwej kariery komputera, a w miarę jak komputery stają się niezbędne, postępu technicznego cywilizacji. Spójrzmy na liczby opisujące typowy zaawansowany procesor główny (CPU, od ang. *Central Processor Unit*) w roku 1975 i 2015. Szybkość działania CPU mierzona jest liczbą najprostszych działań arytmetycznych zmiennoprzecinkowych na sekundę (jednostka zapisywana jako FLOP albo FLOP/s, to właśnie oznacza: floating point operation per second). Wynosiła w roku 1975 około 0,1 MFLOP (0,1 miliona FLOP). Czterdzieści lat później procesor CPU był już milion razy szybszy, 0,1–0,3 TFLOP (teraflop to 10^{12} działań/s), a koprocesory obliczeniowe osiągnęły 1 TFLOP w podwójnej precyzji obliczeń. Liczba tranzystorów w procesorach dochodzi do około 10 miliardów.

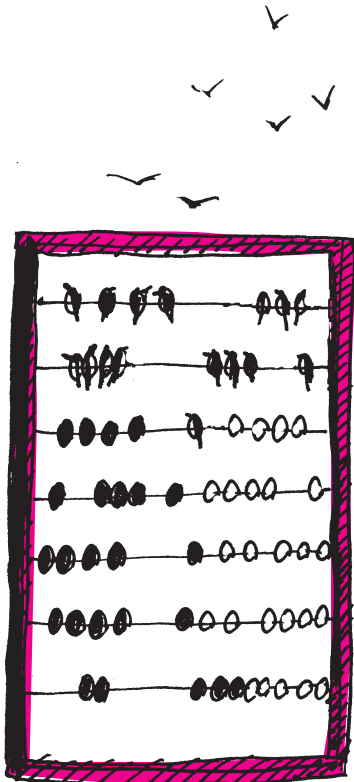
Prawo Moore'a nie było przypadkiem. Przeciwnie, było decyzją przemysłu elektronicznego, umożliwioną szczęśliwie przez prawa fizyki. Projektanci procesorów (np. Intel) miniaturyzowali podzespoły tak, by następna generacja procesora na tej samej powierzchni mieściła dwa razy więcej tranzystorów. To prawo Moore'a. Ale gdyby krok miniaturyzacji wymuszał np. dwa razy większe straty cieplne, to taka miniaturyzacja musiałaby być zatrzymana po kilku etapach miniaturyzacji. Kontynuowana, doprowadziłaby do tego, że obecnie komputer domowy zużywałby megawaty mocy, a telefon komórkowy rozładowywałby baterię w ułamku sekundy. Sprawdźmy zatem konsekwencje skalowania rozmiaru dla mocy pobieranej przez procesor, zamienianej na ciepło wskutek zjawiska oporności elektrycznej. Odprowadzenie znacznej ilości ciepła jest w istocie większym wyzwaniem, niż dostarczenie energii elektrycznej, choć koszty energii są niebagatelne. W przypadku superkomputerów są porównywalne z kosztem ich zakupu.

Bariera energii

Pierwsze mikroprocesory mało się grzały. Za to teraz zderzyliśmy się we wszystkich rodzajach obliczeń, od telefonii do centrum obliczeniowego, z barierą energetyczną. Pojedyncze procesory (CPU, koprocesory arytmetyczne i karty graficzne) nie mogą pobierać więcej niż $P = 200$ W, a najwyżej 300 W, inaczej byłoby trudno je zasilać w budynku mieszkalnym i równie trudno byłoby wentylować pokoje komputerowe. Na limit mocy napotykać też procesory w telefonach komórkowych i tabletach: długość pracy urządzenia bez ładowania baterii jest odwrotnie proporcjonalna do P , co nakłada na moc górne ograniczenie.

Energia elektryczna potrzebna do obliczeń szybko rośnie z częstotliwością f zegara procesora, jak i liczbą N tranzystorów. Bramka logiczna ma jeden lub więcej tranzystorów o efektywnej pojemności elektrycznej C (jest to stosunek ładunku do potencjału elektrycznego, $C = Q/V$). Ma także pewną oporność. Zmiana stanu bramki wymaga naładowania pojemności C poprzez opornik od potencjału zerowego do napięcia operacyjnego V . Przepływając przez źródło napięcia V , ładunek Q uzyskuje energię QV , czyli CV^2 . Sumowanie przyczynków do elektrostatycznej energii potencjalnej tranzystora pokazuje, że niezależnie od wartości oporności bramki połowa energii źródła gromadzona jest na tranzystorze, a połowa zamieniana w czasie ładowania na ciepło. Tylko nieliczne procesory potrafią odzyskać tę pierwszą energię potencjalną, większość w końcu traci całość dostarczonej energii CV^2 . Maksymalna moc rozpraszana w czasie obliczeń prowadzonych f razy na sekundę na N bramkach logicznych procesora zależy więc od f i V :

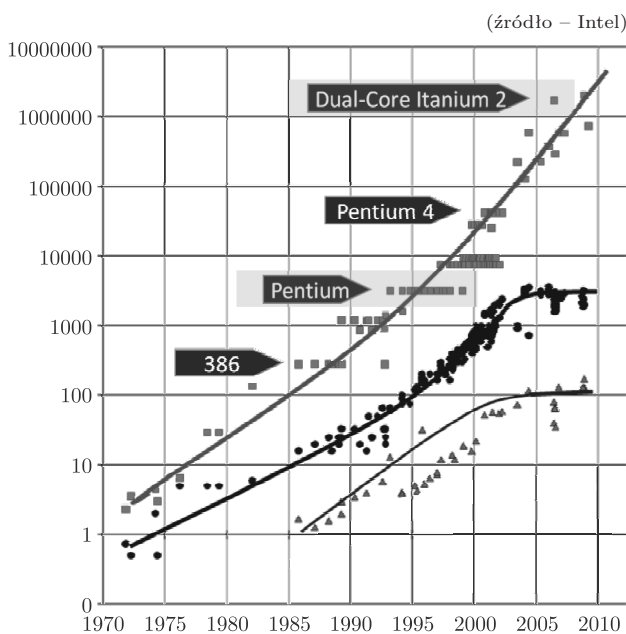
$$P \sim N C V^2 f.$$



Jak zmieniają się parametry tranzystora przy zmianie generacji, tj. zmniejszeniu jego struktury o czynnik liniowy $\sqrt{2}$, a pola powierzchni o czynnik 2? Spójrzmy na kolejne czynniki powyższego wyrażenia na moc P . Liczba tranzystorów N podwaja się. Pojemność C to pewna stała materiałowa razy pole powierzchni elektrod (np. okładek kondensatora), podzielone przez odległość między nimi (to wynika z praw Coulomba i Gaussa w elektrostatyce). Ponieważ pole powierzchni spada dwukrotnie, a wszystkie odległości o czynnik skalowania $\sqrt{2}$, to pojemność C bramki spada o czynnik $\sqrt{2}$. Inżynierowie przy zmniejszeniu rozmiarów tranzystorów zmniejszali też o czynnik $\sqrt{2}$ różnicę potencjału V tak, by pole elektryczne (tzn. potencjał podzielony przez odległość elektrod) pozostało stałe. Przy niezmiennym V pole elektryczne za bardzo by w końcu wzrosło, niszcząc właściwości półprzewodnika. Zmiany stanu półprzewodników w tranzystorze zajmują pewien czas, który jest tym większy, im grubsze są warstwy półprzewodnika. I odwrotnie – zmniejszenie grubości pozwalało na zmniejszenie czasu przeładowania tranzystora. Dlatego częstotliwość f była zwiększana przy miniaturyzacji $\sqrt{2}$ razy. Zmiany wielkości po prawej stronie wyrażenia na P zebrane w jedną liczbę dają stałą wartość równą

$$(2)(1/\sqrt{2})(1/\sqrt{2})^2(\sqrt{2}) = 1.$$

I tu właśnie leży tajemnica trzydziestoletniego sukcesu procedury skalowania: nie wymagało ono dostarczenia większej mocy tranzystorom na centymetrze kwadratowym powierzchni. Pobór mocy procesora rósł z innego powodu, takiego, że pole jego powierzchni w umiarkowanym tempie zwiększano. Liczba tranzystorów powiększała się o więcej niż czynnik 2. Można było użyć tak liczne tranzystory do zwiększania złożoności logiki procesora oraz na wewnętrzną, szybką pamięć zwaną podręczną (ang. *cache*). Było to potrzebne do maskowania narastającej od 2000 r. powolności pamięci operacyjnej RAM (*Random Access Memory*) w stosunku do CPU. Zastosowano coraz bardziej skomplikowane kolejki danych i instrukcji ściąganych przedwcześnie do pamięci podręcznej. Dzięki takim metodom szybkość zegara pomiędzy generacjami procesorów rosła nie o czynnik $\sqrt{2}$, lecz około dwukrotnie. Tempo rozwoju techniki obliczeniowej było fenomenalne.



Rys. 2. Wzrost w kolejnych latach liczby tranzystorów N w tysiącach (■), częstotliwości zegara procesora f w megahercach (●) i mocy elektrycznej jego zasilania P w watach (▲).

Idylla skalowania zakończyła się w latach 2003–2004, kiedy inne niż oporność elektryczna zjawisko dorównało dyssypacji energii w tranzystorze. Jest to prąd ucieczki (ang. *leakage current*), przeciekanie elektronów przez nominalnie nieprzewodzącą warstwę półprzewodnika. Zjawisko było drugorzędne przy dużych początkowo wartościach V , ale po kolejnym zmniejszeniu V dyssypacja energii związana z prądem ucieczki zaczęła dominować. Wynik był natychmiastowy, choć nienagłaśniany. Od tej pory utrzymywano $V \approx \text{const}$. Aby wydzielana moc cieplna P nie przekroczyła bariery energetycznej, konieczne też było praktyczne zamrożenie szybkiego uprzednio przyrostu częstotliwości f . Ilustruje to rysunek 2 pokazujący historię zmian N , f i P w latach 1970–2010. Po roku 2004 szybki przyrost f zakończył się. Dla oszczędności energii redukowano w niektórych urządzeniach f poniżej wartości 2–4 GHz osiągniętych wcześniej. Tak jest np. w nowych procesorach arytmetycznych Intel Xeon Phi oraz procesorach graficznych (GPU), które mają obecnie $f \sim 1\text{--}1,5$ GHz. W cieniu bariery energetycznej wymuszającej $P < 300$ W, $f < 4$ GHz oraz $V \approx \text{const}$, prędkość obliczeń dość trudno jest dalej podnosić. Jest to absolutnie niemożliwe, jeśli zachowamy dawną architekturę procesora z tylko jednym, szybkim, centralnym kalkulatorem.



Mysleć równolegle

Świat (techniki obliczeniowej) uratowała po roku 2004 wielordzeniowość procesorów. To, że sprzedaje się teraz w sklepach procesory o coraz większej liczbie niezależnych rdzeni N_r , przyspiesza komputer proporcjonalnie do N_r , ale tylko pod warunkiem odpowiedniego dostosowania algorytmów. Gdyby nie działała nieubłagana fizyka, w tej chwili moglibyśmy liczyć na jednym rdzeniu CPU o $f = 30$ GHz, zamiast na 10 rdzeniach CPU o $f = 3$ GHz albo na 60 rdzeniach procesora Xeon Phi o zegarze 1,1 GHz. Programowanie byłoby tradycyjne, łatwiejsze, i wszystkie rodzaje zadań wykonywałyby się szybko. Ale w realnym świecie nie moglibyśmy tego robić w budynku mieszkalnym, gdzie nadmierny pobór prądu aktywowałby bezpieczniki!

Obecna ewolucja procesora polega na powielaniu w mniejszej skali przestrzennej rdzenia obliczeniowego o ograniczonej liczbie tranzystorów tak, aby obliczenia szły coraz większą liczbą równoległych torów (wątków programu). To pozwoliło procesorom kontynuować bicie rekordów sumarycznej mocy obliczeniowej (to temat następnego artykułu z tej serii). Ułatwiło procesorom, ale nie wszystkim programistom, i nie we wszystkich zastosowaniach. Rewolucja wielowątkowości została programistom narzucona przez inżynierów, ale – jak widzieliśmy – był bardzo istotny powód: energetyka tranzystora. Niektórzy opierają się do dziś idei programowania kart graficznych do celów obliczeniowych, mimo że do niektórych zastosowań w tej dekadzie będą to urządzenia zdecydowanie najszybsze (m.in. do uczenia maszynowego). Dla nich najlepszym rozwiązaniem mogą być procesory MIC (*Many Integrated Cores*), czyli Liczne Zintegrowane Rdzenie, jak Xeon Phi firmy Intel. Uruchamianie na nich programu, zwłaszcza wcześniej działającego programu sekwencyjnego, trwa krócej niż w przypadku GPU, urządzenia o podobnej liczbie N , ale odmiennej architekturze i hierarchii pamięci. W obu przypadkach nietrywialne programowanie nowoczesnych maszyn obliczeniowych jest jednak ciekawe, gdyż ich moce obliczeniowe nadal bardzo szybko rosną i pozwalają atakować dotychczas nierozwiązywalne problemy. Celem wytyczonym przez ministerstwo energii USA jest budowa do 2020 roku komputera robiącego 10^{18} działań arytmetycznych na sekundę (exaflop). Można też zbudować mini-superkomputer u siebie w domu i prowadzić obliczenia, np. dynamiki gazu, o czym opowie trzeci odcinek z tej serii.

