

1 Informatyka i nauki pokrewne.

Na każdym kroku słyszymy nazwę informatyka. Wszystko, co wiąże się z komputerami, skłonni jesteśmy określać jako informatykę. Wprowadzamy do szkół przedmiot o nazwie „informatyka”... Czym zajmuje się ta dziedzina wiedzy? Artykuły na tematy informatyczne w prasie najczęściej dotyczą technologii budowy komputerów lub oprogramowania użytkowego i tak naprawdę nic wspólnego z informatyką nie mają. W tym krótkim rozdziale spróbuję przedstawić podstawowe koncepcje informatyki, jej związek ze światem komputerów i ich zastosowań. Jest to najbardziej „teoretyczny” z rozdziałów tej książki.

Jak już wspominałem w poprzednim rozdziale informatyka jest dyscypliną całkiem młodą, powstała w połowie lat 60-tych dzięki rozwojowi matematyki i elektroniki. Ten podział istnieje z grubszą do dzisiaj: informatyka teoretyczna, uprawiana na uniwersytetach, związana jest z matematyką, podczas gdy informatyka techniczna, uprawiana na politechnikach, związana jest z mikroelektroniką i zastosowaniami metod komputerowych. Silne zespoły badawcze informatyków pracują również w wielkich firmach komputerowych.



Największą i najstarszą (założoną w 1947 roku) organizacją skupiającą informatyków jest ACM, Association for **C**omputing **M**achinery (dosłownie „Stowarzyszenie zajmujące się maszyną obliczeniową”). Właśnie to stowarzyszenie opublikowało w 1968 roku zalecenia dla nowo powstających programów studiów informatycznych, dając początek „naukom komputerowym” (computer science) w USA i „informatyce” w Europie. Nie są to zbyt szczęśliwe nazwy: nauka o komputerach bardzo zawęża istotę **algorytmiki**, wiedzy, dotyczącej sposobów rozwiązywania zagadnień w oparciu o szczegółowe przepisy, czyli **algorytmy**.

Algorytmika to fundament informatyki, a jednocześnie bardzo ważna dyscyplina dla techniki, ekonomii oraz wszystkich nauk matematyczno-przyrodniczych. Do sprawdzenia, dokąd doprowadzi nas dany algorytm, najłatwiej jest używać komputerów, ale w zasadzie można to również zrobić przy pomocy ołówka i kartki papieru. W tym sensie informatyka teoretyczna wcale nie zależy od istnienia komputerów a „prawdziwy informatyk” komputerami się brzydzi (do niedawna była to dość często spotykana postawa wśród informatyków, dzisiaj prawie wszyscy używają komputerów przynajmniej jako maszyn do pisania).

Z drugiej strony **informatyka** skupia się na pojęciu „informacji”, również nie oddającej istoty zagadnienia. Definicja encyklopedyczna głosi: „informatyka zajmuje się całokształtem przechowywania, przesyłania, przetwarzania i interpretowania informacji. Wyróżnia się w niej dwa działy, dotyczące sprzętu i oprogramowania”. Większość z tego, co robią informatycy nie bardzo do tej definicji pasuje. Nowsza definicja, opracowana w 1989 roku przez ACM, mówi: „Informatyka to systematyczne badanie procesów algorytmicznych, które charakteryzują i przetwarzają informację, teoria, analiza, projektowanie, badanie efektywności, implementacja i zastosowania procesów algorytmicznych. Podstawowe pytanie informatyki to: co można (efektywnie) zalgorytmizować”. Jest to więc nauka o tym, co można osiągnąć przy pomocy procesów przetwarzania informacji (procesów obliczeniowych) i w tym sensie jest niezależna od sprzętu, umożliwiającego wykonywanie obliczeń. W pismach popularnonaukowych jak i w prasie codziennej pod określeniem „informatyka” rozumie się najczęściej zupełnie coś innego. Omawiając odkrycia minionego roku w astronomii, biologii, chemii, fizyce i informatyce nawet w pismach na wysokim poziomie czytamy z jednej strony o czarnych dziurach i kwarkach, a więc prawdziwych odkryciach naukowych, a z drugiej o wprowadzeniu nowych mikroprocesorów czy nowego systemu operacyjnego firmy Microsoft. Z prawdziwą informatyką nie ma to jednak wiele wspólnego.

1.1 Czym zajmuje się informatyka?

Zainteresowania informatyki jako gałęzi nauki skupiają się wokół rozwiązywania **zadań algorytmicznych**, czyli takich zadań, dla których można znaleźć jakieś formalne sposoby postępowania, przepisy, prowadzące do ich rozwiązania. Na pierwszy rzut oka prawie wszystko da się zalgorytmizować (tj. znaleźć algorytm dla danego zagadnienia), chociaż nie zawsze znane algorytmy zapewniają dokładne rozwiązania. Wiele zagadnień matematycznych da się zalgorytmizować, większość zagadnień związanych z wyszukiwaniem, przesyłaniem i przetwarzaniem informacji, przetwarzaniem dźwięku i obrazu również. Nie zawsze jednak można znaleźć **algorytmy efektywne**, a więc prowadzące do rozwiązania w rozsądnym czasie. Niektóre algorytmy są wprawdzie proste, ale wymagają ogromnych obliczeń, w praktyce trwających nieskończenie długo. Informatyka zajmuje się więc oceną **złożoności obliczeniowej** algorytmów, poszukiwaniem efektywnych algorytmów, oraz testowaniem i dowodzeniem ich **poprawności**.

W tych zagadnieniach, dla których ściśle rozwiązania nie są znane, poszukuje się przybliżonych metod rozwiązywania czyli szuka się **algorytmów heurystycznych**, metod, które nie gwarantują rozwiązania ale często mogą być pomocne. Ten ostatni rodzaj algorytmów rozwijany jest przez specjalistów od **sztucznej inteligencji**, działu

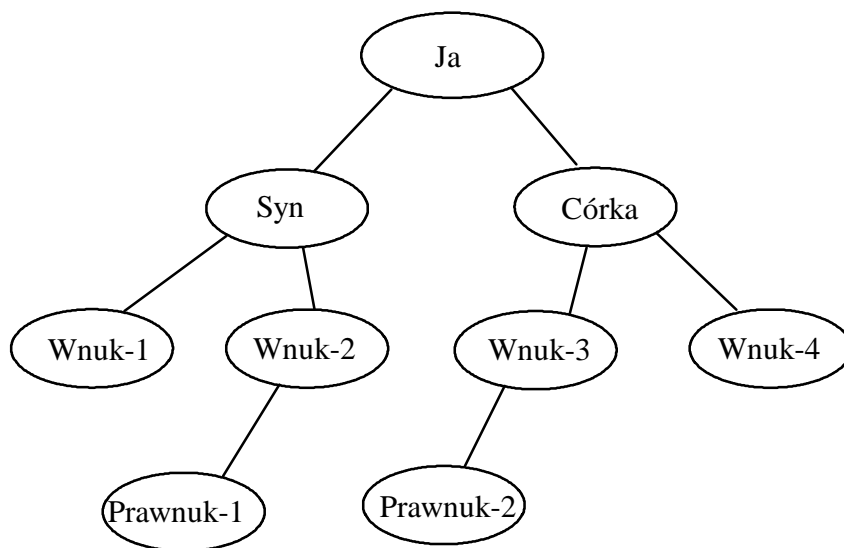
informatyki, zajmującego się procesami dla których nie znamy (być może nie istnieją) efektywnych algorytmów rozwiązań. Sztucznej inteligencji poświęciłem osobny rozdział.

