gdzie h_0 jest wysokością początkową (w chwili $t = 0$), a $e \simeq 2{,}72$ jest podstawą logarytmów naturalnych. Wygodnie jest zapisać ten wzór w postaci:

$$\log_{10} h = \log_{10} h_0 - ABt \log_{10} e.$$

Widać, że logarytm h zmienia się w czasie liniowo. Przedstawiając więc logarytm zmiennej wysokości h na wykresie jako funkcję czasu, możemy przekonać się, czy otrzymany wzór jest słuszny, czy nie (jeśli tak — otrzymamy linię prostą). Z braku miejsca nie przytoczę podobnego rozumowania, z którego wynika, że przy kwadratowej zależności oporu od prędkości (a więc przy prędkości wypływu proporcjonalnej do pierwiastka z ciśnienia) zależność h od czasu powinna być kwadratowa:

$$h(t) = h_0(t-t_0)^2,$$

gdzie h_0 jest wysokością początkową, a t_0 czasem, po którym wysokość spadnie do zera. W tym przypadku przedstawiając pierwiastek z h jako funkcję czasu powinniśmy otrzymać na wykresie prostą. Jak jest w rzeczywistości — przekonajcie się sami.

Niezależnie jednak od tego, co Wam wyjdzie, pamiętając wyniki doświadczenia z ubiegłego miesiąca macie prawo zapytać:

CZY ZALEŻNOŚĆ KWADRATOWĄ MOŻNA WYJAŚNIĆ TEORETYCZNIE?

W każdym razie łatwo wyjaśnić odstępstwa od zależności liniowej. U podstaw naszych rozwiązań teoretycznych leżało bowiem założenie, że ruch cieczy odbywa się w warstwach. Taki ruch nazywa się **laminarny.** Doświadczenie uczy, że powyżej pewnej prędkości zaczynają tworzyć się wiry, ośrodek silnie się miesza — występuje przepływ **burzliwy,** czyli **turbulentny.** Wtedy część potencjalnej energii ciśnienia przetwarza się na energię kinetyczną płynu w wirze. Przy prędkości zwiększającej się coraz bardziej ta część będzie miała coraz większe znaczenie, aż wreszcie stanie się dominująca. Jaka jest zależność oporu od prędkości przepływu przy całkowitej zamianie energii ciśnienia na kinetyczną, możecie wywnioskować z zasady zachowania energii. Znane szkolne zadanie rozważa wypływ cieczy z naczynia przez małą dziurkę. Wiadomo, że otrzymuje się wtedy prędkość proporcjonalną do pierwiastka z ciśnienia:

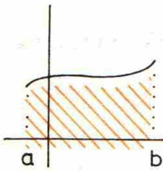
$$v = \sqrt{2gh} = \sqrt{\frac{2p}{\varrho}},$$

gdzie ϱ jest gęstością cieczy.

Odpowiada to kwadratowej zależności oporu od prędkości. Widać więc, że przepływ laminarny i skrajnie burzliwy są pewnymi wyidealizowanymi przypadkami granicznymi. W obszarze przejściowym zależność oporu od prędkości jest bardzo skomplikowana, a jej opis teoretyczny bardzo trudny.

Rozwiązanie zadania M 48

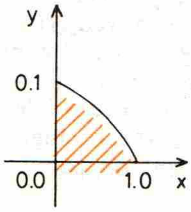
Całkę funkcji ciągłej i nieujemnej można interpretować jako pole. Jak wiadomo, $\int_a^b f(x)\,dx$ jest równa polu zakreślonemu:



Aby zastosować tę uwagę, zmieńmy w pierwszej całce oznaczenia z x na y. Mamy więc wykazać

$$\int_0^1 \sqrt[m]{1-y^n}\,dy = \int_0^1 \sqrt[n]{1-x^m}\,dx.$$

Rysując odpowiedni wykres stwierdzamy, że w obu przypadkach liczymy pole tego samego obszaru, który jest ograniczony odcinkami



$$0 \leqslant x \leqslant 1,$$
$$0 \leqslant y \leqslant 1$$

i krzywą

$$x^m + y^n = 1,$$

raz przyjmując za zmienną niezależną y, raz — x.

O wyszukiwaniu informacji

Prof. dr Zdzisław PAWLAK

Książka telefoniczna, katalog w bibliotece, lista płacy w przedsiębiorstwie, spis lokatorów domu, mały rocznik statystyczny, karty pracy robotników jakiegoś przedsiębiorstwa, katalog znaczków pocztowych, kartoteka przestępców, kartoteka pacjentów w przychodni lekarskiej — wszystko to są przykłady zbiorów informacji.