A co dalej? Komputer kwantowy

Przewodniki w najnowszych CPU mają grubość 14 nm albo 40 warstw atomowych krzemu. Zmniejszenie ich 10 razy spowoduje, że procesor znacznie zachowywać się dziwnie, nieprzewidywalnie, gdyż elektrony i atomy będą przejawiały cechy kwantowe. Koncepcję obliczeń kwantowych sformułował fizyk, Richard Feynman, w 1981 roku. W wielu laboratoriach na świecie już teraz fizycy i inżynierowie dokonali w praktyce przeskoku do warstw jednoatomowych i koniecznego przy tym przejścia do fizyki kwantowej. Oczekujemy, że za około dziesięć lat zbudowane zostaną pierwsze użytkowe egzemplarze komputerów kwantowych. Cząstki reprezentować mogą jednocześnie wartości logiczne zera i jedynki na zasadzie superpozycji stanów kwantowych (w takim podwójnym stanie jest kot w słynnym eksperymencie myślowym Schrödingera). Jest więc pewne, że można będzie niesłychanie szybko sprawdzać wyniki wszystkich możliwych kombinacji zer i jedynek w obliczeniu probabilistycznym bądź kryptologicznym, jak i w niektórych problemach optymalizacji. Problem rozkładu dużej liczby naturalnej na czynniki pierwsze (ważny przy deszyfracji kodów) został kwantowo-algorytmicznie rozwiązany 23 lata temu. Nie jest tylko jasne, do ilu problemów spoza kombinatoryki i probabilistyki komputer kwantowy będzie się nadawał, czy będzie miał funkcjonalność komputera von Neumanna. Jeśli tak, to czeka nas zmiana paradygmatu obliczeń. Entuzjastyczne badania trwają, ale przez najbliższe 25 lat nie pozbywałbym się jeszcze komputera klasycznego.

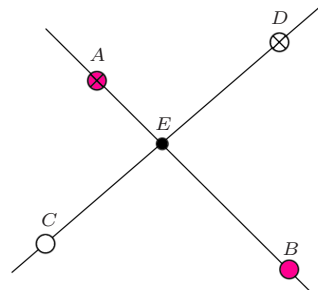


Kolorowanie płaszczyzny, prostych i okręgów

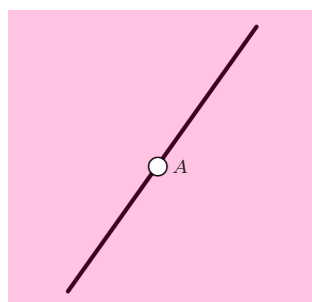
Jadwiga CZYŻEWSKA*

Na drugim etapie VIII Olimpiady Matematycznej Gimnazjalistów (obecnie Olimpiady Matematycznej Juniorów) pojawiło się następujące zadanie:

Każdy punkt płaszczyzny należy pomalować na pewien kolor w taki sposób, aby każda prosta była jednokolorowa lub dwukolorowa. Jaka jest największa możliwa liczba kolorów, których można użyć do pomalowania punktów tej płaszczyzny? Odpowiedź uzasadnij.



Rys. 1



Rys. 2

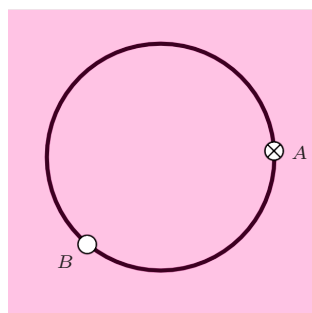
Odpowiedź brzmi: trzy kolory. Udowodnijmy najpierw nie wprost, że nie możemy użyć czterech kolorów. Załóżmy, że można użyć czterech kolorów – wybierzmy zatem cztery różnokolorowe punkty. Zauważmy, że żadne trzy z nich nie są współliniowe, bo w innym przypadku znaleźlibyśmy trzykolorową prostą. Oznaczmy wybrane przez nas punkty A, B, C, D w taki sposób, by proste AB i CD się przecinały. Ich wspólny punkt oznaczmy jako E (rys. 1). Jeśli punkt E ma taki sam kolor co punkt A czy B , to prosta CD jest trzykolorowa. W innym przypadku prosta AB jest trzykolorowa. Dowiedliśmy, że kolorując płaszczyznę przy użyciu czterech kolorów, zawsze znajdziemy trzykolorową prostą. Teraz pokażę kolorowanie płaszczyzny trzema kolorami, spełniające warunki zadania. Pokolorujmy całą płaszczyznę jednym kolorem, np. zielonym. Wybierzmy dowolną prostą i pokolorujmy ją innym kolorem, np. niebieskim. Na prostej wybierzmy dowolny punkt i pokolorujmy go trzecim kolorem, np. czerwonym. Wtedy każda prosta jest co najwyżej dwukolorowa (rys. 2).

Po zawodach zadałam sobie następujące pytanie. Załóżmy, że każda prosta jest co najwyżej trzykolorowa. Jak wiele kolorów mogliśmy wtedy użyć do pokolorowania punktów płaszczyzny? Odpowiedź jest w tym przypadku zaskakująca: mogliśmy użyć nieskończenie wiele kolorów! Pokolorujmy całą płaszczyznę jednym kolorem. Wybierzmy dowolny okrąg i pokolorujmy każdy jego punkt innym kolorem. Jeśli prosta nie ma punktów wspólnych z okręgiem, jest jednokolorowa. Jeśli jest do niego styczna, jest dwukolorowa, a jeśli prosta przecina okrąg, jest trzykolorowa. Możemy postawić analogiczne pytania dla okręgów: na ile kolorów możemy pokolorować punkty płaszczyzny, jeśli każdy okrąg może być co najwyżej dwu- czy trzykolorowy. Rozwiązania tych problemów są bardzo podobne do tych postawionych dla prostych, dlatego znalezienie odpowiedzi pozostawiam Czytelnikowi.

Moja praca odpowiada na jeszcze bardziej złożone pytanie niż to postawione podczas olimpiady: na ile kolorów można pokolorować punkty płaszczyzny, jeśli każda prosta może być co najwyżej m -kolorowa, a każdy okrąg co najwyżej n -kolorowy. W pracy udało mi się znaleźć odpowiedź dla prawie wszystkich wartości m i n – wyjątkiem jest przypadek $m = n = 3$ – umiem pokazać, że w takim przypadku możemy użyć czterech kolorów, ale nie sześciu. W dalszej części artykułu zaprezentuję dowody dotyczące tego problemu. Przypadek pięciu kolorów jest otwarty.

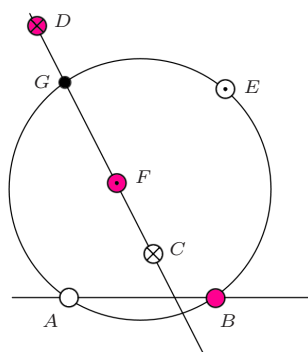
Twierdzenie. *Niech k będzie maksymalną liczbą kolorów, na które możemy pokolorować punkty płaszczyzny tak, że każda prosta i każdy okrąg będą co najwyżej trzykolorowe. Wtedy $4 \leq k \leq 6$.*

Na początku pokażę kolorowanie płaszczyzny czterema kolorami, takie, że każda prosta i każdy okrąg są co najwyżej trzykolorowe. Pokolorujmy całą płaszczyznę jednym kolorem, np. fioletowym. Wybierzmy okrąg i pokolorujmy go innym kolorem, np. brązowym. Na okręgu wybierzmy dwa punkty i pokolorujmy je na jeszcze inne kolory, np. czerwony i zielony. Wtedy każdy okrąg i każda prosta będą co najwyżej trzykolorowe (rys. 3). Stąd $k \geq 4$.

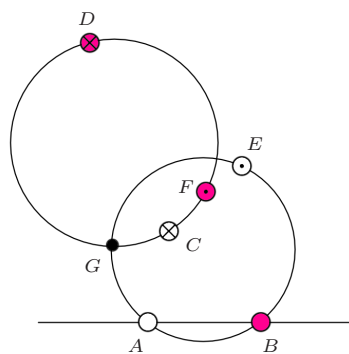


Rys. 3

Założmy teraz, że możemy pokolorować płaszczyznę sześcioma kolorami – wybierzmy sześć różnokolorowych punktów. Zauważmy, że możemy znaleźć prostą przechodzącą przez co najmniej dwa punkty, taką że pozostałe wybrane punkty będą leżeć po jednej stronie tej prostej. Z drugiej strony zauważmy, że na takiej prostej mogą leżeć co najwyżej trzy wybrane punkty, bo w innym przypadku znaleźlibyśmy czterokolorową prostą. Rozpatrujemy zatem dwa przypadki ze względu na liczbę wyróżnionych punktów na wybranej prostej. Poniżej przedstawię jeden z nich (rozumowanie dla drugiego z nich jest analogiczne i pozostawiam je Czytelnikowi).



Rys. 4



Rys. 5

Założmy, że na prostej leżą dokładnie dwa wybrane punkty – oznaczmy je jako A i B , a pozostałe jako P_1, P_2, P_3, P_4 . Spośród kątów AP_1B, AP_2B, AP_3B i AP_4B wybierzmy ten o największej mierze – AP_iB (zauważmy, że gdyby były dwa takie kąty, to znaleźlibyśmy czterokolorowy okrąg). Punkt P_i oznaczmy jako C . Spośród pozostałych kątów wybierzmy ten o minimalnej mierze – AP_jB (jest tylko jeden taki kąt). Punkt P_j oznaczmy jako D . Pozostałe punkty oznaczmy jako E i F . Zauważmy, że punkt C leży wewnątrz okręgu AEB . Rozpatrzmy teraz dwa przypadki:

1. Punkty C, D i F są współliniowe. Jeden z punktów przecięcia prostej CDF z okręgiem AEB oznaczam jako G . Jeśli punkt G ma taki sam kolor co któryś z punktów C, D lub F , to okrąg AEB jest czterokolorowy. W innym przypadku prosta CDF jest czterokolorowa (rys. 4).
2. Punkty C, D, F nie są współliniowe. Jeden z punktów przecięcia okręgu CDF z okręgiem AEB oznaczam jako G . Jeśli punkt G ma taki sam kolor co któryś z punktów C, D, F , to okrąg AEB jest trzykolorowy. W innym przypadku okrąg CDF jest trzykolorowy (rys. 5).

W każdym przypadku (także dla trzech wybranych punktów na prostej) znajdujemy czterokolorową prostą lub okrąg. Zatem nie możemy wykorzystać sześciu kolorów do pokolorowania płaszczyzny w taki sposób, żeby każda prosta i każdy okrąg były co najwyżej trzykolorowe. Zatem $4 \leq k \leq 6$. Co ciekawe, w przypadku gdy kolorujemy punkty płaszczyzny z zachowaniem warunku, że każda prosta jest co najwyżej trzykolorowa, a każdy okrąg co najwyżej czterokolorowy, możemy użyć nieskończenie wielu kolorów. Aby to wykazać, potrzebujemy następującego lematu:

Lemat. *Istnieje taki zbiór punktów \mathcal{X} o nieskończonej liczbie elementów, że żadne trzy punkty należące do \mathcal{X} nie leżą na jednej prostej i żadne cztery nie leżą na jednym okręgu.*

Zdefiniujmy przez indukcję ciąg zbiorów $\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_n, \dots$, które spełniają następujące warunki:

1. $\mathcal{X}_1 \subseteq \mathcal{X}_2 \subseteq \dots \subseteq \mathcal{X}_n \subseteq \dots$,
2. dla każdego n żadne trzy punkty należące do zbioru \mathcal{X}_n nie są współliniowe,
3. dla każdego n żadne cztery punkty należące do zbioru \mathcal{X}_n nie leżą na jednym okręgu,
4. zbiór \mathcal{X}_n zawiera $n + 3$ punkty.

Zbiór \mathcal{X}_1 zawiera wierzchołki dowolnego trójkąta i jego ortocentrum, zatem spełnia wszystkie żądane warunki. Nowy zbiór \mathcal{X}_{n+1} ze zbioru \mathcal{X}_n tworzymy w następujący sposób: przez każde dwa punkty zbioru \mathcal{X}_n prowadzimy prostą, a przez każde trzy okrąg. Niech narysowane obiekty tworzą razem zbiór \mathcal{Z} . Zauważmy, że liczba prostych i okręgów jest skończona. Wybierzmy prostą, która nie należy do zbioru \mathcal{Z} . Proste i okręgi przecinają ją w skończonej liczbie punktów. Wybierzmy zatem punkt P , który nie leży na żadnej prostej czy okręgu należącym do zbioru \mathcal{Z} . Wtedy $\mathcal{X}_{n+1} = \mathcal{X}_n \cup \{P\}$. Rozważmy zbiór \mathcal{X} , będący sumą wszystkich zbiorów \mathcal{X}_n . Zauważmy, że zbiór \mathcal{X} zawiera nieskończenie wiele elementów. Ponadto żadne trzy punkty należące do tego zbioru nie są współliniowe, a żadne cztery nie leżą na jednym okręgu. Zatem zbiór \mathcal{X} jest szukanym przez nas zbiorem.