Załóżmy, że chcemy rozwiązać zagadnienie, w którym mamy do czynienia z pewną liczbą N elementów. Problemy algorytmiczne dzieli się na łatwo rozwiązywalne, czyli takie, dla których istnieją (choć nie muszą być znane, to czasami można udowodnić, że istnieją) algorytmy efektywne, pozwalające rozwiązać problem w czasie nie większym niż jakaś potęga liczby N , np. proporcjonalnym do N lub do kwadratu N . Oczywiście należy szukać najbardziej efektywnych algorytmów, których zależność od liczby elementów jest najniższa. Dla przykładu, jest wiele algorytmów sortowania (porządkowania) listy obiektów. Typowe algorytmy wymagają około N^2 operacji dla N obiektów, są jednak i takie, które porządkują w znacznie krótszym czasie, proporcjonalnym do niższej potęgi N . Odkryto jednak blisko 1000 interesujących problemów trudnych, czyli takich, dla których czas potrzebny do ich rozwiązania rośnie szybciej niż dowolna potęga N . Takie trudne zagadnienia określa się jako NP-trudne. NP to skrót od „nie-wielomianowa” (zależność czasu rozwiązania od rozmiaru problemu).

Wiele prostych w sformułowaniu zagadnień optymalizacji należy do zagadnień trudnych, np. „problem wędrującego komiwojażera”: jak znaleźć najkrótszą drogę, łączącą N miast odwiedzając każde tylko jeden raz. Ciekawą klasą złożoności są problemy „NP-zupełne”. Dodanie przymiotnika „zupełne” wiąże się ze zdumiewającym twierdzeniem, iż wszystkie te problemy można by rozwiązać w łatwy sposób, gdyby chociaż jeden dał się łatwo rozwiązać! Czy takie rozwiązanie istnieje? Do dzisiaj nie udało się tego nikomu udowodnić. Wiadomo natomiast, że istnieją zagadnienia, dla których nie można wcale znaleźć algorytmu rozwiązania, zagadnienia nierozstrzygalne, a nawet problemy wysoce nierozstrzygalne. Czytelników zainteresowanych problemami klas złożoności muszą jednak odesłać do literatury dotyczącej podstaw informatyki.

Algorytmy wymagają danych początkowych i tworzą nowe dane, będące rozwiązaniem problemu. Dane mogą mieć skomplikowaną strukturę, stąd jedna z gałęzi informatyki zajmuje się **strukturami danych**, czyli organizacją danych różnych typów w logicznie powiązane ze sobą struktury. Do najprostszych struktur danych należą tablice liczb, do bardzo złożonych należą struktury danych zapisujące informacje o całym przedsiębiorstwie, czyli wzorce, według których zorganizowana jest informacja przechowywana w bazach danych. Przedmiotem badań informatyków są sposoby reprezentacji danych, szyfrowanie i sposoby zabezpieczenia danych przed odczytaniem przez niepowołane osoby a także określeniem ilości informacji, zawartej w danych.

Do najprostszych danych należą liczby, czyli dane numeryczne: mogą to być liczby całkowite należące do ograniczonego zakresu, mogą to być liczby wymierne zapisane z określoną dokładnością (określane też niezbyt precyzyjnie jako „liczby rzeczywiste”),



Często stosowaną strukturą w informatyce są drzewa, wychodzące od węzła zwanego korzeniem, od którego wybiegają węzły potomne. Węzły na najniższym poziomie, nie mające potomstwa, nazywa się liśćmi.

mogą to być również tablice liczb. Jednowymiarowe tablice nazywamy **wektorami**, dwuwymiarowe **tablicami** a wielowymiarowe po prostu **tabelami**. Bardzo ogólną strukturą danych jest **lista**. Uproszczone listy, dla których dozwala się jedynie na operacje dokładania lub usuwania ostatniego elementu nazywa się **stosami** (podobnie jak stos naczyń), natomiast listy, dla których dozwala się na operacje dokładania elementów na końcu i usuwania elementów na początku listy nazywamy **kolejkami**. Kolejki nazywa się też listami FIFO, od angielskiej nazwy „First-In-First-Out”, czyli ten element, który pierwszy wszedł na listę zejdzie z niej pierwszy.

Dane mogą mieć bardziej złożone struktury, których elementy powiązane są ze sobą w skomplikowanych sposób. Często stosowaną strukturą jest **drzewo**, podobne do drzewa genealogicznego (por. rysunek), zawierające powiązane ze sobą węzły. Jeden wyróżniony węzeł początkowy nazywa się **korzeniem**. Odchodzi od niego **węzły potomne**, mające z kolei swoje potomstwo, aż dochodzimy do **liści drzewa**, czyli węzłów bez potomstwa. Drzewo, w którym dany węzeł ma co najwyżej dwa węzły potomne nazywa się drzewem binarnym. Drzewa nie mogą zawierać pętli wewnętrznych, mogą się jedynie rozrastać. Bardziej złożonymi strukturami, dopuszczającymi dowolne powiązania węzłów ze sobą, są **grafy**. Linie łączące ze sobą węzły grafu nazywa się łukami. Mogą one mieć

wyróżniony kierunek określający dozwoloną kolejność przechodzenia od węzła do węzła przy wędrowaniu po grafie. Informatycy wyróżniają wiele rodzajów grafów.

Jedną z ważniejszych dyscyplin informatyki jest **teoria języków programowania**, zwana również teorią języków algorytmicznych, teorią języków formalnych lub teorią języków sztucznych. Formalne zapisanie algorytmu wymaga zdefiniowania jakiegoś zestawu symboli-instrukcji wraz z regułami ich użycia. Reguły te nazywa się gramatyką lub składnią i razem z zestawem symboli tworzą one pewien język programowania. Algorytm zapisany w takim języku nazywa się **programem**. Naukę reguł składni i symboli języka programowania, konstruowania struktur danych dopuszczalnych przez te reguły i zapisywania algorytmów w tym języku formalnym nazywa się **nauką programowania**. Formalne określenie wymagań, dotyczących struktur danych nazywa się **specyfikacją**. Z pisaniem specyfikacji i dowodzeniem, że dany program je spełnia związanych jest wiele ważnych problemów.

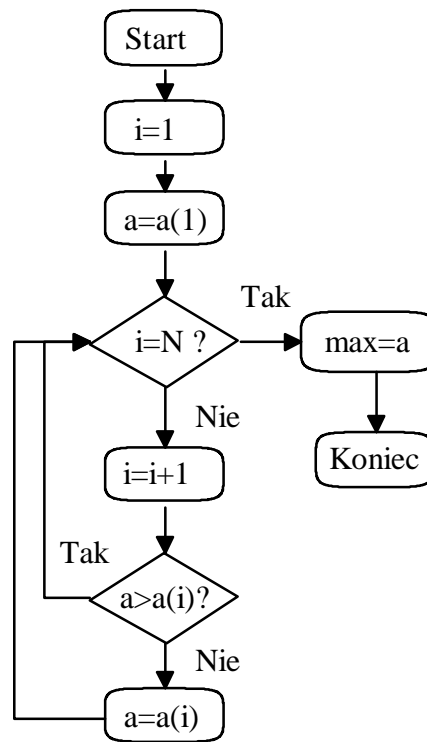


Diagram przedstawiający algorytm szukania maksymalnego elementu dla N -wymiarowego wektora $a(i)$. Test przedstawione są w ramkach prostokątnych.