Chcąc poznać numer telefonu kolegi, znaleźć interesującą nas książkę w bibliotece czy określić wartość nowo zdobytego znaczka pocztowego, należy zajrzeć do książki telefonicznej, katalogu bibliotecznego bądź filatelistycznego.

Wyszukiwanie potrzebnych nam informacji jest jedną z najczęściej występujących czynności we wszelkiego rodzaju poczynaniach: w życiu codziennym, pracy, nauce, rozrywce.

Aby jednakże znaleźć to, czego szukamy w interesującym nas katalogu czy spisie, musimy wiedzieć, o co nam chodzi; inaczej mówiąc, musimy jakoś scharakteryzować poszukiwane informacje przez podanie nazwiska i adresu kolegi, autora i tytułu książki bądź serii i kraju, z którego pochodzi znaczek. Proces znajdowania interesującej nas informacji jest we wszystkich podanych przykładach bardzo prosty, i sprowadza się do przeszukania odpowiednich spisów (katalogów).

Sprawa się znacznie komplikuje, gdy zbiory informacji są bardzo duże. Np. gdybyśmy chcieli stworzyć i wykorzystywać spis wszystkich obywateli kraju, w którym każdy obywatel jest dokładnie scharakteryzowany, albo centralny krajowy katalog książek czy też światowy rozkład lotów samolotów, sprawa nie byłaby tak prosta — przede wszystkim z uwagi na bardzo duży czas potrzebny na wyszukanie interesującej nas informacji.

W tak skomplikowanych przypadkach przychodzą z pomocą współczesne maszyny liczące. Można w nich gromadzić olbrzymie zbiory informacji, oraz bardzo szybko — praktycznie prawie natychmiast — wyszukiwać to, czego potrzebujemy. To nowe zastosowanie maszyn liczących rozwinęło się powszechnie od niedawna i co ciekawsze — o czym może nie wszyscy wiedzą obecnie stanowi ono największy procent wszelkich zastosowań (około 70%). A więc współczesna maszyna licząca to przede wszystkim olbrzymi katalog wszelkiego rodzaju informacji, a nie bardzo szybki arytmometr, jak było to jeszcze kilkanaście lat temu.

Apetyt rośnie w miarę jedzenia. Możliwości, które oferuje informatyka, są bardzo duże, ale specjaliści chcieliby, aby były one jeszcze większe, aby zbiory dostępnych informacji były praktycznie nieograniczone, aby proces wyszukiwania jeszcze bardziej usprawnić i przyspieszyć, aby jak najbardziej ułatwić praktyczne wykorzystywanie maszyn liczących do tego celu.

Procesu tego nie da się zrealizować wyłącznie drogą ulepszeń technicznych; konieczne jest również znalezienie nowych metod gromadzenia i wyszukiwania informacji. W formułowaniu i badaniu takich metod bardzo ważną rolę odgrywa matematyka. Pozwala ona na badanie wielu własności systemów informacyjnych, jeszcze zanim zostaną one zrealizowane praktycznie.

W artykule tym przedstawię pewne problemy matematyczne, które powstały w związku ze stosowaniem maszyn liczących do gromadzenia i wyszukiwania informacji.

Rozważania nasze rozpoczniemy od sprecyzowania pojęcia systemu wyszukiwania informacji. Uważamy, że system taki jest określony, gdy dany jest pewien zbiór X, zwany zbiorem obiektów (np. zbiór książek, znaczków czy ludzi), oraz metoda charakteryzowania elementów zbioru; ściślej: dany jest pewien zbiór A, elementy którego nazwiemy deskryptorami, oraz funkcja φ: $A \to \mathfrak{P}(X)$ ($\mathfrak{P}(X)$ jest zbiorem potęgowym). Funkcja φ przypisuje każdemu deskryptorowi $a \in A$ zbiór tych wszystkich elementów zbioru X, które posiadają własność wyrażoną przez deskryptor a. Np. jeżeli X jest zbiorem uczniów w klasie, zaś b oznacza własność posiadania blond włosów, to $\varphi(b)$ jest zbiorem wszystkich blondynów z klasy.