Jak, mając taki zbiór, możemy pokolorować płaszczyznę tak, by każda prosta była co najwyżej trzykolorowa, a każdy okrąg co najwyżej czterokolorowy? Pokolorujmy całą płaszczyznę jednym kolorem. Następnie weźmy taki zbiór punktów \mathcal{X} o nieskończonej liczbie elementów, że żadne trzy punkty, które do niego należą, nie są współliniowe, żadne cztery nie leżą na jednym okręgu. Każdy punkt należący do zbioru \mathcal{X} pokolorujmy innym kolorem. Wtedy każda prosta będzie co najwyżej trzykolorowa, a każdy okrąg co najwyżej czterokolorowy. Rozwinięte zadanie z Olimpiady Matematycznej Juniorów okazało się bardzo ciekawym problemem. Obecnie pracuję nad jego uogólnieniem w trzecim wymiarze.

Kończąc, chciałabym bardzo podziękować opiekunowi mojej pracy, Panu Wojciechowi Guzickiemu za zaproponowanie mi tego tematu oraz pomoc przy tworzeniu pracy.



Zadania

Redaguje *Lukasz BOŻYK*

M 1528. Każde pole tablicy o wymiarach 8×8 należy pomalować na czarno albo na białą. Na ile sposobów można uczynić to tak, aby każdy kwadrat 2×2 zawierał parzystą liczbę czarnych pól?

Rozwiązanie na str. 1

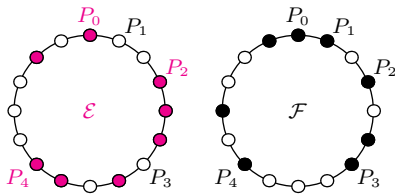
M 1529. Sześcian przecięto płaszczyzną, uzyskując w przekroju pięciokąt opisany na okręgu. Udowodnić, że ten pięciokąt ma oś symetrii.

Rozwiązanie na str. 2

M 1530. Na okręgu o długości 2^n ($n \geq 2$) znajdują się punkty $P_0, P_1, \dots, P_{2^n-1}$ będące wierzchołkami 2^n -kąta foremego, oznaczone w taki sposób, że długość łuku $P_{i-1}P_i$, mierzonego zgodnie z ruchem wskazówek zegara, jest równa i dla każdego $i = 1, 2, \dots, 2^n - 1$. Niech

$$\mathcal{E} = \{P_i : 2 \mid i\} \quad \text{oraz} \quad \mathcal{F} = \{P_i : i < 2^{n-1}\}.$$

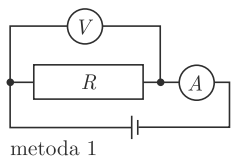
Udowodnić, że \mathcal{E} i \mathcal{F} są przystające (jako podzbiory płaszczyzny).



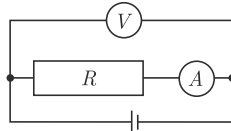
Uwaga. Można udowodnić, że dla każdego $n \geq 2$ istnieje etykietowanie wierzchołków 2^n -kąta foremego o opisanych własnościach – jest to równoważne zadaniu 2 z I etapu LX OM, którego rozwiązanie można znaleźć na stronie archom.ptm.org.pl/?q=node/9. Rysunek przedstawia przypadek $n = 4$ z zaznaczonymi zbiorami \mathcal{E} oraz \mathcal{F} .

Rozwiązanie na str. 8

Przygotował *Andrzej MAJHOFER*



metoda 1



metoda 2

F 927. Mamy do dyspozycji woltomierz o oporze wewnętrznym R_V i amperomierz o oporze wewnętrznym R_A . W celu wyznaczenia nieznannej wartości oporu R opornika możemy zestawić obwód pomiarowy na dwa różne sposoby (rysunek). Wykonując pomiary metodą 1, odczytaliśmy wartość prądu równą I_1 i napięcie równe U_1 , a pomiary metodą 2 dały wyniki I_2 i U_2 . W jakich warunkach wartość obliczona na podstawie uproszczonego wzoru $R_1 = U_1/I_1$ jest dokładniejszym przybliżeniem dokładnej wartości R niż wartość $R_2 = U_2/I_2$? Zakładamy, że każdorazowo woltomierz poprawnie pokazuje różnicę potencjałów na jego zaciskach, a amperomierz poprawnie podaje wartość płynącego przezeń prądu.

Rozwiązanie na str. 4

F 928. Oszacuj, o ile zmienia się masa człowieka w ciągu dobrze przespanej nocy (8 godzin snu). Dorosły człowiek w spoczynku wykonuje średnio 12 oddechów na minutę, a w każdym oddechu „wymienia” około 0,5 l powietrza. Skład wdychanego powietrza to (objętościowo) w 78% azot i 21% tlen, a wydychanego powietrza w 78% azot, 17% tlen i 4% dwutlenek węgla.

Rozwiązanie na str. 5

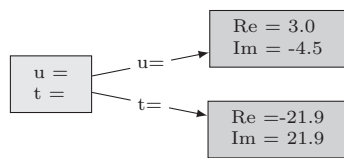
Wiszące referencje.

Czy można wyeliminować to zagrożenie?

Andrzej SALWICKI*

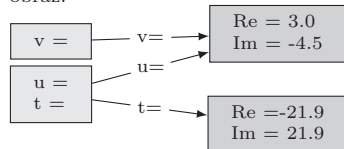
Na nasze potrzeby możemy przyjąć uproszczony obraz obiektów.

Obiekty są prostokątami (pojemnikami, pudełkami, ...). Zmienne występują tylko w obiektach. Wartością zmiennej jest albo liczba, albo obiekt.



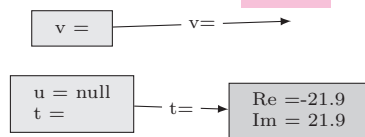
Zmienne t i u wskazują na dwa obiekty.

Po wykonaniu polecenia `v ← u` mamy taki obraz.



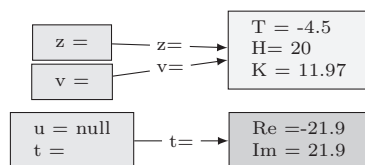
Zmienne u i v wskazują na ten sam obiekt. Wartości wyrażeń `u.Im` oraz `v.Im` są równe. Wykonanie polecenia `v.Re ← 50` spowoduje, że od tej pory wartość wyrażenia `u.Re` będzie równa 50. W takim przypadku mówi się o *aliasingu* zmiennych u i v.

Po wykonaniu polecenia `delete(u)`.



Od tej pory obiekt wskazywany przez zmienną u nie istnieje (`u=null`). A zmiennej v nie odpowiada żaden obiekt! Próba wyznaczenia wartości `u.Re` lub `u.Im` spowoduje zgłoszenie błędu. Wartości wyrażeń `v.Re` i `v.Im` są fałszywe. Obliczenie jest kontynuowane, bez ostrzeżenia o błędzie!! Na tym polega błąd **DANGLING REFERENCE**. Groźne jest to, że program nie sygnalizuje błędu. Programista może stracić miesiące na zrozumienie, co zaszło, i wykrycie takiego błędu. Koszty?

Co się dalej może zdarzyć? Sprzeczna interpretacja danych, gdy na zwolnionym miejscu pojawi się inny obiekt.



Widac to wyraźnie z rysunku: wartość `v.Re` jest nie tylko fałszywa numerycznie, typy usuniętego obiektu i obiektu z są różne.

W tej pracy przedstawimy groźne zjawisko – błąd wiszących referencji – jakie występuje w programowaniu obiektowym, np. w C++, Pascalu, C. W kolejnym artykule omówimy rozwiązanie pozwalające wyeliminować ten błąd. Zaczniemy od krótkiej ekspozycji problemów, jakie napotykamy podczas zarządzania pamięcią obiektów w każdym języku programowania obiektowego.

Obiekty są 1° tworzone, 2° współdzielone, 3° wykorzystywane i 4° ewentualnie stają się niepotrzebne.

Będziemy tutaj abstrahować od wielu szczegółów, takich jak rozmiar obiektu, jego typ, sposoby zarządzania odzyskiwaną pamięcią, itp.

Obiekty tworzymy, wykonując polecenie `x ← new(...)`, które tworzy nowy obiekt *o* i przypisuje go jako wartość zmiennej *x*. Np. `u ← new(Re ← 3.0, Im ← -4.5)`.

Obiekt *o* utworzony przez polecenie `new` może stać się wartością kilku zmiennych. Np. w ten sposób: `y ← x; ...; z ← y`. Ma wtedy miejsce współdzielenie (ang. *aliasing*, *sharing*) obiektu *o* przez zmienne *x, y, z*.

Współpracę z obiektem *o*, który jest wartością zmiennej *x*, można sprowadzić do trzech przypadków:

- *inspekcja* – ma miejsce wtedy, gdy odczytujemy wartość zmiennej zapisanej w obiekcie, np. trzeba wyznaczyć wartość wyrażenia `x.attr`.
- *uaktualnienie* – gdy przypisujemy nową wartość zmiennej w obiekcie *x*, np. `x.at ← 17`.
- *serwis* – gdy wykonujemy polecenie wykonania usługi zdefiniowanej w obiekcie, np. `call x.NarysujOkrąg(p,12)`.

Wszystkie trzy przypadki powinny rozpoczynać się od sprawdzenia, czy wartością zmiennej *x* jest jakiś żywy obiekt, czy też specjalna wartość `null`.

Na koniec sytuacja, gdy program zaniecha współpracy z obiektem *o* i gdy przestaje on być wartością jakiegokolwiek zmiennej. Taki obiekt nazywany jest *śmieciem*. Mądry programista pozbędzie się śmieci, uwalniając pamięć zajęta przez te obiekty.

Jakie *problemy* stwarza zarządzanie pamięcią obiektów? Dwa główne zagrożenia dla obliczeń naszego programu to:

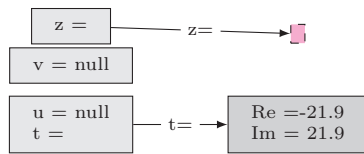
- wiszące referencje i
- zaśmiecanie pamięci.

Najwięcej szkód wyrządza błąd wiszących referencji (ang. *dangling reference error*). Można go nazwać *cichym zabójcą*, ponieważ program, w którym usadowił się taki błąd, może wydawać się bezpieczny przez długie miesiące i lata, zanim użytkownicy zorientują się, że program działa nieprawidłowo. Koszty nieprawidłowego działania mogą okazać się bardzo wysokie. (Zdarzało się, że nieprawidłowe działanie programu powodowało nawet śmierć ludzi!) Koszty wykrycia błędu i zlokalizowanie miejsca w programie, w którym ten błąd wystąpił, też nie są bagatelne. Nie umiem podać rzetelnych statystyk. Na pewno jednak tylko finansowe straty w ostatnich trzydziestu latach to kwoty rzędu setek milionów euro. Czy można skonstruować taki program, który analizowałby programy i wykrywał miejsce wystąpienia błędu wiszących referencji w programie?

Niestety, taki algorytm nie istnieje. Wytlumaczenie jest proste: gdyby istniał taki algorytm *A*, to moglibyśmy skonstruować inny algorytm *H*, który wykrywałby, czy dowolny dany tekst programu zakończy obliczenia w czasie skończonym, czy też obliczenie będzie przedłużane bez ograniczenia (czyli program zapętlił się). Jednakże – jak wiadomo – ten problem jest nierozstrzygalny.

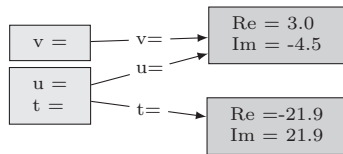
Mając to na względzie, twórcy języka programowania Java postanowili: *ponieważ błąd wiszących referencji jest tak groźny i ponieważ nie jest możliwe skonstruowanie kompilatora, który by sygnalizował wystąpienie tego błędu przed uruchomieniem programu, to w języku Java instrukcji delete(u) nie będzie*. Jak postanowili, tak uczynili.

Inny błąd – *opóźnionej destrukcji* występuje, gdy zostanie wykonane polecenie delete(v).

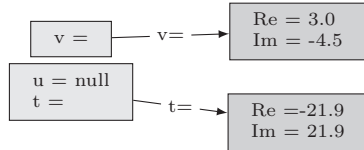


W efekcie usunięto poprawny i potrzebny obiekt z, zamiast wcześniej usuniętego obiektu u.

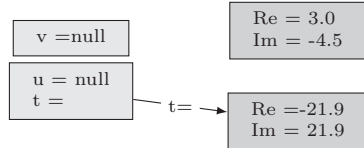
Jak powstają obiekty śmieci?



Po wykonaniu polecenia u ← null



Po wykonaniu v ← null



Na tym obrazku pojawił się **śmieć!** Wskaż go!

Ale czyniąc tak, naśladową pewnego króla Albanii.

Rzecz działa się dawno, ponad sto lat temu. Pomiędzy stolicą Tiraną i portem Dürres istniała jedyna w kraju linia kolejowa, długa na 40 kilometrów i opadająca stromo w dół. Otóż jednego dnia król wydał takie rozporządzenie: *Ponieważ na linii kolejowej często zdarzają się katastrofy i ponieważ w katastrofach tych wykolejeniu ulega ostatni wagon, przeto zarządzam, by od dziś do pociągu nie doczepiać ostatniego wagonu*.

