Teoria złożoności obliczeniowej

Algorytm za milion dolarów?

Marek Kubale
Wydział Elektroniki, Telekomunikacji
i Informatyki Politechniki Gdańskiej
członek Komitetu Informatyki Polskiej Akademii Nauk
kubale@eti.pg.gda.pl
Prof. Marek Kubale jest profesorem Katedry Algorytmów
i Modelowania Systemów Politechniki Gdańskiej

Komputery zdobyły już sobie prawo obywatelstwa niemal wszędzie. Ale komputer to nie tylko sprzęt, lecz także oprogramowanie. Niektóre wynalazki w dziedzinie oprogramowania wywarły bezpośredni wpływ na jakość naszego życia

Zasadnicze znaczenie ma tu pojęcie algorytmu. Słowo to pochodzi od przydomka Muhammeda ibn Musa, średniowiecznego matematyka perskiego, którego z arabska zwano al-Chwarizmi, co po łacinie brzmiało Algorismus. Przez wieki nie było formalnej definicji algorytmu, ale od czasów starożytnych tworzono opisy rozwiązywania różnych problemów (nie wyłączając dietetycznych). Obecnie te opisy nazywamy algorytmami. Dzisiaj powiedzielibyśmy, że algorytm to jednoznaczny sposób przetworzenia w skończonym czasie pewnych danych wejściowych do pewnych danych wyjściowych, czyli wyników.

2300 lat algorytmiki

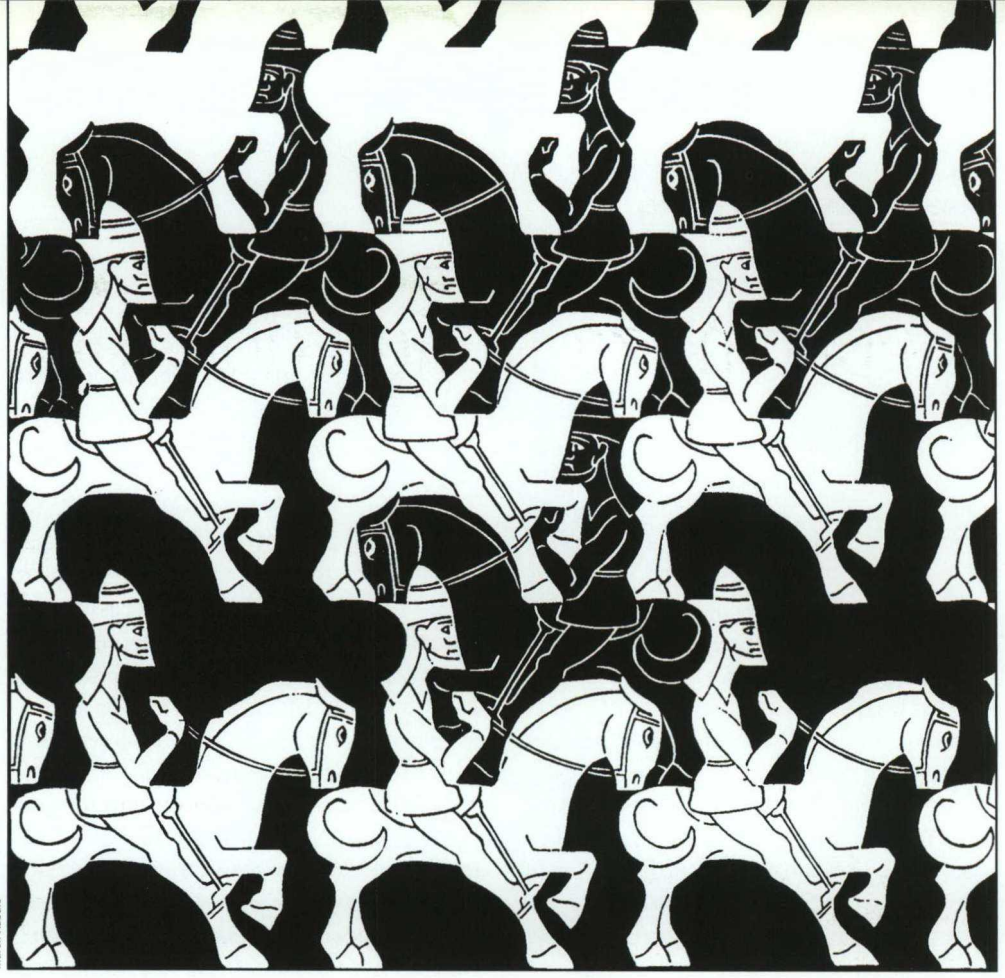
Za pierwszy nietrywialny algorytm uznaje się przepis Euklidesa służący do obliczania największego wspólnego dzielnika dwóch liczb naturalnych. Co prawda był on znany już wcześniej, ale opisał go dopiero Euklides około roku 300 p.n.e. w „Elementach".