Programy realizowane są przez **procesory**. Dla teoretyka informatyki procesory są matematycznie opisanymi urządzeniami przyjmujące skończoną liczbę różnych stanów, nazywanymi **automatami skończonymi** lub automatami Turinga, gdyż do zagadnień obliczalności wprowadził je jeszcze w 1936 roku angielski matematyk Alan Turing. Teoria automatów skończonych należy do podstaw informatyki i odpowiada na pytania, jakiego rodzaju problemy można przy pomocy obliczeń rozwiązać, oraz jakie klasy urządzeń są do tego potrzebne. Chociaż prawdziwe mikroprocesory spotykane w komputerach są bardzo złożone ich możliwości obliczeniowe nie odbiegają od tych, które

mają proste konstrukcje matematyczne zwane uniwersalnymi automatami skończonymi. Nowa i bardzo modna klasa zagadnień dotyczy algorytmów realizowanych współbieżnie, to znaczy na wielu procesorach jednocześnie.

Algorytmy wygodnie jest zapisywać w postaci **diagramów** czy schematów blokowych. Takie schematy przedstawiają w formie graficznej każdy krok wykonywanego algorytmu. Analiza algorytmów jest więc związana z teorią grafów. Analiza grafów i drzew jest jedną z ulubionych metod informatyki. Oprócz teorii obliczeń, na którą składa się teoria abstrakcyjnych automatów i teoria złożoności obliczeniowej informatycy interesują się również klasycznymi działami matematyki, przede wszystkim logiką (istnieje nawet taki dział jak logika algorytmiczna, stworzony specjalnie dla informatyków), matematyką dyskretną oraz analizą numeryczną.

Oprócz teorii, której dotyczy większość przedmiotów wykładanych na studiach informatyki, w kręgu zainteresowań informatyków pojawiają się również zagadnienia bardziej praktyczne, takie jak organizacja i architektury systemów komputerowych, systemów operacyjnych i sieci komputerowych, teoria baz danych, inżynieria oprogramowania, optymalizacja działania programów, zapewnienie bezpieczeństwa i niezawodności działania, probabilistyczne teorie wspomagania decyzji. Niektóre uniwersytety i politechniki uczą również takich zagadnień, jak analiza obrazu, teoria symulacji komputerowych i zastosowania metod komputerowych w różnych dziedzinach nauki i techniki.

Przeciętny kandydat na informatyka zupełnie błędnie utożsamia sobie **zastosowania komputerów z informatyką** i ma najczęściej całkiem fałszywe wyobrażenie o tym, czym rzeczywiście jest informatyka. Interesują go komputery, oprogramowanie użytkowe (przede wszystkim gry) i języki programowania. Mamy tu do czynienia z sytuacją podobną, jak w przypadku studiów astronomicznych, mających niewiele wspólnego z tym, o czym marzy każdy miłośnik astronomii: patrzeniem w niebo. Informatyka nie jest nauką o komputerach tak jak astronomia nie jest nauką o teleskopach. Prawdziwy informatyk niekoniecznie zna się na komputerach. Gorzej, bo nie ma dotychczas zgodności nawet co do tego, czy informatyka jako nauka przydatna jest informatykom-praktykom! Zaciekle dyskusje nad rolą matematycznych teorii, takich jak teoria logiki algorytmicznej, specyfikacji programów czy złożoności obliczeniowej w kształceniu inżynierów-programistów toczą się w prasie fachowej od lat. Siłą napędową informatyki jest bez wątpienia rozwój rynku komputerowego i zastosowań komputerów. Nie będę tu się nad tym problemem zatrzymywał, postaram się raczej odpowiedzieć na pytanie, gdzie szukać należy specjalistów od zastosowań komputerów?

1.2 Informatyka i „nauki komputerowe”

Informatyka jest bardzo młodą dziedziną a już z jej pnia wyodrębnia się szereg innych dziedzin. O tym, jak młoda jest to dziedzina świadczyć może fakt, że do połowy lat 90-tych nie mieliśmy jeszcze w Polsce uprawnień zawodowych dla informatyków, takich jak dla inżynierów specjalności budowlanej czy elektrycznej.

Nauki zwykło się dzielić na teoretyczne i doświadczalne. Pojawienie się komputerów o dużej mocy obliczeniowej stworzyło nową jakość, nie pasującą do tego podziału. Prawidłowości, poszukiwane przez nauki przyrodnicze czy społeczne, są algorytmami określającymi zachowanie się systemów. Programy komputerowe pozwalają na zbadanie konsekwencji zakładanych praw, pozwalają na symulację rozwoju skomplikowanych systemów i określanie własności systemów stacjonarnych. Komputery pozwalają na robienie doświadczeń w sytuacjach zbyt skomplikowanych, by możliwa była uproszczona analiza teoretyczna. Doświadczenia komputerowe nie są przy tym ograniczone przez prawa natury istniejącego świata, lecz tylko przez fantazję programisty. Być może świat, w którym żyjemy, jest istotnie jedynym możliwym logicznie niesprzecznym światem, jak chcą tego najnowsze teorie fizyczne - możemy to zbadać symulując światy, w których obowiązują inne prawa niż w naszym. Obliczenia komputerowe, pozwalające znaleźć odpowiedź na pytanie: „co by było, gdyby ...”, przekraczają też granice nauk teoretycznych. Praw natury poszukiwaliśmy dotychczas tworząc modele w oparciu o konstrukcje matematyczne typu funkcji. Algorytmy, wykonywane przez programy komputerowe, obejmują jednak znacznie szerszą klasę modeli niż te, które dają się sformułować w tradycyjnym języku. Wystarczy wspomnieć o dwóch bardzo obecnie modnych dziedzinach komputerowej matematyki: teorii **fraktali** i teorii **automatów komórkowych**.

Słowo „teoria” nie całkiem tu pasuje, gdyż są to przede wszystkim komputerowe eksperymenty a sposób pracy matematyków komputerowych bardzo jest odmienny od tradycyjnego. Komputer jest cudownym narzędziem pozwalającym śledzić oddziaływania cząstek elementarnych, procesy wewnątrz jąder, zagadnienia chemiczne, biologii molekularnej, modelować procesy w komórkach, procesy ekonomiczne, socjologiczne, ekosystemy czy zjawiska astrofizyczne włącznie z Wielkim Wybuchem. W rozbite na wąskie specjalności nauki komputer wprowadza pewną jedność metody. Poziom techniczny, potrzebny do rozwiązania niektórych zagadnień, osiągnięty został całkiem niedawno, w wielu zaś przypadkach jest on w dalszym ciągu niezadowolający. Przewodniczący amerykańskiej organizacji rządowej „National Science Board”, Dr Roland Schmitt, określił obliczenia na superkomputerach mianem „trzeciej rewolucji naukowej”. W każdym razie jest to trzeci sposób poznawania świata, oprócz opisu teoretycznego i badań doświadczalnych. Wiedza zdobywana w szkole to podstawy, proste modele nie spotykane w codziennym życiu, np. omawia się ruch jednej planety wokół pojedynczej gwiazdy. Na uniwersytetach dodatkowo uwzględnienia się krzywiznę

czasoprzestrzeni, ale nadal zagadnienie rozwiązuje się tylko dla jednej planety, bez wzajemnego wpływu planet na siebie, gdyż uwzględnienie dodatkowych oddziaływań jest zbyt trudne. Opisuje się zjawiska falowe, ale nie takie, które zachodzą w basenie czy zatoce morskiej; oblicza się energie atomu wodoru ale nie energie atomów wieloelektronowych. Omawianie zagadnień modelowych jest niezbędne ale prawdziwe problemy są zawsze bardzo złożone i nie da się ich rozwiązać analitycznie, na kartce papieru, bo wymagają zbyt wielu skomplikowanych obliczeń.