W przypadku Javy okazało się, że na programistę uwolnionego od potwora Charybdy czeka inny potwór Scylla (pamiętasz, Czytelniku, jak Odyseusz żeglował pomiędzy tymi potworami?). Okazało się, że programy w Javie uwolnione od zagrożenia błędem wiszących referencji wpadły w pułapkę *zaśmiecania pamięci* (ang. *memory leakage*).

Twórcy Javy zaoferowali programistom narzędzie znane od dawna: *odśmiecać* (ang. *garbage collector*). Algorytm ten uruchomiony podczas wykonywania programu potrafi zlokalizować wszystkie istniejące obiekty-śmieci i je usunąć. Odzyskaną pamięć można przeznaczyć na nowe obiekty. Koszt odśmiecania nie jest bagatelny, ale pozbywamy się zagrożenia wiszącymi referencjami (twierdzili) i dajemy Ci narzędzie pozwalające kontrolować koszt pamięciowy Twojego programu. Ty określisz, co jest śmieciem, a my pozbędziemy się wszystkich śmieci.

Zapanowała euforia, ale trwała krótko. Niebawem okazało się, że mnożą się sytuacje podobne do trzeciego rysunku po lewej stronie. Programista wykonał polecenie u ← null, ale zapomniał, że obiekt o jest nadal wskazywany przez zmienną v. Odśmiecać nie ma podstaw, by go usunąć! Zresztą, nie zawsze błąd taki leży po stronie programisty. Programiści wykorzystują klasy napisane przez kogoś innego, a tam może ukrywać się polecenie podobne do v ← u z naszego przykładu.

Podsumujmy sytuację:

Charybda	Wiszące referencje stanowią realne zagrożenie.	Nie istnieje algorytm wykrywania takiego błędu w kodzie źródłowym programu.
Scylla	Zjawisko <i>zaśmiecania pamięci</i> zagraża zmniejszeniem szybkości wykonywania obliczeń, a nawet zablokowaniem pracy programu, gdy zabraknie miejsca na kolejny nowy obiekt.	Odśmiecać jest bezsilny, gdy w trakcie wykonywania programu obiekt już niepotrzebny nadal jest wartością jakiejś (zapomnianej) zmiennej.

Co pozostaje? Programiści Javy mogą korzystać z efemerycznych *słabych referencji*. Pojęcie to pojawiło się trzy lata później niż Java. Nie zyskało uznania i w najnowszej wersji Javy pozostał po nim tylko ślad. Programistom w C++ oferowane są różne narzędzia do monitorowania pracy programu (np. *debugging*).

A przecież ponad 35 lat temu Antoni Kreczmar (1945–1996), profesor informatyki Uniwersytetu Warszawskiego, wynalazł *bezpieczny i tani* (w eksploatacji) system przekształcania błędów wiszących referencji w ostrzeżenia o ich obecności. Przypomnijmy sytuację na obrazku z poprzedniej strony, trzecim od góry. Błędem (groźnym!) jest korzystanie ze zmiennej v, ponieważ zmienna ta wskazuje na nieistniejący już obiekt. Natomiast próba

Zobacz str. 374 w Java Language Specification v. 8

wykorzystania zmiennej u zakończy się ostrzeżeniem: *referencja wskazuje na null*. Ostrzeżony programista może łatwo odnaleźć błąd i (z pewnym trudem) go naprawić.

Zadanie, jakie postawił sobie Kreczmar:

1) Zapewnić prawdziwość wszystkim formułom postaci

$$\underbrace{((f_1 = \dots = f_k) \wedge f_1 \neq \mathbf{none})}_{\text{warunek początkowy}} \Rightarrow \underbrace{[\text{kill}(f_i)]}_{\text{instrukcja}} \underbrace{(f_1 = \dots = f_k = \mathbf{none})}_{\text{warunek końcowy}}$$

Formułę tę należy czytać tak: *jeśli na pewien obiekt o wskazuje k zmiennych f_1, \dots, f_k , to po wykonaniu instrukcji $\text{kill}(f_i)$ wszystkie zmienne przyjmują wartość **none**, a miejsce zajmowane przez obiekt zostaje zwrócone do puli wolnych miejsc.*

- 2) Wykonanie instrukcji kill ma zająć tyle samo czasu, niezależnie od liczby k zmiennych współdzielących obiekt.
- 3) Koszt sprawdzenia, czy obiekt jest żywy, ma być stały $O(1)$ i niewielki. Ma to znaczenie, ponieważ sprawdzenie takie dokonuje się przy każdym dostępie do obiektu.

Zajrzyj do hasła Antoni Kreczmar w Wikipedii.

Kreczmar nie tylko rozwiązał ten problem, ale także był bardzo ważnym współtwórcą języka programowania Loglan'82 i zaprogramował maszynę wirtualną tego języka.

Parę słów o projekcie badawczym Loglan

Prace nad językiem zostały podjęte w roku 1977 w Zakładzie Teorii Obliczeń Instytutu Informatyki UW. Rok później Zjednoczenie MERA, producent minikomputerów Mera 400, zawarło z nami umowę na stworzenie języka programowania obiektowego i współbieżnego Loglan oraz na zaprogramowanie kompilatora tego języka na komputery Mera 400. Otrzymaliśmy niezbyt dużą kwotę pieniędzy (rzędu 40 milionów ówczesnych złotych, z czego większą część pochłonął zakup 2 komputerów Mera 400), i co najważniejsze, dwa komputery do dyspozycji Instytutu. Prace zakończyły się powodzeniem: opublikowaliśmy raport zawierający definicję języka Loglan, przekazaliśmy kompilator Loglanu dla komputerów Mera 400, zorganizowaliśmy dwie konferencje międzynarodowe (1983, 1984), szkołę PTI nt. Loglanu w 1985, nawiązaliśmy owocną współpracę z Uniwersytetami w Kilonii i w Rzymie, ...

Wiele informacji o języku Loglan, jego kompilatorach dla Linuxa i Windows (działają do dziś!), o problemach badawczych sformułowanych i rozwiązanych w związku z Loglanem znajdziesz w repozytorium:

lem12.uksw.edu.pl

Pragnę podkreślić, że:

- W projekcie uczestniczyło kilkanaście osób z mojego Zakładu, wiele z nich to dziś profesorowie uczelni w Polsce, Niemczech, Kanadzie, USA, Meksyku, Szwecji, ... Niestety, nie ma dziś z nami Antoniego Kreczmara.
- Bardzo cenny wkład wnieśli doktoranci i studenci. Student Bolek Ciesielski (1988) wymyślił i zrealizował nowy protokół współpracy pomiędzy obiektami procesów – *alien call*. Jego odkrycie wciąż czeka na uznanie szerszej publiczności. Doktorant Oskar Świda (1996) zrealizował koncepcję łączenia loglanowskich maszyn wirtualnych w rozproszony, sieciowy klastrowy loglanowski. Z braku miejsca nie wspomnę o wielu innych osobach.
- Lista problemów badawczych, które sformułowano i rozwiązano, by uzyskać odpowiednią jakość języka i jego kompilatora, zawiera dziewięć pozycji. Tu omówiliśmy problem zapobiegania błędowi wiszących referencji.
- Loglan świetnie nadaje się jako platforma nauczania programowania, ponieważ zawiera prawie wszystkie mechanizmy programowania obiektowego i rozproszonego.
- Loglan jest dobrym punktem wyjścia do prowadzenia badań nad kolejnymi problemami, np. jak zarządzać obliczeniami w komputerze wieloprocesorowym. Żaden z obecnie stosowanych języków programowania nie ma odpowiednich do tego narzędzi.

Jak to rozwiązać?

Czy potrafisz odgadnąć, jak działa system zarządzania pamięcią obiektów zbudowany przez Antoniego Kreczmara? Pamiętasz wymagania? Spróbuj swych sił! Jeżeli wymyślisz **coś lepszego**, to – Autor obiecuje – otrzymasz nagrodę 50 EUR.



Po co komu dowód? (albo słów kilka o sztucznej inteligencji)

Drogi Czytelniku Majsterkowiczu! W tym odcinku Migawki wreszcie będzie ciekawie. Opiszemy, jak samodzielnie skonstruować prawdziwy zegarek mechaniczny!

Postawmy jednak sprawy bardzo uczciwie: nie mamy zielonego pojęcia, jak taki zegarek dokładnie działa. Nasza wiedza sprowadza się do bardzo prostych faktów: taki zegarek składa się z: kilku (a może kilkunastu?) zębatek o pewnej (zapewne nie za dużej) liczbie zębów; pewnej liczby metalowych rurek (o różnych średnicach i długościach); sprężyny zwiniętej spiralnie (także opisanej kilkoma parametrami); wskazówek, pokręteł, obudowy itp. itd.

Generalnie: cel w konfrontacji z naszą wiedzą wydaje się zupełnie nieosiągalny. Spróbujmy więc do sprawy podejść metodą tak zwaną *brutalnej siły*. To znaczy: spróbujmy (po kolei!) skonstruować wszystkie możliwe maszyny złożone z części opisanych wyżej. Już samych elementów (uwzględniając parametry) jest bardzo dużo, a do tego jeszcze możemy bardzo różnie ustawiać je wzajemnie w przestrzeni. Innymi słowy: ilość maszyn-kandydatek na działający zegarek jest astronomiczna. Większość z nich w ogóle nic nie robi, ALE przynajmniej jedna z nich będzie działała zgodnie z naszymi oczekiwaniami.

Nasza metoda jest więc następująca: buduj maszyny po kolei; każdą *testuj*, sprawdzając, czy aby nie *zbudował się* nam prawdziwy zegarek; jeśli test wypadnie pozytywnie – przerwij dalsze poszukiwania.

Nie da się zaprzeczyć, że powyższy algorytm opisuje poprawną metodę konstrukcji zegarka mechanicznego. Ma ona, oczywiście, ewidentne wady. Pierwsza jest taka, że metoda działa niesłychanie wręcz wolno. Druga, może nieco bardziej subtelna, bierze się z faktu, że maszynę uznajemy za poprawną, gdy ją (jak to określiliśmy) przetestowaliśmy. Można by długo dyskutować, co to właściwie znaczy i kiedy jakąś maszynę wolno nam zweryfikować jako Prawdziwy Zegarek Mechaniczny. Niemniej jednak – postawiony cel osiągnęliśmy!

Powyżej opisany sposób postępowania nie jest li tylko małym udanym żartem. Daje on pewne intuicje o tym, na jakiej zasadzie działa tak zwana sztuczna inteligencja, a bardziej konkretnie – uczenie maszynowe oparte o sieci neuronowe.

Opiszmy jakiś przykład. Proponuję rozważyć problem rozpoznawania cyfr napisanych odręcznie przez człowieka. Chcielibyśmy mieć algorytm, który potrafi to zadanie rozwiązywać automatycznie. Innymi słowy, chcemy stworzyć program komputerowy, który na wejściu dostanie zdjęcie kartki papieru, a na wyjściu określi, jaka cyfra została sfotografowana (upraszczamy problem do rozpoznawania jednego znaku).

Aby taki algorytm stworzyć, postąpimy analogicznie jak przy konstrukcji zegarka mechanicznego. To znaczy: skonstatujemy na początku, że nie mamy konkretnego planu, jak by się do tego zadania zabrać wprost. Jakąś tam wiedzę ogólną jednak mamy. Wiemy, że pewna bardzo skomplikowana struktura – mózg ludzki – radzi sobie świetnie z takim zadaniem. Niestety, nie umiemy zaimplementować w komputerze ludzkiego mózgu. Z drugiej strony, ludzki mózg wykonuje zadania znacznie bardziej skomplikowane niż rozpoznawanie cyferek, więc pewnie jakaś o wiele prostsza struktura, ale jednak do mózgu choć trochę podobna, ma szansę zadziałać.

Zakładamy więc, że nasz algorytm będzie działał na obraz i podobieństwo BARDZO uproszczonego mózgu.

Zacniemy od definicji klocków konstrukcyjnych. Będą nimi *sztuczne neurony*, które będziemy mogli dowolnie łączyć, tworząc sieć. Pojedynczy sztuczny neuron działa bardzo prosto. Na przykład w modelu McCullocha–Pittsa zakładamy, że może on mieć wiele wejść (ale jedno wyjście), którym przekazywane są jakieś liczby, a wyjście jest ich jakąś (bardzo prostą) funkcją. Nasz docelowy algorytm będzie zawierał opis sieci takich prostych neuronów wraz z opisem ich parametrów (np. opisem ich wewnętrznej funkcji). Samo obliczenie będzie polegać na tym, że wejściowe zdjęcie przerobimy na ciąg bitów, które będą z kolei wejściem do pewnego (ustalonego) zbioru neuronów. Następnie wszystkie neurony wewnętrzne dokonają obliczeń zgodnych ze swoim opisem i zastanym wejściem, aż neuron wyjściowy poda wynik. Oczywiście, cała sztuka polega na znalezieniu właściwej sieci i właściwego zbioru parametrów.

Sam algorytm poszukiwania właściwego algorytmu (sic!) mógłby działać następująco: buduj po kolei różne algorytmy (kandydatów); każdy kolejny testuj, aż uda się znaleźć jakiś działający poprawnie.