Drugim algorytmem wartym wzmianki był zastosowany około roku 100 n.e. sposób sprawdzania liczby żołnierzy przez chińskich generałów. Dzisiaj algorytm ten, znany pod nazwą chińskiego twierdzenia o resztach, ma o wiele więcej zastosowań.

Na przełomie XIX i XX wieku matematyków zainteresowały pytania ogólne: co można obliczyć, jakie funkcje są obliczalne, dla jakich problemów istnieją algorytmy, i szerzej - czy wszystkie twierdzenia matematyczne można udowodnić lub obalić? W 1900 roku wielki uczony niemiecki Dawid Hilbert wśród 23 wyzwań dla matematyków jako dziesiąty problem sformułował pytanie: czy istnieje algorytm, który dla dowolnego równania wielomianowego wielu zmiennych o współczynnikach całkowitych ma rozwiązanie w liczbach całkowitych? Dopiero po 70 latach matematyk rosyjski Jurij Matijasiewicz odpowiedział negatywnie na to pytanie. Dziesiąty problem Hilberta wywołał olbrzymie zainteresowanie obliczalnością - dziedziną, która zajmuje się poszukiwaniem odpowiedzi m.in. na pytanie,

Komputery stają się coraz bardziej wszechobecne w naszym codziennym życiu

www.sxc.hu



Marek Kubale

Kafelkowanie oparte na motywach grafiki Regular Space Division III Mauritsa Cornelisa Eschera

jakie problemy mają rozwiązanie w postaci algorytmu, a jakie go nie mają.

W roku 1936 matematyk brytyjski Alan Turing opublikował rozprawę „O liczbach obliczalnych z ich zastosowaniem dla problemu rozstrzygalności". Odpowiedział w niej na najważniejsze pytanie, jakie stawiali sobie na początku XX wieku matematycy i filozofowie matematyki: czy da się stworzyć maszynę, która będzie rozstrzygać automatycznie, bez udziału człowieka, prawdziwość twierdzeń matematycznych? Wykazał, że taka maszyna nigdy nie powstanie, a w dowodzie wykorzystał pojęcie abstrakcyjnej maszyny obliczeniowej nazwanej potem jego imieniem. Dziś uczą się o niej studenci informatyki jako o teoretycznym modelu komputera, który w rzeczywistości powstał dopiero kilka lat później.

Pierwsze elektroniczne maszyny cyfrowe (komputery) zostały skonstruowane w czasie II wojny światowej. Miały znaczne rozmiary, a ich moc obliczeniowa była miliardy razy mniejsza niż we współczesnych komputerach osobistych, jednak dzięki nim algorytmika zyskała nowe bodźce rozwoju. Beneficjentami nowego wynalazku były wszystkie dziedziny nauki, w szczególności zaś kombinatoryka i teoria grafów.

Przełomowy dla algorytmiki współczesnej był rok 1971, kiedy informatyk amerykański Steven Cook wykazał, że problem spełnialności formuł zdaniowych jest trudny obliczeniowo. W ten sposób stworzył podwaliny bardzo ważnej klasy problemów, zwanych NP-trudnymi (nieprzeciętnie trudnymi?). Obok kilku tysięcy problemów kombinatorycznych należą do niej także problem komiwojażera i łamanie szyfrów.

Siedem problemów milenijnych

Naukowcy z Clay Mathematics Institute w Massachusetts (USA) w roku 2000, wzorując się na pomyśle Hilberta sprzed 100 lat, sformułowali 7 otwartych problemów matematycznych dotychczas opierających się fachowcom. Za rozwiązanie każdego z tych problemów milenijnych ufundowano nagrodę w wysokości miliona dolarów. Na czele listy znajduje się pytanie algorytmiczne o to, czy P=NP. Mówiąc w dużym uproszczeniu, dylemat ten oznacza, że nie wiemy, czy niektóre trudne problemy obliczeniowe, takie jak na przykład problem komiwojażera, nie mogą być szybko rozwiązane dlatego, iż jest to niemożliwe (P≠NP), czy też przeciwnie - jest to możliwe, lecz ludziom nie starcza pomysłowości, aby taki szybki algorytm zaimplementować (P=NP).