Informatycy zajmują się ważnymi i ciekawymi zagadnieniami, zupełnie się jednak nie pokrywającymi z zainteresowaniami „komputerowców”, którzy szukają rozwiązań konkretnych problemów naukowych. W latach 80-tych nietrywialne zastosowania komputerów zaczęły się na dobre. Nauki komputerowe to gałęzie istniejących nauk i nowo powstające nauki interdyscyplinarne, które powstały i mogą istnieć tylko dzięki komputerom. Dostrzegają to wielcy twórcy nauki. Murray Gell-Mann (Nobel 1969 za teorię kwarków), przemawiając w czasie „Complex Systems Summer School” w Santa Fe, powiedział:

Transformacja społeczeństwa przez rewolucję naukową XIX i XX wieku zostanie wkrótce przyćmiona przez jeszcze dalej idące zmiany, wyrastające z naszych rosnących możliwości zrozumienia złożonych mechanizmów, które leżą w centrum zainteresowania człowieka. Bazą technologiczną tej nowej rewolucji będą niewyobrażalnie potężne komputery razem z narzędziami matematycznymi i eksperymentalnymi oraz oprogramowaniem, które jest niezbędne by zrozumieć układy złożone... Przykładami adaptujących się, złożonych systemów jest ewolucja biologiczna, uczenie się i procesy neuronalne, inteligentne komputery, chemia białek, znaczna część patologii i medycyny, zachowanie się ludzi i ekonomia.

Tego typu zastosowania znajdują się na peryferium zastosowań informatyków, jest to gałąź, która powoli zaczyna przerastać całe drzewo... Dokonam teraz krótkiego przeglądu możliwości zastosowań metod komputerowych w nauce.

1.2.1 Matematyka komputerowa

Większość matematyków uważa komputery za nieco bardziej inteligentne maszyny do pisania chociaż w kursie zastosowań matematyki, od wielu lat prowadzonym w Warszawie, niewiele już matematyki pozostało: naucza się na nim obsługi takich pakietów programowych jak bazy danych, arkusze kalkulacyjne czy podstaw systemów

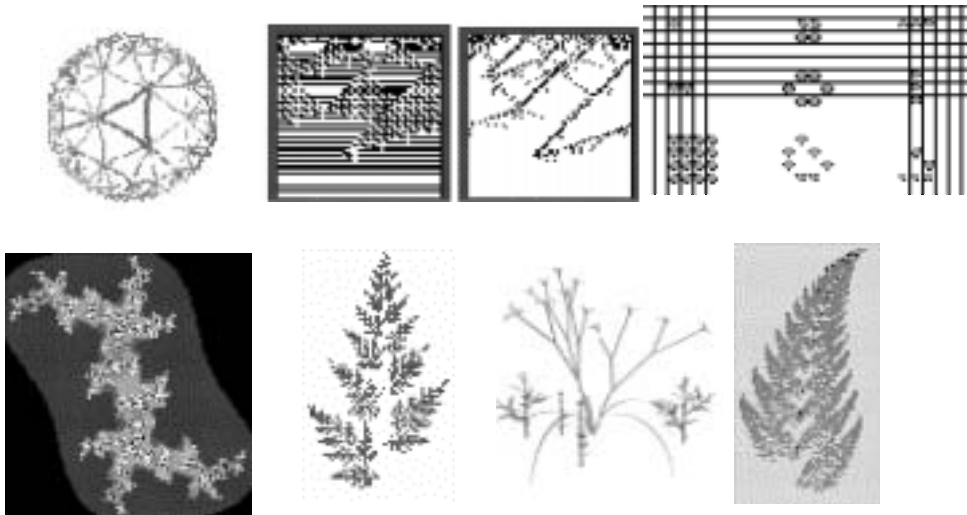
operacyjnych. Matematyka na poziomie politechnicznym, w znacznej mierze XIX-wieczna, daje się uprawiać przy pomocy programów komputerowych służących do manipulacji matematycznymi symbolami.

Są pewne działy matematyki, w których komputerów używa się od dawna, np. w teorii liczb naturalnych, gdzie pewne hipotezy, trudne do udowodnienia, sprawdza się szukając możliwych wyjątków przy pomocy komputerów. Interesujące są próby faktoryzacji różnych wielkich liczb pierwszych w oparciu o obliczanie rozproszone po całym świecie. Każdy komputer wykonuje niewielką część pracy (robi się to w nocy na PC i Mac-ach) i odsyła rano wyniki. Tak rozłożono na czynniki pierwsze 116-cyfrową liczbę i 155-cyfrową 9-tą liczbę Fermata, $2^{512}+1$. Największa znana liczba to $2^{756839}-1$, ma ona 227 832 cyfry (sprawdzono to na superkomputerze Cray w 1992 roku). Innym, bardzo znanym zagadnieniem matematycznym, w którym zastosowanie komputerów było dość istotne, była klasyfikacja grup prostych (grupa jest prosta jeśli nie ma nietrywialnych podgrup normalnych), pełniących podobną rolę w teorii grup jak liczby pierwsze w teorii liczb. Znana jest pełna lista tych grup, największa grupa-potwór liczy w przybliżeniu $8 \cdot 10^{53}$ elementów! Prace nad tą klasyfikacją trwały ponad ćwierć wieku; ocenia się, że pełny dowód, gdyby zebrać wszystkie cząstkowe prace, zabierze 5-10 tysięcy stron druku!

Żadna z tych prac nie kwalifikuje się jako matematyka komputerowa, gdyż komputery grały w nich jedynie rolę pomocniczą. Nieco bliżej matematyki komputerowej umieścić można dowód zagadnienia czterech barw:

Każdą mapę narysowaną na kartce papieru można pokolorować za pomocą czterech barw w taki sposób, że państwa mające wspólną granicę otrzymają różne kolory.

Twierdzenie to czekało na dowód od 1852 roku, a prace nad nim doprowadziły do znacznego rozwoju takich gałęzi matematyki jak teoria grafów. Kiedy wreszcie doczekało się dowodu okazał się on tak skomplikowany, iż nikt nie może go sprawdzić! Problem sprowadzono do zbadania bardzo dużej liczby grafów: w 1970 roku oceniano, że potrzeba na to 11 lat ciągłej pracy szybkiego komputera. Wydaje się jednak, że dowód o umiarkowanej długości nie istnieje i każdy dowód tego zagadnienia będzie zawierał niesprawdzalne (w sensie tradycyjnego dowodu na papierze) przez człowieka elementy. Dowód przeprowadzono przy użyciu trzech komputerów, w latach 1970-76, zużywając na obliczenia ponad 1000 godzin i prowadząc w istocie dialog z programem. Ponieważ w używanych komputerach pojawiały się czasem przekłamania istnieje pewne, choć bardzo



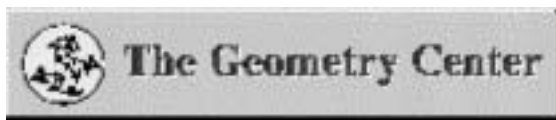
Fraktale, automaty komórkowe i L-systemy to niektóre przykłady matematyki komputerowej.

małe, prawdopodobieństwo, że dowód jest błędny (stąd późniejsze kontrowersje). Komentując swoje wyniki K. Appel i W. Haken stwierdzili²

Od czasów Platona do późnego średniowiecza metody matematyczne uważane były jako przewyższające metody eksperymentalne, a fizyka doświadczalna była nie do przyjęcia dla poważnych naukowców. To bardzo mocno zaważyło na rozwoju pewnych gałęzi fizyki. Na przykład, prawa swobodnego spadku ciał pod wpływem siły ciężkości zostały niepoprawnie wypowiedziane przez Arystotelesa, który spróbował wyprowadzić je w sposób teoretyczny, a błąd nie został poprawiony przez 2000 lat, dopóki proste obserwacje i doświadczenia Galileusza nie wyjaśniły kwestii i nie rozpoczęły szybkiego rozwoju dynamiki mechanicznej. Gdy tylko przekonano się o ważności badań eksperymentalnych (i gdy poznano jeszcze mocniejsze ograniczenia jakie się stosują do metod czysto matematycznych) osiągnięto bardzo owocny rozwój fizyki w wyniku łączenia tych dwóch metod. Zatem fakt, że nasze wyniki wskazują na nieco mocniejsze ograniczenia czysto matematycznych metod niż to chcieliby widzieć niektórzy matematycy, winien być interpretowany nie jako rezultat negatywny, ale jako wskazanie kierunku rozwoju.