(W rzeczywistości działamy nieco sprytniej: wcale nie przeszukujemy kandydatów zupełnie losowo po kolei. Parametry kolejnych rozpatrywanych algorytmów są zależne od wcześniejszych testów. Np. przyjmujemy (heurystycznie), że gdy jakiś kandydat działa *już całkiem nieźle*, to staramy się dalej szukać w jego pobliżu, a więc nieznacznie modyfikując jego parametry.)

Dodatkowo musimy wyjaśnić, czym jest w tym przypadku testowanie. Otóż, aby to wszystko miało sens, musimy mieć (najlepiej bardzo dużą) bazę danych zdjęć cyfr napisanych odręcznie przez ludzi wraz z poprawną informacją, o jakie cyfry chodzi. Wówczas testowanie to po prostu weryfikacja, jak nasz algorytm poradziłby sobie z daną bazą danych.

To wszystko wyżej brzmi pewnie trochę surrealistycznie. A jednak – okazuje się, że gdy rozważane sieci są odpowiednio duże, nasze komputery odpowiednio szybkie, nasze heurystyki w wyborze kolejnych kandydatów trafione, a nasze bazy danych odpowiednio duże – to po prostu działa! W rozsądnym czasie potrafimy znaleźć algorytm, który radzi sobie nieźle z postawionym problemem. Więcej, opisana wyżej metoda działa zaskakująco dobrze z jeszcze trudniejszymi zadaniami, takimi jak rozpoznawanie przeszkód z obrazu kamerki samochodowej czy dobieranie właściwej reklamy do użytkownika serwisu internetowego.

To, co jednak powinno przerażać (szczególnie matematyka), to fakt, iż informatyk, który *znalazł* jakiś algorytm w stogu neuronowego siana, ani nie wie *tak naprawdę*, jak on działa, ani tym bardziej nie potrafi nic o nim udowodnić...

Tomasz KAZANA

PS. Drugi rozdział swojej książki „O Matematyku, rycerzu Gwiazdy Pitagorejskiej” nasz redaktor naczelny, Marek Kordos, kończy tak: *I z pewnym niepokojem należy uświadomić sobie, że dzisiejsza, prawie już dyrygująca światem informatyka też odstępuje od dowodzenia swoich algorytmów, zadowolając się przetestowaniem powstających programów.* Po przeczytaniu tych słów miałem, oczywiście, mieszanie uczucia. Że nie do końca, że nie zawsze. Ale w sztucznej inteligencji przyznaję: do końca i zawsze.

Informatyczny kącik olimpijski (105): Odtwarzanie drzewa

W tym miesiącu omówimy zadanie *New Year and Forgotten Tree*, które pojawiło się na konkursie Good Bye 2015 na platformie Codeforces.

Standardowym opisem nieskierowanego drzewa o n wierzchołkach (ponumerowanych od 1 do n) jest podanie $n - 1$ par liczb reprezentujących jego krawędzie. Na wejściu analizowanego zadania dany jest taki właśnie opis, jednakże każda cyfra została zastąpiona znakiem '?'. Na przykład, krawędź między wierzchołkami 90 i 123 byłaby reprezentowana jako „?? ???”. Naszym zadaniem jest odtworzyć jakiegokolwiek drzewo pasujące do danego opisu bądź stwierdzić, że takie drzewo nie istnieje.

Niech k oznacza największą możliwą liczbę cyfr etykiety wierzchołka (czyli $k \approx \log_{10} n$). Wierzchołki o tej samej liczbie cyfr są, oczywiście, reprezentowane przez ten sam napis, więc w naturalny sposób podzielimy wierzchołki na k grup, zgodnie z liczbą cyfr etykiety wierzchołka. Dla każdej pary takich grup znamy żadaną liczbę krawędzi między wierzchołkami tych grup.

Okazuje się, że próba prostego zachłannego odtwarzania naszego drzewa nie rozwiązuje zadania. Również (być może narzucające się) użycie twierdzenia Halla (dla 2^k zbiorów grup) też nie jest poprawne. Dobrym pomysłem jest natomiast konstrukcja, w której utrzymujemy cały czas jedną spójną składową i dokładamy do niej nowe krawędzie i wierzchołki.

Spróbujemy w każdej grupie wyróżnić jeden wierzchołek w sposób prowadzący do pewnych dwóch własności. Po pierwsze, chcemy, by krawędzie między wyróżnionymi wierzchołkami łączyły je w jedną spójną składową. Równoważnie, między k wyróżnionymi wierzchołkami potrzebujemy $k - 1$ krawędzi.

Rozważmy dowolne rozwiązanie (drzewo pasujące do danego opisu). Wybierzmy jedną dowolną krawędź między dwoma wierzchołkami z różnych grup i wyróżnijmy te dwa wierzchołki. Dopóki wyróżnionych jest mniej niż k wierzchołków, powtarzajmy następujące kroki. Grupę nazwijmy nową, jeśli nie zawiera jeszcze żadnego wyróżnionego wierzchołka. Wierzchołek nazwijmy osiągalnym, jeśli da się do niego dojść z jednego z już wyróżnionych wierzchołków bez przechodzenia przez wierzchołki z nowych grup. Z takich definicji wynika, że istnieje krawędź między osiągalnym wierzchołkiem i wierzchołkiem z nowej grupy. Te dwa wierzchołki oznaczmy jako u i v . Jeśli u jest wyróżnione, to po prostu wyróżniamy też v . W przeciwnym przypadku zastąpmy najpierw krawędź między u i v krawędzią między v i wierzchołkiem wyróżnionym w grupie z u (nazwijmy go u'). Powstały graf wciąż pasuje do tego samego opisu oraz wciąż jest drzewem, bo po usunięciu krawędzi uv wierzchołki u i u' są w jednej składowej (bo u było osiągalne). Teraz możemy wyróżnić v i wciąż spełniony jest żądany warunek, że krawędzie między wyróżnionymi wierzchołkami łączą je w jedną składową.

Po drugie, niech każda krawędź w drzewie dotyka co najmniej jednego wyróżnionego wierzchołka. Jeśli tak nie jest, to spójrzmy na dowolną krawędź uv między dwoma niewyróżnionymi wierzchołkami z różnych grup. Jako u' i v' oznaczmy wyróżnione wierzchołki z grup zawierających odpowiednio u i v . Usuńmy krawędź uv , co dzieli drzewo na dwie składowe. Wiemy, że wyróżnione wierzchołki są ze sobą połączone w jedną składową, zatem u' i v' są wciąż w jednej składowej. Po usunięciu krawędzi uv wierzchołki u i v muszą

być w dwóch różnych składowych, zatem jeden z nich musi być w tej samej składowej co u' i v' .

Bez straty ogólności powiedzmy, że wierzchołek u jest w tej samej składowej co u' i v' . Zatem wierzchołek v jest w innej składowej i możemy go połączyć nową krawędzią z u' . Nowe drzewo wciąż pasuje do danego opisu. Podobne rozumowanie można zastosować dla krawędzi między dwoma niewyróżnionymi wierzchołkami z tej samej grupy, co pozostawiamy Czytelnikowi jako proste ćwiczenie. Skoro możemy pozbywać się krawędzi między niewyróżnionymi wierzchołkami, to musi istnieć rozwiązanie, które takich krawędzi nie ma wcale.

Pokazaliśmy, że jeśli istnieje jakiegokolwiek rozwiązanie, to istnieje też takie, w którym da się wyróżnić po jednym wierzchołku w każdej z k grup w taki sposób, że pewien podgraf drzewa jest drzewem rozpinającym k wyróżnionych wierzchołków oraz że nie ma krawędzi między niewyróżnionymi wierzchołkami. Możemy bez straty ogólności powiedzieć, że wyróżniamy wierzchołki o numerach 10^i . Skoro k jest bardzo małe, to po prostu rozważmy kolejno wszystkie k^{k-2} drzewa rozpinające o tylu wierzchołkach. W zadaniu tym można to zrobić brutalnie na różne sposoby, natomiast zainteresowanym Czytelnikom polecamy zapoznanie się z kodami Prüfera, które opisują wydajną metodę iteracji po wszystkich etykietowanych drzewach o ustalonej liczbie wierzchołków.

Powiedzmy, że ustaliliśmy $k - 1$ krawędzi między k wyróżnionymi wierzchołkami. Pozostałe $n - 1 - (k - 1) = n - k$ krawędzi nazwijmy zwykłymi. Zwykła krawędź dotyka dokładnie jednego wyróżnionego i jednego niewyróżnionego wierzchołka. Każda taka krawędź musi dotykać innego niewyróżnionego wierzchołka, bo takich wierzchołków jest $n - k$, czyli tyle, ile jest zwykłych krawędzi. Kolejność niewyróżnionych wierzchołków w grupie nie ma znaczenia, bo każdy z nich jest reprezentowany przez ten sam napis na wejściu. Pozostaje zatem dla każdej takiej krawędzi zdecydować, w której z dwóch grup ma ona dotykać wyróżnionego wierzchołka, a w której ma dotykać nowego niewyróżnionego wierzchołka. Pamiętajmy, że w każdej grupie znamy docelową liczbę niewyróżnionych wierzchołków.

Tę ostatnią część zadania możemy rozwiązać za pomocą algorytmu maksymalnego przepływu. Tworzymy k wierzchołków g_1, \dots, g_k reprezentujących grupy. Prowadzimy z nich do ujścia krawędzie o przepustowościach równych liczbom niewyróżnionych wierzchołków w odpowiednich grupach. Tworzymy też $\frac{k \cdot (k+1)}{2}$ wierzchołków $e_{i,j}$, do których ze źródła prowadzimy krawędzie o przepustowościach równych liczbom zwykłych krawędzi między grupami i i j . Z $e_{i,j}$ prowadzimy krawędzie do g_i i do g_j o nieskończonych przepustowościach. Przepływ o wartości x z $e_{i,j}$ do g_i należy interpretować jako decyzję, by x krawędzi między grupami i i j dotykało niewyróżnionych wierzchołków w i .

Rozważyć musimy wszystkie k^{k-2} drzew rozpinających na k wierzchołkach i dla każdego znaleźć przepływ w grafie o $O(k^2)$ wierzchołkach i krawędziach. Przykładowa złożoność takiego programu to $O(k^{k-2} \cdot k^6 + n) = O(k^{k+4} + n)$, gdzie obecność składnika n wynika z konieczności właściwego odtworzenia i wypisania drzewa o n wierzchołkach.

Zainteresowanym proponujemy zastanowić się, czy możemy zamiast metody maksymalnego przepływu użyć twierdzenia Halla.

Kamil DĘBOWSKI

Bizony, tury, żubry

O żubrach mówi się częściej, gdy dzieje się cokolwiek wokół Puszczy Białowieskiej. Dziś – w obliczu planowanej jej wycinki i rozmów o odstrzale żubrów (podobno chorujących na gruźlicę). Gdy chce się dowiedzieć więcej o historii genetycznej żubra, sprawa komplikuje się, również z powodu polskiego nazewnictwa (żubr, bizon, tur), które bez łacińskich nazw gatunkowych łatwo wprowadza w zakłopotanie (w dodatku, gdy się nie jest specjalistką, jak ja). Pomaga częściowo Wikipedia.

Żubr europejski, (*Bison bonasus*) – gatunek łożyskowca z rodziny wołowatych, rzędu parzystokopytnych. W 2013 roku światowa liczebność gatunku wynosiła 5249 osobników, z czego 3626 żyło w wolnych i w półwolnych populacjach. Według danych z 2013 roku w Polsce żyło 1138 osobników w stadach wolnościowych, ponad połowa w Białowieży. Ten gatunek żubra jest jedynym ze swojego rodzaju, który ostał się w Europie w holocenie (epoka współczesna, liczona od 12 tysięcy lat).

Parę lat temu naukowcy polscy proponowali projekt odtworzenia różnych wymarłych gatunków – jednym z nich był polski tur, ostatni odstrzelony w XVII wieku. Wydaje się, że wraz z zanikiem zainteresowania klonowaniem ssaków projekt ten także się zakończył.

Na granicy plejstocenu i holocenu, jak mówią o tym badania paleogenetyczne, doszło do skrzyżowania pierwotnego bizona z turem. Stwierdzenie, że hybrydyzacja ssaków może doprowadzić do powstania nowego gatunku, było ważnym odkryciem genetycznym ostatnich lat (2016/17). Sformułowano tę tezę dzięki badaniom mitochondrialnego DNA i ocalałych sekwencji jądrowego DNA 64 bizonów późnego plejstocenu z trzech rejonów Europy: Uralu, Kaukazu i wybranych miejsc zachodniej Europy (tu dane z Polski), które porównano z danymi dla bizonów Ameryki Północnej. Bardzo interesującym komentarzem do tych współczesnych i nowoczesnych badań była także analiza naskalnych rysunków bizonów wykonanych ponad 15 tys. lat temu na ścianach jaskiń południowej Europy. Zarówno ogólny zarys ciała, jak i rodzaj rogów, dobitnie świadczą o istnieniu różnych gatunków bizonów na tych terenach.