Dziś spośród 7 problemów milenijnych otwartych pozostało 6, gdyż jeden z nich (hipoteza Poincarego) został rozwiązany w roku 2003. Nagroda nie została wypłacona, gdyż autor rozwiązania, matematyk

rosyjski Grigori Perelman, odmawia nie tylko przyjmowania wszelkich nagród, ale także kontaktów z mediami.

Rodzaje problemów

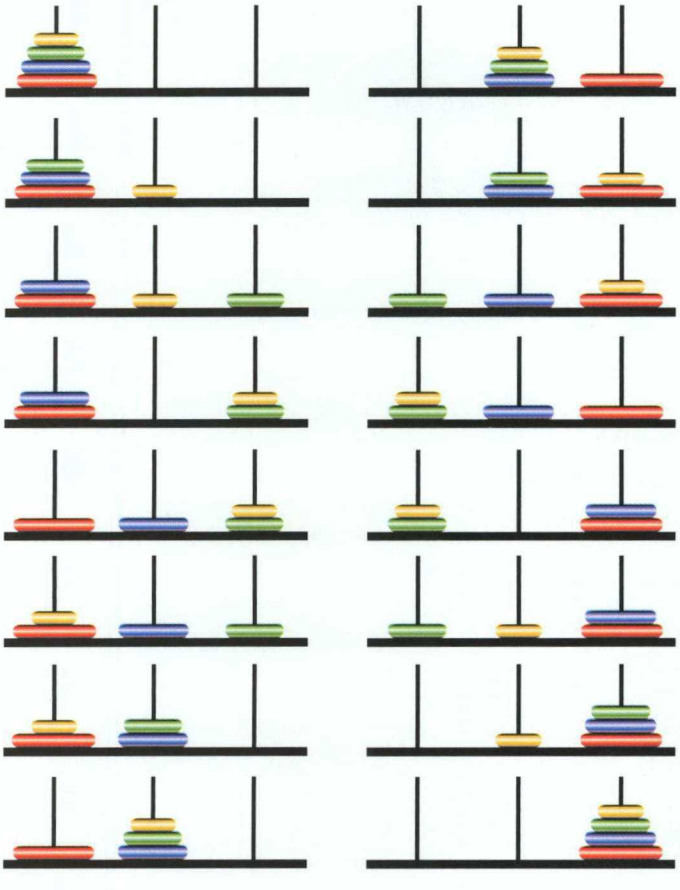
Powiedzenie, że problem może być rozwiązany za pomocą algorytmu, oznacza, że można napisać program komputerowy, który w skończonym czasie da poprawną odpowiedź dla każdych poprawnych danych wejściowych przy założeniu dostępu do nieograniczonych zasobów pamięciowych. Bardzo ważny jest wymóg efektywności czasowej. Dlatego rozważane problemy można podzielić na pięć klas.

Pierwsza to problemy niealgorytmiczne, które nie mogą być rozwiązane za pomocą programów komputerowych. Przykładem jest problem kafelkowania, polegający na rozstrzygnięciu, czy można pokryć płaszczyznę identycznymi kopiami danego wielokąta. Istnienie problemów niealgorytmicznych jest dowodem na to, iż umysł ludzki potrafi robić coś więcej, niż mogą wykonywać komputery – może pracować niealgorytmicznie. Tym samym stworzenie sztucznej inteligencji dorównującej inteligencji właściwej człowiekowi nie jest możliwe. Drugi typ to problemy przypuszczalnie niealgorytmiczne, dla których nie udało się dotychczas podać algorytmu skończonego, ale brak też dowodu, że taki algorytm nie istnieje.