²

Zagadnienie czterech barw, w: Matematyka współczesna, 12 esejów, WNT 1983



W ostatnich latach tego typu dowody zdarzają się często, powstał nawet *Geometry Supercomputer Project*, czyli „superkomputerowy projekt badań geometrii”, pod kierownictwem Alberta Mardena z Uniwersytetu Minnesoty, skupiający kilkunastu matematyków z całego świata. Clement Lam przeprowadził poszukiwania „skończonej płaszczyzny rzutowej 10 rzędu”, wymagające zbadania 100 bilionów (10^{14}) przypadków, używając kilku tysięcy godzin superkomputera Cray-1A. Jego konkluzja jest bardzo ciekawa³:

Wprowadzenie komputerów dostarczyło matematykom bardzo potężnego narzędzia pracy. Pozwala nam ono atakować problemy złożone. Niewielka niepewność, nieodłączna w dowodach wymagających komputerów nie powinna nas zniechęcać przed ich używaniem. Jak powiedział laureat nagrody Nobla w fizyce, Richard Feynman, zwracając się w 1955 roku do National Academy of Science, „Wiedza naukowa jest zbiorem stwierdzeń o różnym stopniu pewności - niektóre bardzo wątpliwe, inne prawie pewne, ale żadne nie są absolutnie pewne.” Tak jak fizycy nauczyli się żyć z niepewnością, tak i my powinniśmy nauczyć się żyć z „niepewnymi” dowodami. Może moglibyśmy sobie podkraść jeszcze jedną kartę z podręcznika eksperymentalnej fizyki - oni mają swoje akceleratory, czemu my nie mielibyśmy marzyć o matematycznych superkomputerach?

Nie tylko matematycy marzą o takich superkomputerach. W praktyce dowody przeprowadzone przy pomocy komputera są często znacznie pewniejsze niż dowody klasyczne, w których jest sporo błędów. Przykładem mogą być tablice całek, w których programy do algebry symbolicznej znajdowały od 10 do 25% błędnie podanych całek. Wielce ciekawy aspekt komputerowej matematyki to używanie metod sztucznej inteligencji w dowodzeniu twierdzeń i wysuwaniu hipotez matematycznych.

Prawdziwa matematyka komputerowa wiąże się z zagadnieniami obliczeniowymi, które trudno jest traktować analitycznie, np. wyrasta z zabaw Mandelbrota z komputerem i fraktalami (Benoit Mandelbrot od dawna nie uprawia już „klasycznej” matematyki), symulacji automatów komórkowych czy prób wyobrażenia sobie 3-wymiarowych przekrojów wielowymiarowych obiektów, możliwych tylko dzięki wizualizacji komputerowej. Jakie odkrycie polskiego matematyka zrobiło największą karierę? W 1946 roku Stanisław Ulam, układając w szpitalu pasjansa, wpadł na pomysł rachunku Monte Carlo. Jest to metoda uniwersalna, pod warunkiem, że mamy pod ręką narzędzie które

³

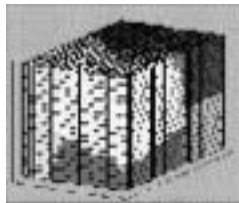
The Mathematical Intelligencer 12 (1990) 8

potrafi w ciągu sekundy zbadać miliony przypadkowych możliwości. Jedynie komputery dają nam takie możliwości.

Odrębnym zagadnieniem matematyki komputerowej jest kwestia automatycznego dowodzenia twierdzeń i sprawdzania poprawności już istniejących dowodów matematycznych. Wiąże się to z jednej strony z programem formalizacji matematyki a z drugiej strony ze sztuczną inteligencją. Na razie niewiele jest rezultatów osiągniętych przy pomocy komputerów, które można uznać za interesujące. W ramach japońskiego projektu komputerów piątek generacji udało się rozstrzygnąć pewne nowe zagadnienia matematyczne (udowodnić istnienie pewnych kwazigrup). Znacznie bardziej zaawansowane są prace nad automatycznym sprawdzaniem poprawności dowodów wykonanych przez ludzi. Niemała jest tu zasługa polskich matematyków od lat rozwijających system MISAR. Jednak najbardziej ambitnym, zakrojonym na długie lata projektem podjętym przez liczne grupy badawcze rozproszone po całym świecie jest projekt QED (od „Quod Erant Demonstratum”). Zmierza on do zbudowania komputerowego systemu w którym zgromadzona zostanie cała wiedza ludzkości o matematyce. Wszystkie dowody zostaną w pełni sformalizowane i sprawdzone w ścisły sposób.



Projekt ten podjęto z pięciu głównych powodów. Po pierwsze, stopień komplikacji wiedzy matematycznej stał się tak wielki, że żaden człowiek nie jest w stanie objąć jej istotnej części, np. zrozumieć szczegółów dowodów klasyfikacji grup prostych. Odszukanie w morzu twierdzeń matematycznych przydatnych rezultatów powinno być możliwe przy pomocy odpowiedniego oprogramowania. Po drugie, złożoność obecnie budowanych mikroprocesorów, samolotów czy robotów staje się tak wielka, że konieczne będzie budowanie ogromnie złożonych modeli matematycznych tych systemów. Struktury wiedzy w systemie QED powinny pozwolić na budowanie takich modeli. Trzecim powodem budowy takiego systemu jest jego potencjalnie wielka użyteczność w nauczaniu matematyki. Dzięki interakcyjnemu sprawdzaniu poprawności dowodów studenci będą mogli sami eksperymentować z różnymi dowodami. Czwartą motywacją rozwoju systemu są względy estetyczne i kulturowe. Jest to wysiłek porównywalny być może z budową

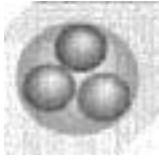


piramid, w który musi wziąć udział wielka liczba matematyków, lecz jego efektem będzie struktura niezrównanej piękności. Ostatnim argumentem jest chęć zabezpieczenia całego gmachu matematyki przed korupcją, przed niepewnością niektórych przypuszczeń i zapomnieniem niektórych gałęzi. Zamiast pytania: „czy jest dowód” będzie można zapytać: czy system QED to sprawdził? Pierwsza konferencja na temat tego projektu odbyła się w USA a druga w Warszawie przy końcu 1995 roku. Szczegółowe informacje o projekcie QED znaleźć można pod

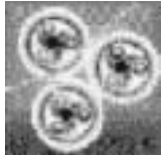
adresem:

<http://www.mcs.anl.gov/qed/>

1.2.2 Fizyka komputerowa



Dziedziną nauk komputerowych, która osiągnęła największą tożsamość, jest fizyka. Widać to choćby po tytułach czasopism fizycznych: *Computers in Physics*, *Computer Physics Communications*, *Computer Physics Reports*, *Journal of Computational Physics*, *EPS Computational Physics Group Newsletter*. We wrześniu 1990 na konferencji w Amsterdamie stwierdzono: „Fizyka komputerowa to fizyka teoretyczna studiowana metodami eksperymentalnymi”



Steven Wolfram, nazywany przez kolegów w Princeton następcą Einsteina, pracował nad teorią cząstek elementarnych i przyczynił się do rozwoju teorii automatów komórkowych. Jego największym sukcesem jest jednak system algebry symbolicznej **Mathematica**. Wolfram tak pisze o komputerach w fizyce⁴:

Lata 80 wspomiane będą jako dekada, w której komputery powszechnie wkroczyły do metod nauk fizycznych. W dawnych czasach fizyka była gałęzią filozofii. Za czasów Galileusza nastąpiła rewolucja i fizyka jest nauką doświadczalną. Teraz mamy kolejną rewolucję tj. symulowanie świata przez komputery. Jest to fundamentalna zmiana sposobu myślenia o nauce. Praca nad automatami komórkowymi i fraktalami pokazała, jak proste fizyczne modele prowadzą do niezwykle złożonych zachowań.

W nadchodzących latach zobaczymy związki pomiędzy pytaniami fizycznymi i teorią obliczeń, fizyka statystyczna pojawi się w projektowaniu systemów współbieżnych.

Inny fizyk, Geoffrey Fox, pracując na Politechnice Kalifornijskiej (CalTech w Pasadenie) stworzył jeden z pierwszych programów studiów w dziedzinie fizyki komputerowej. Program dotyczy fizyki komputerowej i programowania komputerów współbieżnych a znany jest pod nazwą C³ P (Caltech Concurrent Computation Program). Fox tak pisze o przyszłości tej dziedziny⁵:

Fizyka komputerowa jest trzecią siłą, ale nadal słabą. Dopiero stacje robocze o szybkościach gigaflopów i superkomputery o szybkościach teraflopów⁶ to zmieniają. Wielkim wyzwaniem dla uniwersytetów będzie kształcenie ludzi, którzy z takiego

⁴ *Computers in Physics* 7/1989

⁵ *Computers in Physics* 1/1990

⁶ **Gigaflop** to miliard operacji na sekundę a **teraflop** to bilion operacji na sekundę.

narzędzia będą umieli korzystać. Kto ma to robić? Informatycy czy fizycy? Może stosowani matematycy? Potrzebne są programy akademickie w zakresie złożonych systemów w fizyce, chemii czy biologii, programy interdyscyplinarne. Trudno by specjaliści np. w genetyce mogli nauczać zaawansowanych metod komputerowych już teraz niezbędnych w ich pracy. Przemiana świadomości jest znacznie trudniejszym zadaniem niż budowa Teraflopowego komputera.

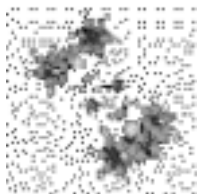
Problem kształcenia specjalistów w zakresie metod komputerowych pojawił się całkiem niedawno. Dopiero w 1994 roku odbyła się w Albuquerque, USA, pierwsza konferencja poświęcona programom nauczania takich zagadnień. W związku z tworzeniem się w Polsce pierwszych ośrodków superkomputerowych w 1993/94 roku problem ten pojawił się również i u nas i będzie dotyczy nie tylko fizyki ale i wielu innych dziedzin nauki. W 1994 roku w Polsce po raz pierwszy sprzedano więcej stacji roboczych i komputerów o większej mocy obliczeniowej instytucjom przemysłowym niż ośrodkom naukowym. Świadczy to nie tylko o niskich nakładach finansowych na naukę ale i o nowych tendencjach w przemyśle, które powinny się utrzymać. Nawet firmy produkujące opony samochodowe zakupiły superkomputery, by dorównać konkurencji światowej.

Najbardziej popularnym podejściem (w znacznej mierze wywodzącym się z prac laureata nagrody Nobla, Kennetha Wilsona) do rozwiązywania równań podstawowej teorii materii, jaką jest chromodynamika kwantowa, jest ich symulacja na sieciach czasoprzestrzennych (nazywa się to fachowo „QCD on lattices”). Chociaż osiągnięte obecnie moce komputerów są do tego celu daleko niewystarczające już teraz ponad 20% czasu superkomputerów amerykańskich przeznaczają się na takie właśnie obliczenia. Wilson twierdzi, że osiągnięcie dokładności doświadczalnych w tej dziedzinie wymaga znacznego rozwoju algorytmów i zwiększenia mocy komputerów o czynnik rzędu stu milionów! Pewne użyteczne rezultaty osiągnąć można będzie „już” przy pomocy maszyn o szybkości teraflopów - 40 fizyków z USA napisało wspólnie projekt budowy takiego komputera do celów QCD. Duża grupa naukowców zamierza więc zbudować szybki komputer do specjalnych obliczeń. Czy taki sposób działania - budowanie wyspecjalizowanego przyrządu - nie przypomina fizyki teoretycznej studiowanej metodami eksperymentalnymi?

Podobne próby w tej dziedzinie podjęto już wcześniej. Przez wiele lat naukowcy z firmy IBM budowali specjalny komputer GF-11 do obliczeń QCD i w maju 1994 roku, po roku nieprzerwanych obliczeń tego superkomputera, uzyskano pierwsze rezultaty. Udało się wyznaczyć stosunki mas dla 8 cząstek elementarnych z dokładnością do kilku procent. Wyniki te mają niezwykle ważne znaczenie poznawcze, gdyż równań QCD nie udało się rozwiązać analitycznie nawet dla zagadnienia czystej próżni i nie wiadomo było, czy naprawdę opisują one poprawnie cząstki elementarne.

Doświadczalna fizyka wysokich energii również nie może się obyć bez fizyków komputerowych. Oprogramowanie analizujące dane z akceleratorów liczy setki tysięcy

wierszy, w ciągu sekund napływają gigabajty danych⁷. Inni „pożeracze” mocy obliczeniowych to fizycy symulujący zachowanie plazmy w różnych warunkach eksperymentalnych, fizycy atmosfery symulujący zachowanie dziury ozonowej, zanieczyszczenia atmosferyczne, astrofizycy badający ewolucję gwiazd, ewolucję wszechświata i galaktyk a specjaliści od mechaniki kwantowej własności ciał stałych takich jak materiały półprzewodnikowe, własności pojedynczych cząsteczek i atomów. Symulacje w zakresie fizyki jądrowej odgrywają niezwykle ważną rolę nie tylko w technice wojskowej - można przy ich pomocy wyeliminować czasami potrzeby próbných wybuchów jądrowych - lecz również w rozwoju energetyki jądrowej. Fizycy i astronomowie zużywają ponad połowę czasu superkomputerów w ośrodkach akademickich.