Paleogenetycy zainteresowali się także historią opanowania Ameryki Północnej przez bizony. Uważa się, że pojawiły się tam w dwu falach imigracyjnych z Syberii, ponad 130 tysięcy lat temu i 45–21 tysięcy lat temu. W publikacjach paleontologów amerykańskich z 2017 roku – po pobraniu i analizie DNA 42 bizonów (dwudziestu sześciu północnoamerykańskich sprzed 400 i 45 000 lat, dziesięciu współczesnych i sześciu syberyjskich) – ustalono, że po przybyciu na te tereny bizony szybko rozprzestrzeniły się. Porównywano DNA mitochondrialny kości pobierany z lokalizacji w kanadyjskiej prowincji Yukon, i (znaleziony wcześniej) w Colorado, w tej samej warstwie wulkanicznej, przypisywanej okresowi sprzed 124 tysięcy lat. Dokładna analiza DNA sugeruje, że bizony Yukonu to bizony stepowe, *B. priscus*, a te z Colorado to giganty o wielkich rogach, (*B. latiphrons*). Mając te dane, można było zaproponować drzewo genealogiczne. DNA bizonów z Yukonu i z Colorado lokują je blisko podstawy drzewa genealogicznego, co sugeruje, że były bliskimi potomkami pierwszych kolonizatorów z Syberii. Ich wspólny żeński przodek datowany jest na okres 195–135 tysięcy lat temu. Potomkowie, szybko zmieniając środowisko życia, ulegli też szybkiemu różnicowaniu na odrębne gatunki.

Współczesne badania wskazują na istnienie dziś wymarłych czterech gatunków, z których trzy zamieszkiwały Amerykę Północną, czwarty spotykano właściwie na całej północnej półkuli (są o tym świadczące wykopaliska z terenów dzisiejszej Japonii). Jak wiadomo, bizony w Ameryce Północnej wymarły z powodów nienaturalnych – zostały wybite przez białych kolonizatorów w celu odciążenia miejscowej ludności od źródła pokarmu.

Koło Matematyków Studentów Uniwersytetu Jagiellońskiego organizuje w dniach 17–23 września 2017 roku XX Międzynarodowe Warsztaty dla Młodych Matematyków „Teoria Liczb”. Wśród zaproszonych gości są:

Jörg Brüdern (Mathematisches Institut Georg-August Universität Göttingen),

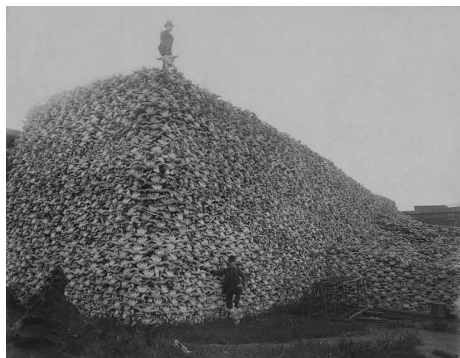
Sander Dahmen (Department of Mathematics, Vrije Universiteit Amsterdam),

Florian Luca (School of Mathematics, University of Witwatersrand, Johannesburg),

Bartłomiej Bzdęga (Uniwersytet Adama Mickiewicza, Poznań),

Bartosz Naskręcki (School of Mathematics, University of Bristol).

Zapraszamy do rejestrowania się na stronie:
kmsuj.im.uj.edu.pl/warsztaty/



Czaszki amerykańskich bizonów przed zmieleniem ich na nawóz (fotografia z 1870 roku). Bez komentarza.

Magdalena FIKUS

Bardzo szybki dostęp, i to na zimno, prosto z Białegostoku

Poszukiwanie nowych sposobów przechowywania informacji jest nie tylko ciekawe poznawczo, ale może mieć konkretne zastosowanie praktyczne. Z tego ostatniego punktu widzenia pożądane są rozwiązania o dużym stopniu niezawodności, szybkie, o dużym stopniu upakowania informacji oraz zużywające jak najmniej energii. W ostatnim kryterium chodzi nie tylko o oszczędność, lecz także o problemy z chłodzeniem.

Ostatnio ukazała się praca [1], w której zademonstrowano rozwiązanie spełniające trzy z tych czterech warunków, z realnymi możliwościami spełnienia również czwartego (upakowania). Wykorzystuje ono pojedynczy femtosekundowy impuls laserowy do zapisu lub odczytu bitu w przezroczystej warstwie ferromagnetycznego dielektryka. Zmiana polaryzacji liniowej impulsu pozwala przełączyć oś magnetyzacji, zapisując w ten sposób bit informacji powtarzalnie, trwale i odwracalnie. Próby wykorzystywania światła do modyfikacji magnetycznej orientacji ferromagnetyków trwają od wielu lat. Interakcja światła z metalem zawsze prowadzi jednak do silnego grzania. Metale są nieprzezroczyste, bo wolne elektrony tylko czekają, żeby każde światło pochłonąć. Kluczowe okazało się użycie zamiast metalu dielektryka, w którym można zmienić orientację za pomocą światła bez znaczącego podgrzewania próbki. Zwykle dielektryki nie wykazują jednak silnych własności magnetycznych. Trzeba je w nie wzbogacić.

Użyтым materiałem była cienka warstwa itrowo-żelazowego granatu domieszkowanego jonami kobaltu. Właśnie *jony kobaltu w tym granacie są odpowiedzialne za silne sprzężenie momentu magnetycznego elektronu z momentem pędu jego ruchu orbitalnego, tzw. sprzężenie spinowo-orbitalne. Światło liniowo spolaryzowane może skutecznie zmienić ruch orbitalny elektronów w jonach, tym samym zmieniając kierunek momentu magnetycznego. Zjawiskiem odpowiedzialnym za te zmiany jest tak zwany efekt fotomagnetyczny, obecność którego w warstwach granatów już w temperaturze pokojowej zademonstrowano [2] w Zakładzie Fizyki Magnetyków Uniwersytetu w Białymstoku.* [3]

Uzyskano metodę, która pozwala na niezwykle szybką manipulację orientacją magnetyczną. Przełączenie następuje w około 20 pikosekund. To trzy rzędy wielkości szybciej niż w najlepszych z obecnie stosowanych pamięci RAM. Natomiast straty energii związane z rozpraszaniem ciepła są nawet cztery rzędy wielkości mniejsze niż w obecnie stosowanych rozwiązaniach. Przezroczystość warstwy pozwala myśleć o istotnym zwiększeniu nierewelacyjnego na razie upakowania poprzez wykorzystanie wielowarstwowości. Droga do komercyjnego wykorzystania pomysłu jest jednak jeszcze dość długa i wyboista. Miejmy nadzieję, że istotny wkład białostockiego zespołu współpracującego z grupą z Holandii przełoży się na sukces technologiczny, a może i na komercyjny, kto wie.

Egzotyczne trajektorie przez trzy szczeliny

Słynny dwuszczelinowy eksperyment Younga, który przekonywał o falowej naturze światła, tak naprawdę dowodzi jego kwantowej natury. Interferencja nie znika, kiedy przepuszczamy przez szczeliny pojedyncze fotony, a nawet gdy zastąpimy je pojedynczymi obiektami tak skomplikowanymi jak fulereny.

W paradygmacie humanisty zadomowiło się jednak pojęcie dualizmu korpuskularno-falowego, pomimo tego że pochodzi ono z przedkwantowego słownika. A może właśnie dlatego?

Natomiast według obecnie najlepiej funkcjonującego modelu rzeczywistości interferencja bierze się z dodania amplitud odpowiadających wszystkim trajektoriom, nawet tym najbardziej egzotycznym. Im dziwniejsza droga, tym przyczynek od niej mniejszy. Dlatego w eksperymencie z dwiema szczelinami wystarczy uwzględnić amplitudy dróg przez każdą ze szczelin.

Podobnie jest w przypadku trzech szczelin. Wtedy jednak łatwo zauważyć pewne umiarkowanie egzotyczne trajektorie. Zamiast rozpatrywać dodatkową drogę, np. dookoła układu alfy Centauri (środek dydaktyczny Feynmana, propagatora całek po trajektoriach), wystarczy wyobrazić sobie, że foton np. może przejść przez pierwszą szczelinę, wrócić drugą, by ostatecznie przejść trzecią. Pamiętając ze szkoły, że światło (w próżni) rozchodzi się po liniach prostych, wydaje się, że możemy tę umiarkowanie egzotyczną marszrutę uznać za egzotyczną w stopniu wystarczającym, żeby jej wpływ zignorować.

Okazuje się, że zbyt pochopnie. Autorom pracy [4] udało się zaobserwować jej wpływ na rejestrowany obraz interferencyjny. Osiągnęli to dzięki wzmocnieniu efektu poprzez plazmoneczne wzbudzenia na powierzchni metalowej przesłony, co znacząco zwiększyło prawdopodobieństwo podążania fotonu egzotyczną trajektorią. Interferencję obserwowali wtedy, gdy oświetlana była tylko jedna szczelina, co pozornie jest sprzeczne z zasadą superpozycji kwantowej. Pozornie, bo po uwzględnieniu w obliczeniach egzotycznych trajektorii uzyskuje się przewidywanie zgadzające się z obserwowanym obrazem.

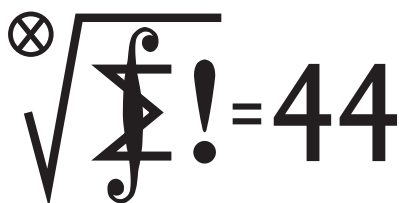
Jest to kolejny spektakularny przykład nieadekwatności niekwantowej wyobraźni. Wzięcie tego efektu pod uwagę może mieć niebagatelne znaczenie tak w badaniach testujących kwantową rzeczywistość, jak i ich praktycznych zastosowaniach.

Piotr ZALEWSKI

- [1] A. Stupakiewicz, K. Szerenos, D. Afanasiev, A. Kirilyuk oraz A. V. Kimel, *Ultrafast nonthermal photo-magnetic recording in a transparent medium*, *Nature* (2017), doi:10.1038/nature20807
- [2] A. Stupakiewicz, A. Maziewski, I. Davidenko oraz V. Zablotskii, *Light-induced magnetic anisotropy in Co-doped garnet films*, *Phys.Rev.B* **64**, 064405 (2001)
- [3] Andrzej Stupakiewicz, <http://1abfiz.uwb.edu.pl/and/?Badania>
- [4] O.S. Magaña-Loaiza, I. De Leon, *i inni*, *Ezotic looped trajectories of photons in three-slit interference*, *Nature Communications* **7**, 13987 (2016), doi:10.1038/ncomms13987

Skrót regulaminu

Każdy może nadsyłać rozwiązania zadań z numeru n w terminie do końca miesiąca $n + 2$. Szkice rozwiązań zamieszczamy w numerze $n + 4$. Można nadsyłać rozwiązania czterech, trzech, dwóch lub jednego zadania (każde na oddzielnej kartce), można to robić co miesiąc lub z dowolnymi przerwami. Rozwiązania zadań z matematyki i z fizyki należy przesyłać w oddzielnych kopertach, umieszczając na kopercie dopisek: **Klub 44 M** lub **Klub 44 F**. Oceniamy zadania w skali od 0 do 1 z dokładnością do 0,1. Ocenę mnożymy przez współczynnik trudności danego zadania: $WT = 4 - 3S/N$, gdzie S oznacza sumę ocen za rozwiązania tego zadania, a N – liczbę osób, które nadesłały rozwiązanie choćby jednego zadania z danego numeru w danej konkurencji (**M** lub **F**) – i tyle punktów otrzymuje nadsyłający. Po zgromadzeniu 44 punktów, w dowolnym czasie i w którejkolwiek z dwóch konkurencji (**M** lub **F**), zostaje on członkiem **Klubu 44**, a nadwyżka punktów jest zaliczana do ponownego udziału. Trzykrotne członkostwo – to tytuł **Weterana**. Szczegółowy regulamin został wydrukowany w numerze 2/2002 oraz znajduje się na stronie deltami.edu.pl



Termin nadsyłania rozwiązań: 31 VII 2017

Zadania z matematyki nr 741, 742

Redaguje Marcin E. KUCZMA

741. Niech W będzie wielościanem wypukłym, środkowo-symetrycznym, i niech π będzie ustaloną płaszczyzną, przechodzącą przez środek symetrii. Przekrój wielościanu W płaszczyzną π jest zawarty w kole o promieniu r . Udowodnić, że przekrój wielościanu W każdą płaszczyzną, równoległą do π , jest zawarty w pewnym kole o promieniu r – lub podać przykład, pokazujący nieprawdziwość takiego stwierdzenia.

742. Niech p będzie liczbą pierwszą postaci $p = 4k + 1$. Dowieść, że istnieje liczba całkowita dodatnia s , mniejsza od p , dla której różnica $sp - \lfloor \sqrt{sp} \rfloor^2$ jest kwadratem liczby całkowitej dodatniej.

Zadanie 742 zaproponował pan Tomasz Ordowski.

Rozwiązania zadań z numeru 1/2017

Przypominamy treść zadań:

733. Wierzchołek czworościanu nazwijmy *ciekawym*, jeśli z trzech wychodzących zeń krawędzi nie da się zbudować trójkąta.

- (a) Czy istnieje czworościan, którego wszystkie wierzchołki są ciekawe?
(b) Czy istnieje czworościan, mający dokładnie jeden ciekawy wierzchołek?