Trzecia kategoria to problemy wykładnicze, dla których czas dojścia do rozwiązania rośnie wykładniczo ze wzrostem rozmiaru rozwiązywanego zadania. Przykładem tego typu jest słynny problem wież Hanoi, łamigłówka w postaci klocków, znana wszystkim przedszkolakom. Problem polega na przenoszeniu krążków z pałeczki pierwszej na trzecią, przy ograniczeniu, że nie wolno położyć krążka większego na mniejszy. Do jego rozwiązania trzeba wykonać 2^n-1 operacji. Związana jest z nim legenda głosząca, że w pewnym klasztorze w Hanoi mnisi buddyjscy przenoszą 64 złote krążki w tempie 1 krążek na sekundę. Z chwilą przeniesienia ostatniego krążka nastąpi koniec świata. Ile zatem zostało nam jeszcze czasu?

Czwarty typ to problemy przypuszczalnie wykładnicze, dla których nie udało się dotychczas podać algorytmu wielomianowego, ale brak też dowodu, że taki algorytm nie istnieje. Przykładem takiego problemu jest faktoryzacja, czyli znalezienie rozkładu danej liczby na czynniki pierwsze. Ostatnia kategoria to problemy wielomianowe, dla których istnieją algorytmy rozwiązujące je w czasie ograniczonym wielomianowo, jak na przykład zagadnienie sortowania, które jest bardzo ważnym problemem algorytmicznym. Ktoś kiedyś zauważył, że przypuszczalnie ponad połowa komputerów na świecie zajmuje się sortowaniem.

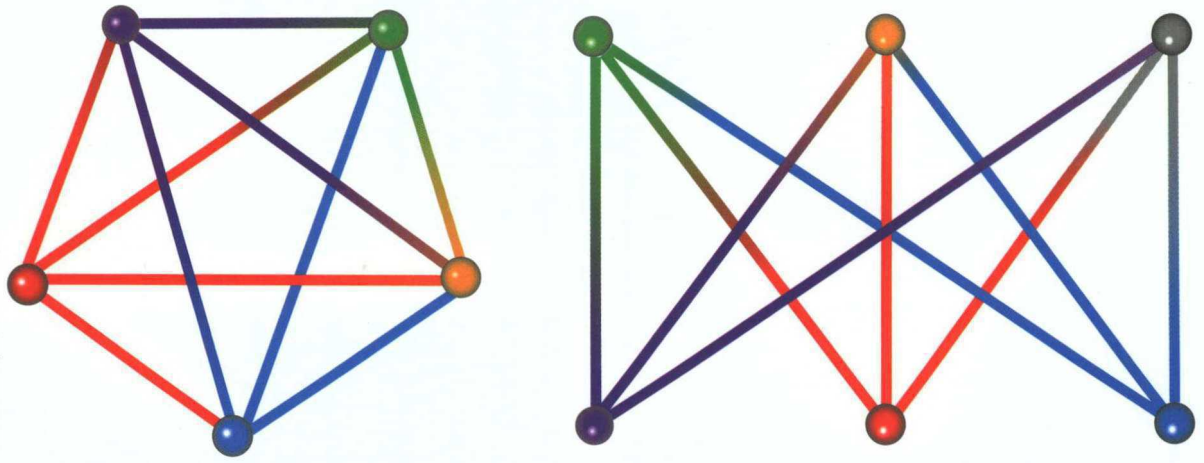
Wieże Hanoi – zabawka przedszkolaka czy poważny problem algorytmiczny? Zadanie polega na przeniesieniu krążków z pierwszej pałeczki na trzecią w taki sposób, aby nigdy nie kłaść większego krążka na mniejszym



Paweł Adamów na podstawie http://mathworld.wolfram.com/

Planowanie optymalne

Historia algorytmiki obfituje w spektakularne pomysły algorytmiczne. Jednym z nich jest zagadnienie programowania liniowego, ważne z uwagi na liczne zastosowania praktyczne w planowaniu produkcji, przydziale zasobów, szeregowaniu zadań i problemach transportowych. W okresie II wojny światowej pojawił się problem optymalnej diety. Chodzi o to, że produkty spożywcze zawierają substancje odżywcze i witaminy w różnych proporcjach. Znając wielkości zapasów i ceny rozmaitych produktów, można określić sposób zaspokojenia potrzeb żywnościowych przy minimalnych kosztach. W roku 1947 Amerykanin George Dantzig opracował metodę simpleksową do



Paweł Adamów na podstawie Marek Kubale

Grafy Kuratowskiego K5 i K3,3 (oznaczane też żartobliwie K5azimierz K3,3uratowski). Są to najmniejsze grafy nieplanarne

rozwiązania tego problemu. Jej nazwa pochodzi od simpleksu, czyli bryły wypukłej, będącej wielowymiarowym uogólnieniem trójkąta. Teoretycznie jest to metoda wykładnicza, lecz praktycznie bardzo wydajna. Dantzig podał przykład problemu przydziału 70 pracowników do 70 stanowisk, który ma 70! rozwiązań dopuszczalnych. Jest to liczba przeogromna, ale jego algorytm daje rozwiązanie optymalne prawie natychmiast.