Zbiorem potęgowym zbioru X nazywa się rodzinę wszystkich podzbiorów tego zbioru i oznacza symbolem $\mathfrak{P}(X)$. Jeśli na przykład $X = \{1,2\}$, to $\mathfrak{P}(X) = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}$ i jest zbiorem 4-elementowym. Ogólnie, jeśli X ma n elementów, to $\mathfrak{P}(X)$ ma 2^n elementów.

Przez system wyszukiwania informacji będziemy rozumieli trójkę

$$S = \langle X, A, \varphi \rangle,$$

gdzie X — jest zbiorem obiektów w S, A — jest zbiorem deskryptorów, zaś funkcja φ: $A \to \mathfrak{P}(X)$ nazywa się charakterystyką w S.

W zbiorze A możemy wprowadzić pewien podział, łącząc elementy tego zbioru w klasy w taki sposób, że elementy należące do tej samej klasy charakteryzują pewną cechę obiektu, np. kolor, datę, wagę, etc. Dla przykładu: rubryki w dowodzie osobistym — nazwisko, adres, data urodzenia etc. — są cechami obiektów, konkretne zaś nazwiska, adresy, daty są deskryptorami. Np. „Kowalski, ul. Polna 29, 30 stycznia 1958" są deskryptorami należącymi kolejno do klasy: nazwisko, adres, data urodzenia.

Aby móc opisywać elementy zbioru X w systemie S, wprowadzimy pewien język formalny L_s. Język ten będzie bardzo prosty. Elementami tego języka będą, po pierwsze, wszystkie deskryptory (ściślej ich nazwy, gdyż np. kolor niebieski w różnych językach ma różne nazwy) oraz znaczki: $1, 0, \vee, \wedge, \sim$.

Za pomocą deskryptorów elementarnych i podanych znaczków tworzyć będziemy wyrażenia poprawne naszego języka (zwane termami) w następujący sposób:

1°. Wyrażenia 1,0 oraz deskryptory są termami (elementarnymi) języka L_s.

2°. Jeżeli t, t' są termami, to również termami są wyrażenia:

$(t \vee t')$ — czyt. „t lub t'",
$(t \wedge t')$ — czyt. „t i t'",
$\sim (t)$ — czyt. „nie t".

3°. Tylko wyrażenia otrzymane przez zastosowanie reguł 1° i 2° są termami.

Przykłady termów: niech dla uproszczenia zbiór A będzie zbiorem liter alfabetu łacińskiego. Wtedy wyrażenia $1 \vee a$, $1 \wedge 0$, $(a \vee b) \wedge c$, ~ 1, $\sim ((a \wedge c) \wedge b)$ są termami.

W ten sposób określiliśmy gramatykę języka L_s, wiemy bowiem, jaką postać mają wyrażenia poprawne w naszym języku. Dla zdefiniowania jakiegokolwiek języka nie wystarczy określenie jego struktury gramatycznej; należy określić także znaczenie jego wyrażeń lub, jak to się mówi, określić semantykę języka.

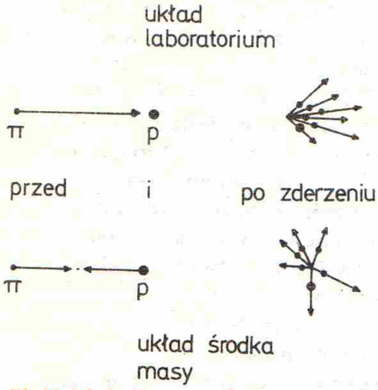
Semantykę języka L_s określimy jak następuje: wprowadzimy funkcję φ^*, którą nazwiemy semantyką, zaś wartość tej funkcji dla dowolnego termu $a \in A$ nazwiemy wartością termu a w systemie S. Funkcja jest określona tak:

1. $\varphi^*(a) = \varphi(a)$,
2. $\varphi^*(a \vee b) = \varphi^*(a) \cup \varphi^*(b)$,
3. $\varphi^*(a \wedge b) = \varphi^*(a) \cap \varphi^*(b)$,
4. $\varphi^*(\sim a) = X - \varphi^*(a)$,
5. $\varphi^*(0) = \phi$,
6. $\varphi^*(1) = X$.

($\cap$, $\cup$ oznaczają odpowiednio iloczyn oraz sumę zbiorów, zaś ϕ jest zbiorem pustym).

dc. F 16

Najwygodniej rozwiązać nasz problem w układzie środka masy zderzających się cząstek. W układzie środka masy, jak wynika z definicji tego układu, wypadkowy pęd cząstek przed i po reakcji wynosi zero. Przebieg zderzenia πp, obserwowany w układzie laboratorium i w układzie środka masy, przedstawiamy obrazowo na rys. 1.