734. Udowodnić, że dla każdej liczby naturalnej $n > 1$ oraz dla dowolnych liczb rzeczywistych dodatnich p, q, a_1, \dots, a_n zachodzi nierówność

$$\sum_{k=1}^n a_k^{p+q} a_{k+1}^{-q} \geq \sum_{k=1}^n a_k^p \quad (\text{przyjmujemy } a_{n+1} = a_1).$$

Czołówka ligi zadaniowej **Klub 44 M** po uwzględnieniu ocen rozwiązań zadań 727 ($WT = 3,22$) i 728 ($WT = 1,60$) z numeru 10/2016

Marek Spychała	Warszawa	42,75
Witold Bednarek	Łódź	42,32
Zbigniew Skalik	Wrocław	41,22
Roksana Słowik	Knurów	38,41
Franciszek S. Sikorski	Warszawa	38,08
Adam Dzedzej	Gdańsk	35,68
Patryk Jaśniewski	Gdańsk	31,78

733. (a) Weźmy dowolny czworościan oraz jego najdłuższą krawędź. Przyjmijmy, że ma ona długość a , zaś dwie przyległe do niej ściany mają krawędzie długości a, b, c oraz a, d, e , przy czym krawędzie a, b, e mają wspólny koniec oraz krawędzie a, c, d mają wspólny koniec. Wówczas $a < b + c$ oraz $a < d + e$; stąd $2a < b + c + d + e$, wobec czego musi zachodzić co najmniej jedna z nierówności $a < b + e$ lub $a < c + d$. Zatem co najmniej jeden z końców krawędzi a nie jest wierzchołkiem ciekawym.

(b) Dokładnie jeden wierzchołek ciekawy jest możliwy. Rozpocznijmy od przykładu czworościanu zdegenerowanego do czwórki punktów na płaszczyźnie w konfiguracji: ABC – trójkąt prostokątny; D – środek przeciwprostokątnej AB ; $|BC| = a$, $|CA| = b$, $|AB| = |BD| = c$, przy czym $a < c < b < 2a$ (np. $a = 16$, $b = 30$, $c = 17$). Z punktu A wychodzą krawędzie długości $b, c, 2c$; z punktu C : a, b, c ; z punktu D : c, c, c ; każda z tych trójek spełnia warunek trójkąta. Pozostaje wierzchołek B – jedyny ciekawy (wspólny koniec krawędzi $a, c, 2c$). Teraz wystarczy wyjść w przestrzeń i nieznacznie przemieścić wierzchołek D , usuwając go

prostopadle z płaszczyzny ABC . Wychodzące zeń krawędzie trochę się wydłużą. Przy małym przemieszczeniu rozważane nierówności (ostre) pozostaną w mocy; punkt B nadal będzie jedynym wierzchołkiem ciekawym.

734. W nierówności Bernoulliego $(1 + t)^{\alpha+1} \geq 1 + (\alpha + 1)t$ (słusznej dla $t \geq -1$, $\alpha \geq 0$) wykonujemy „przesunięcie zmiennej” $1 + t = x$, uzyskując postać

$$x^{\alpha+1} \geq (\alpha + 1)x - \alpha \quad (\text{dla } x \geq 0, \alpha \geq 0).$$

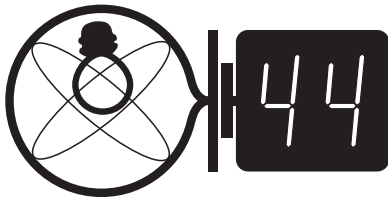
Podstawiając $\alpha = q/p$ oraz $x = (a_k/a_{k+1})^p$, dostajemy

$$\left(\frac{a_k}{a_{k+1}}\right)^{p+q} \geq \frac{p+q}{p} \left(\frac{a_k}{a_{k+1}}\right)^p - \frac{q}{p} \quad (\text{dla } k = 1, \dots, n).$$

Mnożymy stronami przez a_{k+1}^p , otrzymując

$$\frac{a_k^{p+q}}{a_{k+1}^q} \geq \left(1 + \frac{q}{p}\right) a_k^p - \frac{q}{p} a_{k+1}^p.$$

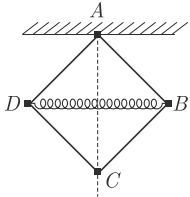
Suma tych nierówności (dla $k = 1, \dots, n$) to teza zadania.



Zadania z fizyki nr 638, 639

Redaguje Elżbieta ZAWISTOWSKA

Termin nadsyłania rozwiązań: 31 VII 2017



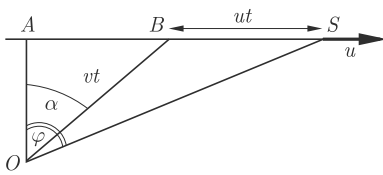
Rys. 1

638. Układ składa się z czterech jednakowych, lekkich prętów o długości l i lekkiej sprężyny o długości $2l$ (rys. 1). Pręty połączone są przegubowo za pomocą małych kulek o jednakowych masach. Układ zamocowany jest w punkcie A i znajduje się w polu ciężkości. W stanie równowagi pręty tworzą kwadrat. Znaleźć częstość małych drgań układu, przy których punkt C porusza się po linii pionowej.

639. Lis biegnie po linii prostej ze stałą prędkością v_1 . Lisa goni pies, którego prędkość ma stałą wartość v_2 i skierowana jest cały czas na lisa. W chwili, gdy prędkości v_1 i v_2 są do siebie prostopadłe, odległość między lisem a psem wynosi l . Jakie jest w tym momencie przyspieszenie psa?

Rozwiązania zadań z numeru 1/2017

Przypominamy treść zadań:



Rys. 2

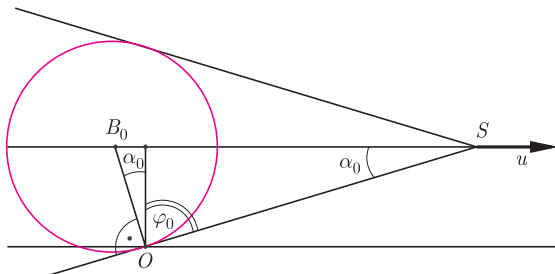
630. Samolot leci z prędkością u po prostej poziomej, przechodzącej nad głową obserwatora. W pewnej chwili obserwator widzi samolot w kierunku, który tworzy z pionem kąt φ . Jaki kąt z pionem tworzy w tej samej chwili kierunek, wzdłuż którego dociera do obserwatora dźwięk silnika samolotu? Prędkość dźwięku wynosi v . Rozważ przypadki $u < v$ oraz $u > v$.

631. Na powierzchni poziomej znajdują się dwa jednakowe, cienkościennie walce o masie m każdy. Osie walców są równoległe, promienie są równe R . Na początku jeden z walców spoczywa, a drugi toczy się bez poślizgu w kierunku pierwszego z prędkością ruchu postępowego v_0 aż do centralnego, sprężystego zderzenia. Współczynnik tarcia kinetycznego walców o podłoże jest równy μ , tarcie między walcami jest zaniedbywalne. Znaleźć maksymalną odległość między walcami po zderzeniu.

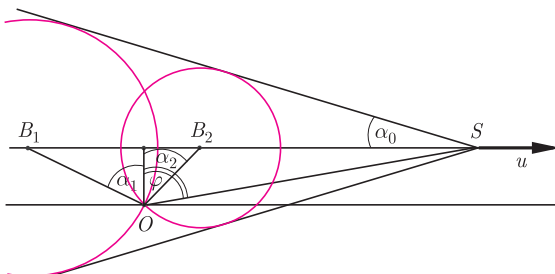
630. Niech obserwator, znajdujący się w punkcie O , widzi samolot w punkcie S i słyszy wysłany przez samolot z punktu B dźwięk, który dotarł do niego po czasie t . Zgodnie z rysunkiem 2

$$(1) \quad \operatorname{tg} \varphi = \operatorname{tg} \alpha + \frac{ut}{vt \cos \alpha},$$

gdzie α jest szukanym kątem. Wprowadźmy oznaczenie $x := \operatorname{tg} \varphi - \operatorname{tg} \alpha > 0$.



Rys. 3



Rys. 4

Zgodnie z (1) mamy

$$x = \frac{u}{v} \sqrt{1 + \operatorname{tg}^2 \alpha} = \frac{u}{v} \sqrt{1 + (\operatorname{tg} \varphi - x)^2}.$$

Otrzymujemy stąd równanie kwadratowe

$$(2) \quad x^2 + \frac{2 \operatorname{tg} \varphi}{\frac{v^2}{u^2} - 1} x - \frac{1}{\cos^2 \varphi (\frac{v^2}{u^2} - 1)} = 0.$$

Rozwiązania tego równania mają postać

$$x_{1,2} = \frac{\operatorname{tg} \varphi}{1 - \frac{v^2}{u^2}} \pm \sqrt{\frac{\operatorname{tg}^2 \varphi}{(1 - \frac{v^2}{u^2})^2} - \frac{1}{\cos^2 \varphi (1 - \frac{v^2}{u^2})}}.$$

Gdy prędkość dźwięku jest większa od prędkości samolotu ($v > u$), wyróżnik równania (2) jest dodatni dla każdego φ , $x_1 > 0$, $x_2 < 0$, zatem $x = x_1$. Dla samolotu naddźwiękowego ($u > v$) wyróżnik równania (2) jest dodatni, gdy $\cos \varphi < \frac{v}{u}$ – równanie (2) ma wtedy dwa pierwiastki dodatnie. Każdy punkt trajektorii samolotu jest źródłem fali kulistej, obwiednia tych fal tworzy powierzchnię stożkową (rys. 3), która przesuwa się z prędkością samolotu u . Po raz pierwszy obserwator słyszy dźwięk wysłany przez samolot z punktu B_0 , gdy $\sin \alpha_0 = \cos \varphi_0 = \frac{v}{u}$. W następnych chwilach do obserwatora docierają czoła dwóch fal kulistych (rys. 4) i wydaje mu się, że z punktu B_0 poruszają się w przeciwnie strony dwa źródła dźwięku.

631. Ponieważ tarcie między walcami jest zaniedbywalne, prędkość kątowna ruchu obrotowego pierwszego walca $\omega_0 = \frac{v_0}{R}$ nie zmienia się w wyniku zderzenia. Z zasad zachowania pędu i energii (zderzenie sprężyste) wynika, że po zderzeniu prędkość ruchu postępowego pierwszego walca wynosi zero, a drugiego v_0 . Po zderzeniu na pierwszy walec działa siła tarcia skierowana do przodu. Jego prędkość ruchu postępowego rośnie liniowo z czasem: $v = \mu g t$, prędkość kątowna ruchu obrotowego maleje liniowo z czasem: $\omega = \omega_0 - \frac{\mu g t}{R}$. Po czasie t_0 , gdy spełniony jest związek $v = \omega R$, rozpoczyna się toczenie bez poślizgu. Stąd $t_0 = \frac{v_0}{2\mu g}$. Na drugi walec działa siła tarcia skierowana do tyłu. Jego prędkość ruchu postępowego maleje liniowo z czasem, prędkość ruchu obrotowego rośnie liniowo z czasem. Po czasie t_0 drugi walec również zaczyna toczyć się bez poślizgu. Od chwili t_0 prędkości ruchu postępowego obu walców wynoszą $\frac{v_0}{2}$. Drogi przebyte przez walce w czasie t_0 wynoszą odpowiednio $s_1 = \frac{v_0 t_0}{2}$ i $s_2 = \frac{3 v_0 t_0}{4}$. Maksymalna odległość, na jaką się oddalą, jest równa $\Delta s = \frac{v_0^2}{8 \mu g}$.

Prosto z nieba: Brakujące ogniwo

W okrągłą pięćdziesiątą rocznicę odkrycia pulsarów dostarczamy garść nowych informacji o tych fascynujących – relatywistycznych i niezwykle gęstych – obiektach. Jedną z podstawowych i wciąż nierozwiązanych zagadek związanych z pulsarami jest mechanizm powstawania promieniowania radiowego. Z obserwacji ponad dwóch tysięcy pulsarów w Galaktyce oraz poza nią wiemy, że do powstania pulsu potrzebna jest szybka rotacja (setki obrotów na sekundę) gwiazdy wyposażonej w ogromne, jak na gwiazdowe standardy, pole magnetyczne: 10^4 T lub większe. Pulsar jest w istocie gigantycznym dipolem magnetycznym o osi niepokrywającej się z osią rotacji gwiazdy. Otoczenie gwiazdy – magnetosfera – jest wypełnione plazmą (naładowanymi cząstkami: elektronami, pozytonami, protonami) „wyrwanymi” z powierzchni gwiazdy przez ogromną różnicę potencjałów wywołaną polem magnetycznym. Pole zmusza plazmę do obrotu z częstotliwością obrotu gwiazdy aż do miejsca, w którym prędkość liniowa plazmy przekroczyłaby prędkość światła. Krytyczna odległość wyznacza „cylinder świetlny”, dzielący magnetosferę na obszary, w których linie pola magnetycznego są zamknięte wewnątrz cylindra (w okolicach równikowych dipola), oraz obszar otwartych linii (w okolicach biegunowych dipola). Plazma obszar otwartych linii jest źródłem promieniowania radiowego. Cząstki są przyspieszane wzdłuż linii pola, emitując m.in. promieniowanie krzywiznowe, oraz oddziałując z fotonami (kreacja i anihilacja par elektron-pozyton); mogą także opuścić okolicę pulsara. Podobnie do światła latarni morskiej snop spójnego promieniowania z okolic biegunowych magnetosfery dociera do dogodnie usytuowanego obserwatora.