Z praktycznego punktu widzenia niezwykle ważne są symulacje meteorologiczne. Dlaczego mamy kiepskie prognozy pogody? Nasze prognozy są niedokładne bo nie są oparte na obliczeniach. Nawet Indie zakupiły do tego celu superkomputer! Sprawdzalność dobowych prognoz pogody podawanych przez Instytut Meteorologii i Gospodarki Wodnej wynosi 81-85%. Opierają się one na wielu danych dochodzących do dyżurnego synoptyka z krajowych obserwatoriów meteorologicznych, mapkach prognostycznych z krajów ościennych, zdjęciach satelitarnych. Ostateczna prognoza oparta jest jednak nie na symulacji zachowania się atmosfery, ale na doświadczeniu i intuicji synoptyka. Trafność prognoz krótkoterminowych opartych na założeniu, że tak jak jest dzisiaj będzie i jutro wynosi 65% a trafność prognoz opartych na symulacjach atmosfery sięga 97%. Niewielki błąd wynika z braku precyzyjnych danych o stanie atmosfery w danym momencie, a więc z niepewności warunków początkowych. Niestety, przewidywania pogody w okresie czasu dłuższym niż kilka dni są niewykonalne, gdyż charakter równań opisujących atmosferę jest taki, iż nawet bardzo mały błąd w warunkach początkowych bardzo szybko rośnie prowadząc do wielkiej niepewności wyników. Liczba superkomputerów pracujących prawie wyłącznie dla potrzeb meteorologii jest na świecie duża.

Fizyka komputerowa korzysta ze specyficznych metod numerycznych, wprowadza również analogie fizyczne, by znaleźć nowe podejścia do rozwiązywania zagadnień innych nauk komputerowych. Przykładem jest tu metoda stopniowego studzenia (simulated annealing) lub metoda minimalnych powierzchni zastosowana w różnych zagadnieniach optymalizacyjnych.

⁷

por. Computers in Physics, vol. 6, no.1 (1992)

1.2.3 Chemia komputerowa

Po etapie intensywnego rozwoju teorii i oprogramowania w latach sześćdziesiątych i siedemdziesiątych chemia kwantowa (dział chemii i fizyki teoretycznej zajmujący się matematycznym opisem cząsteczek chemicznych) stała się w latach 80-tych przydatna chemikom doświadczalnikiem. Obecnie znacznie jest więcej użytkowników wielkich pakietów programów kwantowochemicznych niż ludzi znających się na metodach wykorzystywanych w tych pakietach. Osiągane dokładności obliczeń własności małych, kilkuatomowych cząsteczek są często na poziomie danych doświadczalnych a można je uzyskać znacznie szybciej. Chemicy komputerowi znacznie lepiej znają się na programach i komputerach niż na robieniu doświadczeń czy rozwijaniu teorii. Istnieją nie tylko pisma dotyczące chemii kwantowej lecz również *Journal of Computational Chemistry*, czyli pismo chemii obliczeniowej. W sieci komputerowej Internet działa od lat grupa dyskusyjna chemii komputerowej (computational chemistry group).

W powszechnym użytku jest kilkanaście uniwersalnych systemów programów dla obliczeń kwantowochemicznych, każdy z nich liczący setki tysięcy wierszy programu⁸. Opracowanie takich programów zajmuje dziesiątki osobołat (np. system „Columbus” ocenia się na 25 osobołat). Pewną próbą zebrania wielu z tych pakietów i stworzenia uniwersalnego systemu do „globalnej symulacji świata” jest MOTECC (Modern Techniques in Computational Chemistry), projekt podjęty przez Enrico Clementiego w firmie IBM. Idea globalnej symulacji jest prosta: seria modeli opartych na fundamentalnych równaniach opisujących materię, od modeli mikroskopowych do makroskopowych, korzystających z wyników pośrednich modeli na kolejnych poziomach powinna nam pozwolić wyznaczyć wszystkie własności materii (począwszy od własności atomów) z komputerowych obliczeń. Z technicznego punktu widzenia są to bardzo wyrafinowane obliczenia. System MOTECC 90 składa się z 30 dużych pakietów programów: od mechaniki kwantowej przez mechanikę statystyczną do dynamiki ośrodków ciągłych. Niestety trudności finansowe firmy IBM na początku lat 90-tych odbiły się bardzo niekorzystnie na finansowaniu badań nie mających szybkiego praktycznego zastosowania.

Przykładem symulacji globalnej są obliczenia dla wody. Zaczynamy od atomów, składających się na cząsteczkę wody: dwa wodory i jeden tlen, czyli 3 jądra atomowe i 10 elektronów. Wstępne wyniki dają obliczenia metodą pola samouzgodnionego, następnie dodajemy korelację elektronową, obliczamy hiperpowierzchnie energetyczne i analizujemy wibracje cząsteczki wody. Przechodzimy teraz do badania oddziaływań 2 cząsteczek (tj. dimeru) wody, obliczamy ich potencjały oddziaływania, dodajemy więcej cząsteczek badając małe klastry wody i poprawiając potencjały oddziaływania. Następnym etapem jest mechanika statystyczna: używając metod dynamiki molekularnej lub prostej metody Monte Carlo symulujemy 10^3 - 10^5 cząsteczek w ustalonej objętości i

⁸ Są to programy pisane prawie wyłącznie w języku Fortran

temperaturze, generujemy miliony konfiguracji obliczając energię całego układu używając potencjałów wyliczonych poprzednio, liczymy średnie statystyczne i własności układu, takie jak funkcje korelacji par, przekroje na rozpraszanie promieni X i rozpraszanie neutronów,



translacyjno-rotacyjne funkcje korelacji, widma podczerwone i Ramana, współczynniki dyfuzji, czasy relaksacji NMR, dane termodynamiczne. Dla 100 000 cząsteczek w systemie otwartym

i czasach rzędu piko do nanosekund (bilionowych do miliardowych części sekundy) daje się dość dokładnie obserwować turbulencję i dostać prawidłową liczbę Reynoldsa, opisującą hydrodynamiczne zachowanie płynów. Z tak obliczonych makroskopowych własności można korzystać przy rozwiązywaniu równań dynamiki płynów i badaniu np. ruchu fal w zatoce.

Kilka prostych kresek na początku i mutacje genetyczne stanowiące podstawę ewolucji już po kilkudziesięciu pokoleniach prowadzą do niezwykle zróżnicowanych kształtów, zadziwiająco podobnych do form biologicznych.

Globalna symulacja wymaga dużych mocy obliczeniowych (obliczenia dla wody prowadzone były na specjalnie skonstruowanym systemie, składającym się z wielu współpracujących ze sobą superkomputerów), ale do końca tej dekady takie obliczenia będą możliwe przy pomocy komputera stojącego na biurku. Bardzo intensywnie rozwijaną częścią chemii obliczeniowej jest **modelowanie molekularne**, wymagające szybkich stacji graficznych i programów dynamiki molekularnej. Wkracza to już w chemię białek, należąca do biochemii. Określenie struktury białka metodami krystalograficznymi jest bardzo trudne, wymaga czasami lat pracy i kosztuje około miliona dolarów. Jednocześnie całkiem tanio wyprodukować można białka o zadanej sekwencji aminokwasów. Określenie własności białek na podstawie znajomości sekwencji tworzących je aminokwasów jest trudne, ale do pewnego stopnia możliwe dzięki metodom komputerowym. Jedną z gałęzi biochemii jest **farmakologia kwantowa** zajmująca się zastosowaniem metod komputerowych do projektowania leków. Opracowywanie nowych leków jest niezwykle kosztowne - na 12.000 testowanych substancji chemicznych zaledwie jedna ma szansę na pojawienie się na półkach aptek. Przewidywanie własności takich substancji przy pomocy komputerów pozwala na znaczne oszczędności, stąd firmy farmakologiczne od końca lat osiemdziesiątych zaczęły kupować komputery i zatrudniać chemików komputerowych. Ogromnie dużo informacji na temat zastosowania komputerów w farmakologii znaleźć można pod adresem <http://www.cpb.uokhsc.edu/pharmacy/pharmint.html>. Różne aspekty zastosowania komputerów w chemii można znaleźć w serwerach umieszczonych na liście <http://www.chem.ucla.edu/chempointers.html>

1.2.4 **Biologia i biocybernetyka komputerowa**



Chociaż komputery wykorzystuje się w wielu działach biologii, do gromadzenia informacji i wspomagania eksperymentów to biologia komputerowa dopiero się zaczyna. Z jednej strony mamy symulacje na poziomie makroskopowym, ekologiczne, związane z przepływem substancji i energii w przyrodzie oraz z biologią populacyjną, z drugiej strony mamy symulacje na poziomie molekularnym, w genetyce i biologii molekularnej, w szczególności problemy powstania życia i kodu genetycznego. Jednym z najważniejszych zagadnień jest określenie struktury przestrzennej białek na podstawie sekwencji aminokwasów. Najczęściej próbuje się stosować różne metody heurystyczne, statystyczne lub detektory regularności oparte na modelach sieci neuronowych.