W układzie środka masy cała dostępna energia cząstek początkowych może być w szczególności zużyta na produkcję nowych cząstek, które wówczas spoczywają w tym układzie. A jak taki przypadek wygląda w układzie laboratorium? Wszystkie wyprodukowane cząstki poruszają się z tą samą prędkością, równą prędkości układu środka masy względem laboratorium. Do rozwiązania zadania pozostaje nam tylko wyznaczyć prędkość układu środka masy v_{cm} oraz napisać prawo zachowania pędu lub energii w układzie laboratorium. Nim przystąpimy do rachunków, zauważmy, że dyskutowane cząstki poruszają się z prędkościami bliskimi prędkości światła. Dla przykładu policzmy prędkość nadlatującego pionu:

$$v_\pi = \frac{p_\pi \cdot c^2}{E_\pi},$$

gdzie p_π i E są pędem i energią pionu, a c — prędkością światła. Ponieważ między energią cząstki, jej pędem i masą spoczynkową zachodzi związek relatywistyczny: $E_\pi^2 = p_\pi^2 \cdot c^2 + m_0^2 c^4$, więc dla pionu o pędzie 200 GeV/c porusza się z prędkością

$$v_\pi = (1 - 2{,}5 \cdot 10^{-6})\, c.$$

W naszych obliczeniach musimy stosować wzory mechaniki relatywistycznej. Prędkość środka masy wynosi:

$$v_{cm} = \frac{m_\pi \cdot v_\pi + M \cdot 0}{m_\pi + M} = \frac{p_\pi \cdot c^2}{E_\pi + Mc^2},$$

gdzie M jest masą spoczynkową protonu,

a m_π masą nadlatującego pionu,

$$m_\pi = m_0 \Big/ \sqrt{1 - \frac{v_n^2}{c^2}}$$

Z prawa zachowania pędu, napisanego w układzie laboratorium otrzymujemy związek:

$$p_\pi = \frac{M \cdot v_{cM}}{1 - \frac{V_{cM}^2}{c^2}} + \frac{K m_0 v_{cM}}{\sqrt{1 - \frac{V_{cM}^2}{c^2}}}.$$

Proste przekształcenia prowadzą do ostatecznego wyniku:

$$K = E\left(\frac{1}{m_0}\left(\sqrt{M^2 + m_0^2 + 2ME_\pi/_{c^2}} - M\right)\right)$$

$E(x)$ jest funkcją entier, równą największej liczbie całkowitej mniejszej od x.

Wartość licznika K wynosi 132. Zasada zachowania energii i pędu dopuszcza produkcję znacznie większej liczby cząstek, niż te 16 obserwowane na zdjęciu. Mamy nadzieję, że Czytelnik nie będzie miał kłopotu ze znalezieniem sam odpowiedzi na drugie pytanie postawione w zadaniu (odp 280 MeV/c).

Mówiąc inaczej, wartościami termów w systemie wyszukiwania informacji S są pewne podzbiory zbioru X (w szczególności zbiór pusty oraz cały zbiór X). Mówiąc jeszcze inaczej termy „oznaczają" pewne zbiory.

Język L_s służy nam do opisywania właściwości zbioru obiektów. Oczywiście możemy w języku L_s zamiast symboli używać wyrażeń języka naturalnego. W ten sposób możemy np. podać charakterystykę książek, które nas interesują w związku ze zbliżającym się egzaminem. Za pomocą wyrażeń języka L_s możemy np. napisać, że chcemy, aby bibliotekarka wydała nam książki na temat fizyki jądrowej, autorów Kowalskiego lub Michalskiego, w języku polskim lub rosyjskim, wydane w latach 1950–1971, ale nie w roku 1960 (bo np. z tego roku literaturę już mamy). Język L_s służy więc do opisywania zbiorów informacji.

Język ten ma bardzo ciekawe własności. Jedną z nich przytoczymy tutaj, przedtem jednak wprowadzimy kilka potrzebnych pojęć.

Powiemy, że term a jest atomowy, jeśli ma on postać

$$a = e_0 \wedge \ldots \wedge e_k$$

gdzie e_i $(0 \leqslant i \leqslant k)$ są termami elementarnymi, każdy należy do innej klasy deskryptorów i przy tym w a występują termy ze wszystkich klas deskryptorów systemu S.