Niebo w maju

Dobrym celem obserwacyjnym w trakcie majowych nocy będzie gromada kulista M5 (oznaczana jako Messier 5 lub NGC 5904). Gromadę tę odkrył Gottfried Kirch w 1702 roku, przy okazji obserwacji komety C/1702 H1, natomiast Messier uwzględnił ją w swoim katalogu, jako obiekt mgławicowy, jednak nie dostrzegł tam gwiazd. Dopiero pod koniec XVIII wieku Herschel zaobserwował około 200 gwiazd w gromadzie M5. Obecnie wiemy, że gromada ta zawiera ponad pół miliona gwiazd, co czyni ją jedną z największych i najmasywniejszych gromad na naszym niebie. M5 znajduje się w gwiazdozbiorze węża (łac. *Serpens*), a dokładniej w jego głowie (łac. *Serpens Caput*). Co ciekawe, Wąż jest jedynym gwiazdozbiorem podzielonym na dwie części – Głowę Węża oraz Ogon Węża, rozdzielonych gwiazdozbiorem Wężownika. Szukając M5, warto poszukać najjaśniejszej gwiazdy konstelacji Węża, czyli α Ser (znanej jako Unukalhai) o jasności $2,6^m$. Gromada M5 znajduje się około 10 stopni na południowo-zachód od α Ser, a dokładniej na współrzędnych: rektascensja 15,3 h i deklinacja $+02^\circ 04'$. W maju będzie widoczna od 22:00 do 3:00 rano, czyli praktycznie przez całą noc. Po zmroku M5 znajdować się będzie 31° nad południowo-wschodnim horyzontem, przed świtem natomiast 32° , patrząc w kierunku południowo-zachodniego nieba. Jasność gromady to $5,8^m$, jest to zatem obiekt raczej do obserwacji przez lornetki lub małe teleskopy.

Miłośnikom spadających gwiazd polecamy majowy rój η -Akwarydy. Ten deszcz meteorów dostępny do obserwacji będzie od 24 IV do 20 V, a jego maksimum wypadnie 6 V. η -Akwarydy są jednym z najciekawszych rojów o dużej aktywności, gdyż raptem w ciągu jednej godziny można dostrzec ponad 40 śladów. Radiant znajduje się w konstelacji

Powyższy opis jest z grubsza poprawny w przypadku zwykłych pulsarów w rodzaju pulsara w mgławicy Krab (okres obrotu 0,034 s). Oprócz nich istnieje inna podgrupa gwiazd neutronowych o niezwykle silnym polu magnetycznym, zwanych magnetarami. To obiekty o polach rzędu 10^{11} T, które obracają się o wiele wolniej niż przeciętne radiopulsary: przeciętny okres obrotu to kilka sekund. Charakterystyczną cechą niektórych magnetarów są niezwykle potężne, krótkotrwałe „wybuchy” i emisja promieniowania X i gamma (powtarzalne źródła miękkich promieni gamma, ang. *soft gamma repeaters*), a innych – regularne pulsacje promieniowania X (anomalne pulsary rentgenowskie, ang. *anomalous X-ray pulsars*). Uważa się, że źródłem emisji magnetarów nie jest, jak w przypadku zwykłych pulsarów, energia kinetyczna rotacji gwiazdy, lecz energia zmagazynowana w skomplikowanym i ewoluującym polu magnetycznym.

Do niedawna wydawało się, że klasyczne radiopulsary i magnetary to przedstawiciele zupełnie różnych klas obiektów. Sytuację zmieniły zeszłoroczne obserwacje rentgenowskich teleskopów Fermi, Swift i NuSTAR. Zarejestrowały one serię typowo magnetarowych błysków twardego promieniowania emitowanych przez radiopulsar PSR J1119-6127 (okres obrotu 0,4 s), który do tej pory nie wykazywał się niczym nadmiernie interesującym. Po gwałtownym epizodzie pulsar powrócił w ciągu 10 dni do zwykłej emisji radiowej. Astronomowie wiążą duże nadzieje ze zbadaniem tego niezwykłego „brakującego ogniwa”, które może pomóc w lepszym zrozumieniu mechanizmów pulsarowej emisji w różnych energiach, oraz ewolucji tych obiektów.

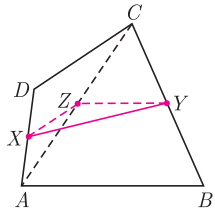
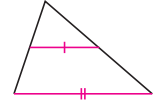
Michał BEJGER

Wodnika na współrzędnych: rektascensja 22,5 h i deklinacja $-01^\circ 04'$. Takie położenie powoduje, że rój widoczny będzie w drugiej połowie nocy, nisko na niebie – raptem 12° nad wschodnim horyzontem. Tak niskie położenie radiantu powoduje, iż można zaobserwować przesuwanie się w górę meteorów, poruszających się z prędkościami około 66 km/s i zostawiających długie smugi. Warto pamiętać, że η -Akwarydy są związane z najsłynniejszą kometą wszechczasów, czyli z kometą Halleya.

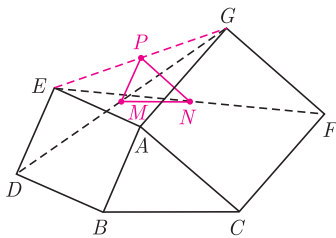
W maju warto obserwować i fotografować trzy bliskie spotkania naszego jedyne naturalnego satelity z planetami Układu Słonecznego. Cztery dni po pełni przypadającej 6 V, w odległości niecałych 2° od Księżyca zobaczymy Jowisza. Para pojawi się zaraz po zachodzie Słońca, 28° nad horyzontem południowo-wschodnim, a jej obserwacje będą mogły potrwać aż do godziny 03:12 nad ranem. Jasność Księżyca wyniesie tej nocy $-12,5^m$, natomiast Jowisza $-2,4^m$, a oba obiekty zobaczymy na tle gwiazdozbioru Panny. W maju Saturn znajdzie się $3^\circ 04'$ od Księżyca. Obserwacje tej pary nie będą proste, zarówno dlatego, że oba ciała wzniosą się tylko 15° nad horyzont, w dodatku najlepszy czas do ich obserwacji wypadnie około godziny 2:52. Saturn osiągnie jasność $0,0^m$, Księżyc zaś $-12,4$, i znajdą się na tle gwiazdozbioru Strzelca. W drugiej połowie miesiąca można również spróbować obserwacji spotkania Księżyca i Neptuna, oddalonych od siebie tylko o $0^\circ 26'$. Jednak układ tych ciał zobaczymy raptem 14° nad horyzontem, dwie godziny przed wschodem Słońca. Jasność Neptuna wyniesie $7,9^m$, natomiast Księżyca (będącego 5 dni przed nowiem wypadającym 25 V) $-11,6^m$. Znajdziemy je na tle konstelacji Wodnika.

Karolina BĄKOWSKA

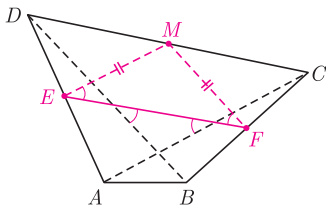
W dowolnym trójkącie odcinek łączący środki dwóch boków jest równoległy do trzeciego boku i dwukrotnie od niego krótszy. Ten prosty fakt okazuje się zadziwiająco przydatny.



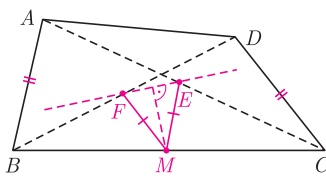
Rys. 1



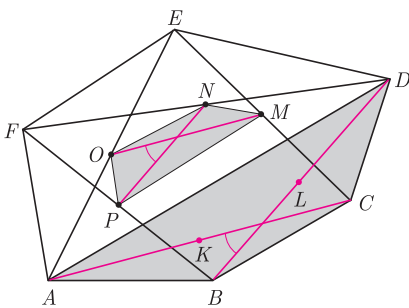
Rys. 2



Rys. 3



Rys. 4



Rys. 5. $[F]$ oznacza pole figury F .

Zadanie 2 pochodzi z LIH Olimpiady Matematycznej, dwa inne rozwiązania opisano w *deltoidach* 17 i 29 (*Delta* 5/2009 i *Delta* 5/2011).

1. Punkty X, Y są środkami odpowiednio boków AD i BC czworokąta wypukłego $ABCD$. Udowodnij, że $XY \leq \frac{1}{2}(AB + CD)$, przy czym równość zachodzi wtedy i tylko wtedy, gdy $AB \parallel CD$.

2. Na bokach AB i AC trójkąta ABC zbudowano, po jego zewnętrznej stronie, kwadraty $ABDE$ i $ACFG$. Punkty M i N są odpowiednio środkami odcinków DG i EF . Wyznacz możliwe wartości wyrażenia $MN : BC$.

3. Przekątne AC i BD czworokąta wypukłego $ABCD$ są równej długości. Punkty E i F są odpowiednio środkami boków AD i BC . Udowodnij, że prosta EF tworzy równe kąty z przekątnymi AC i BD .

4. Czworokąt $ABCD$ nie jest równoległobokiem oraz $AB = CD$. Punkty E i F są odpowiednio środkami przekątnych AC i BD . Wykaż, że rzuty prostokątne odcinków AB i CD na prostą EF są równej długości.

5. W sześciokącie wypukłym $ABCDEF$ o polu 1 punkty K, L, M, N, O, P są środkami odpowiednio przekątnych AC, BD, CE, DF, EA, FB i tworzą sześciokąt wypukły $KLMNOP$. Wyznacz jego pole.

Rozwiązania

R1. Niech Z będzie środkiem przekątnej AC (rys. 1). Wówczas $XZ \parallel CD$, $XZ = \frac{1}{2}CD$, $ZY \parallel AB$ oraz $ZY = \frac{1}{2}AB$. Stąd na mocy nierówności trójkąta dla punktów X, Y, Z mamy $XY \leq XZ + ZY = \frac{1}{2}(AB + CD)$, przy czym równość zachodzi wtedy i tylko wtedy, gdy $XZ \parallel ZY$, czyli gdy $AB \parallel CD$. \square

R2. Niech P będzie środkiem odcinka EG (rys. 2). Wówczas $PM \parallel ED \parallel AB$, $PM = \frac{1}{2}ED = \frac{1}{2}AB$, $PN \parallel GF \parallel AC$ oraz $PN = \frac{1}{2}GF = \frac{1}{2}AC$. Wobec tego trójkąty PMN i ABC są podobne w skali $1 : 2$, a więc $MN : BC = 1 : 2$. \square

R3. Oznaczmy środek boku CD przez M (rys. 3). Wówczas $ME = \frac{1}{2}AC = \frac{1}{2}BD = MF$. Wobec tego trójkąt MEF jest równoramienny i podstawa EF tworzy równe kąty z bokami ME i MF . Jednocześnie $ME \parallel AC$ oraz $MF \parallel BD$, co kończy dowód. \square

R4. Niech M będzie środkiem boku BC . Wówczas $ME \parallel AB$, $MF \parallel CD$ oraz $ME = \frac{1}{2}AB = \frac{1}{2}CD = MF$ (rys. 4). Wobec tego trójkąt MEF jest równoramienny ($E \neq F$, gdyż $ABCD$ nie jest równoległobokiem). Stąd rzut M na prostą EF jest środkiem podstawy EF , a więc rzuty boków ME i MF na prostą EF są równej długości jako połówki podstawy. Wobec tego również dwukrotnie od nich dłuższe rzuty odcinków AB i CD są równej długości. \square

R5. Niech α oznacza miarę nie większego z kątów pomiędzy przekątnymi AC i BD , wówczas $[ABCD] = \frac{1}{2}AC \cdot BD \cdot \sin \alpha$ (rys. 5). Jednocześnie $MO \parallel AC$, $MO = \frac{1}{2}AC$, $NP \parallel BD$ oraz $NP = \frac{1}{2}BD$, zatem kąt pomiędzy odcinkami MO i NP także jest równy α oraz

$$[MNOP] = \frac{1}{2}MO \cdot NP \cdot \sin \alpha = \frac{1}{8}AC \cdot BD \cdot \sin \alpha = \frac{1}{4}[ABCD].$$

Analogicznie $[PKLM] = \frac{1}{4}[DEFA]$, stąd $[KLMNOP] = \frac{1}{4}[ABCDEF] = \frac{1}{4}$. \square

Zadania domowe

6. Niech K, L, M, N będą środkami kolejnych boków czworokąta $ABCD$. Wykaż, że $KLMN$ jest równoległobokiem, że $AC \perp BD \Leftrightarrow KM = LN$, że $AC = BD \Leftrightarrow KM \perp LN$ oraz wyznacz stosunek pól $[KLMN] : [ABCD]$.

7. Dany jest trójkąt ABC o bokach $AB = 2$ oraz $AC = BC = 3$. Punkt K jest środkiem boku BC , punkt L leży na boku AC oraz $KL = 1$. Wyznacz długość odcinka AL .

Wskazówka. $AL = \frac{3}{2}$ to tylko jedna z możliwości.