W roku 1979 Leonid Chaczijan, matematyk ormiańskiego pochodzenia, opublikował tzw. metodę elipsoidalną, która była pierwszym algorytmem stricte wielomianowym. Wynik Chaczijana ma znaczenie głównie teoretyczne, gdyż szacuje się, że przewaga metody elipsoidalnej nad simpleksową ujawnia się dopiero przy 1000 ograniczeniach i n=50000 zmiennych, ale i tak był to przełom w historii programowania liniowego. Wreszcie w roku 1984 informatyk hinduski Narendra Karmarkar opracował metodę punktu wewnętrznego. W trakcie jej realizacji komputer musi wykonać około $l*n^{3,5}$ operacji zmiennoprzecinkowych, gdzie l to liczba bitów potrzebnych do zapisu danych. Jest to aktualnie najszybszy asymptotycznie algorytm do rozwiązywania problemu programowania liniowego.

Banki i grafy

Kolejnym zagadnieniem jest testowanie pierwszości liczb, ważne ze względu na zastosowania w szyfrowaniu informacji w bankowości, wojskowości, telekomunikacji itp. W roku 2002 trzej matematycy hinduscy, Manindra Agrawal, Neeraj Kayal i Nitin Saxena, podali nowy test pierwszości liczby naturalnej i tym samym problem, który nurtował ludzkość od starożytności, został rozwiązany! Obecnie, po pewnych usprawnieniach, algorytm AKS może być wykonany w czasie rzędu n^6, gdzie n to liczba cyfr. Mimo tego sukcesu testy deterministyczne są ciągle znacznie wolniejsze od probabilistycznych.

Kolejnym ważnym zagadnieniem jest tzw. problem spłaszczania grafu, czyli rysowania grafu na płaszczyźnie (np. ekranu monitora) tak, aby jego krawędzie nie przecinały się. Pod pojęciem grafu rozumiemy tutaj strukturę złożoną z wierzchołków (punktów) i krawędzi (linii). Na początku XX wieku matematyków interesowała odpowiedź na pytanie, jakie są warunki konieczne i wystarczające do tego, aby graf był planarny. Problem ten w roku 1930 rozwiązał nasz rodak Kazimierz Kuratowski, udowadniając, że graf jest planarny wtedy i tylko wtedy, gdy nie zawiera podgrafu homeomorficznego z K5 i K3,3.

Oczywiście prawie 100 lat temu Kuratowski nie zajmował się algorytmami. Z jego dowodu można jednak wywieść pewien algorytm testowania planarności grafu o złożoności n^6. Obecnie, w efekcie pracy Johna Hopcrofta i Roberta Tarjana z roku 1974, potrafimy zaprogramować komputery tak, by rysowały grafy planarne w sposób wydobywający wszystkie symetrie w nich zawarte w czasie wprost proporcjonalnym do ich rozmiaru, tj. w czasie liniowym (w praktyce prawie natychmiast).

Czy P=NP?

Obecnie najważniejszym problemem współczesnej informatyki teoretycznej jest pytanie milenijne o to, czy P=NP? W roku 2002 zapytano 100 czołowych informatyków teoretyków, w którym roku i z jakim skutkiem problem ten zostanie według nich rozwiązany. Odpowiedziało 79 profesorów. W tym gronie 61 respondentów sądziło, że P≠NP, 9 przypuszczało, że jest przeciwnie, a 9 udzieliło innej odpowiedzi. Zdaniem większości odpowiedź poznamy przed rokiem 2040. ■

Chcesz wiedzieć więcej?

Kubale M. (2009). *Łagodne wprowadzenie do analizy algorytmów*, Gdańsk: Wydawnictwo Politechniki Gdańskiej.

Kubale M. i in. (2002). *Optymalizacja dyskretna. Modele i metody kolorowania grafów*, Warszawa: WNT.