Powiemy, że term t jest w postaci normalnej gdy

$$t = a_0 \vee \ldots \vee a_p,$$

zaś wszystkie a_i są termami atomowymi.

Powiemy, że dwa termy są równoważne, jeżeli ich znaczenia są jednakowe, tj. $t \approx t'$ wtedy i tylko wtedy, gdy $\varphi^*(t) = \varphi^*(t')$.

Język L_s ma trzy bardzo ważne własności:

1. Znaczenie każdych dwu różnych termów atomowych t, t' jest rozłączne (tj. jeżeli $t \neq t'$, to $\varphi^*(t) \cap \varphi^*(t') = \phi$).

2. Suma znaczeń wszystkich termów atomowych w systemie S jest równa X, tj.

$$\bigcup_{t \in L_S} \varphi^*(t) = X.$$

3. Dla każdego termu t w L_S istnieje w L_S term t' w postaci normalnej równoważny z t: $\varphi^*(t) = \varphi^*(t')$.

Powyższe trzy własności mają bardzo poważne konsekwencje praktyczne i teoretyczne (a może nawet i filozoficzne, o czym na końcu). Wynika z nich mianowicie, że każdy system wyszukiwania informacji w naturalny sposób wyznacza podział zbioru obiektów X na klasy rozłączne, które nazwiemy atomami systemu S. Każdemu atomowi odpowiada wyrażenie atomowe lub, inaczej mówiąc, każdy atom jest jednoznacznie opisany przez wyrażenie atomowe. Własność 2 zaś mówi, że każdy podzbiór zbioru X, który możemy opisać w języku systemu informacyjnego, jest sumą atomów. Znaczy to, że ogólnie biorąc w języku informacyjnym nie możemy opisać dowolnego podzbioru zbioru X!. Ilustruje to może najlepiej następujący przykład: Gdybyśmy potraktowali jakąś bibliotekę jako system wyszukiwania informacji i opracowali dla niej odpowiedni język informacyjny, to w wyniku tego moglibyśmy utworzyć z książek zbiory atomowe, pakując je np. w paczki, i wszelkie życzenia czytelników, które można wyrazić w języku informacyjnym, byłyby zrealizowane tylko poprzez sumy takich paczek.

Własności te mają szczególne znaczenie dla maszynowego realizowania systemów wyszukiwania informacji. Pozwalają one, zamiast szukać interesujących nas informacji w całym zbiorze, od razu określić miejsce, gdzie informacje te się znajdują. Nie potrzeba wyjaśniać, jakie to ma znaczenie dla przyspieszenia procesu wyszukiwania informacji. Warto dla wyjaśnienia dodać, że jak dotychczas aktualnie istniejące systemy wyszukiwania informacji nie są oparte na podanych tu zasadach; prowadzone są dopiero próby realizacji tej idei. Należy się spodziewać, że przyniosą one pozytywne rezultaty.

Warto się przy okazji poruszanych problemów zastanowić nad pytaniem bardziej ogólnej natury. Czy na przykład opisywanie zbiorów informacji możliwe jest jedynie w podany sposób? Co to jest właściwie informacja? Jakie ma ona podstawowe własności? — etc.

Należy dodać, że pojęcie informacji, które jest nam potrzebne do celów podanych w tym artykule, nie ma nic wspólnego z pojęciem informacji występującym w tzw. teorii informacji stworzonej przez Shannona w latach czterdziestych. W istniejącej aktualnie teorii informacji punktem wyjścia określenia informacji były całkiem inne fakty nawiązujące do przesyłania sygnałów w sieciach telekomunikacyjnych. Natomiast w latach trzydziestych naszego stulecia logik i filozof amerykański Rudolf Carnap zastanawiał się nad pojęciem informacji właśnie w sensie zbliżonym do tego, o którym pisaliśmy w tym artykule. Być może uda się stworzyć nową „teorię informacji" wychodzącą z faktów, których dostarczyły współczesne maszyny liczące.

Na zakończenie jeszcze jedna uwaga. Niemal wszystkie problemy, które powstają w związku z konstrukcją i zastosowaniem maszyn liczących, mają bardzo różnorodne aspekty — od aspektów związanych bezpośrednio z zastosowaniami aż do spraw natury, daleko wykraczających poza problematykę maszyn liczących. Jest to prawdopodobnie przyczyną tego, że informatyka — nauka o maszynach liczących i metodach ich użytkowania — jest tak pociągająca.