Jeden z najciekawszych projektów **modelowania ekologicznego** opublikowano w 1986 roku⁹. Dotyczył on plemienia Turkana, nomadów zamieszkujących północnozachodnią Kenię. Plemię to żyje większość czasu na krawędzi głodu. W początkach lat 80-tych *Narodowe Laboratorium Zasobów Ekologicznych* (National Resource Ecology Laboratory) z Uniwersytetu Stanu Colorado rozpoczęło szczegółowe badanie warunków życia i środowiska plemienia Turkana. Przeprowadzono bardzo szczegółową analizę z punktu widzenia przepływu energii w przyrodzie, począwszy od energii słonecznej, przez roślinną, produkty zwierzęce i energię zużywaną na pracę ludzi. Do oceny strat energii potrzebny był bardzo dokładny model, uwzględniający sprawność fotosyntezy, metabolizmu, produkcji białek, czynniki środowiskowe takie jak ilość opadów w różnych okresach roku, wilgotność gleby itp. Informacje, potrzebne do stworzenia takiego szczegółowego modelu zbierano latami. By określić, ile i jakiego pożywienia zjadają różne gatunki zwierząt obserwowano je licząc ile razy w ciągu dnia dane zwierzę coś gryzło. Na całym obszarze, po którym wędruje plemię Turkana policzono liczbę różnych roślin posługując się analizą zdjęć lotniczych. Oceniono jakość i ilość drewna, zużywanego na opał. Uwzględniono różne sposoby zdobywania pożywienia w okresach suszy: gdy brakuje żywności a zwierzęta z powodu głodu dają mało mleka ludzie piją ich krew, wymieniają je na zboże lub zjadają. W poprzednich modelach mówiło się o „łańcuchu pokarmowym”; w tych badaniach była raczej „pajęczyna” lub „sieć” pokarmowa, gdyż zidentyfikowano ponad 30 alternatywnych dróg zdobywania pożywienia. Tak szczegółowy model pozwala na śledzenie i realistyczne przewidywanie reakcji plemienia Turkana na zmiany klimatyczne, np. długotrwałe okresy suszy.

⁹ D.L. Cappock et.al., Journal of Applied Ecology (1986)

Informacje na temat bioróżnorodności, ekologii i zastosowania metod komputerowych znaleźć można pod adresem:

<http://golgi.harvard.edu/biopages/biodiversity.html>



Bardzo ciekawe zagadnienia informatyczne związane są z projektem mapowania ludzkiego genomu. Specjaliści od baz danych oceniają, że problemy związane z sekwencjonowaniem genomów, wymagające zapisu informacji o miliardach par nukleotydów to jedno z największych wyzwań, przed którymi stanęli. Ocenia się, że w 1994 roku szybkość sekwencjonowania wynosiła około 160 milionów par rocznie, a w 1999 roku

wzrosła do 1.6 miliarda (genom człowieka zawiera prawie 3 miliardy par nukleotydów, w około 100.000 genów). Na początku lat dziewięćdziesiątych genom nawet tak dobrze zbadanego organizmu jak *Escherichia coli* zmapowany został zaledwie w 1/6. Coraz więcej biologów zamiast prowadzić badania doświadczalne korzystać będzie z baz danych i symulacji komputerowych w biologii i w medycynie. Przykładem takiego problemu obliczeniowego, gdy już powstanie odpowiednia baza danych, będzie próba rekonstrukcji drzewa ewolucji. Przy okazji dyskusji nad tym projektem¹⁰ stwierdzono, że bardzo brakuje komputerowych biologów. Informatycy nie znają języka biologów i nie potrafią się z nimi porozumieć a biolodzy mają trudności z opanowaniem technik matematycznych i komputerowych potrzebnych do pracy w komputerowej biologii. Biologia molekularna bliska jest biochemii komputerowej ale ma swoje bardzo specyficzne metody, wywodzące się nie tylko z fizyki czy chemii kwantowej ale i z typowo informatycznych zagadnień, np. teorii języków sztucznych. Więcej informacji o komputerach w genetyce znaleźć można pod adresem:

http://www.ornl.gov/TechResources/Human_Genome/genetics.html

Na jednej ze sponsorowanych przez firmę IBM konferencji „Modelowanie i Analiza w Wielkiej Skali” (Large Scale Analysis and Modeling) nagrodzono 2 prace biologiczne: pierwsza dotyczyła trójwymiarowej rekonstrukcji graficznej ucha środkowego, a druga modelowania pracy serca i analizy elektrokardiogramu. Symulacje działania różnych narządów dla potrzeb biologii i medycyny są coraz częstsze. Ważnym działem biologii komputerowej są symulacje działania komórek nerwowych jak i fragmentów układu nerwowego, a w szczególności mózgu. Biologia komputerowa przechodzi tu w **komputerową medycynę**. Należą do niej symulacje działania organizmu ludzkiego i zwierzęcego na różnym poziomie, począwszy od symulacji całych narządów do symulacji cykli biochemicznych pozwalających śledzić na komputerowym modelu co dzieje się w organizmie z podawanym lekiem czy innymi substancjami. Lekarze często nie pamiętają o różnych efektach ubocznych, zwłaszcza gdy są one wywołane kombinacją kilku leków. Dobrze skonstruowany model pozwala na wykrycie potencjalnych problemów zanim jeszcze powstaną. Niestety, zbyt mało jeszcze wiadomo o procesach zachodzących w

¹⁰

Communications of the ACM 11 (1991) 45

żywych organizmach by można było tworzyć dobre modele opisujące szerszą klasę zjawisk. Jest to wyzwanie zarówno dla biologów i medyków prowadzących badania doświadczalnie jak i dla komputerowych biologów i medyków tworzących coraz bardziej złożone modele. Zainteresowanym komputerami w medycynie polecam serwer WWW: <http://golgi.harvard.edu/biopages/medicine.html>

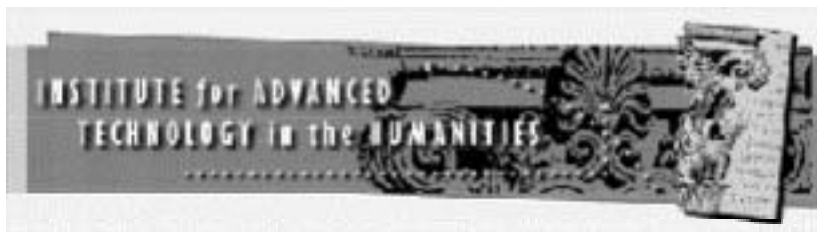
Przykładem wpływu biologii na metody komputerowego rozwiązywania problemów są algorytmy genetyczne, pozwalające na znalezienie optymalnych rozwiązań dla szerokiej klasy zagadnień, oparte na analogii z genetycznymi mechanizmami selekcji naturalnej. Obszerne informacje na ten temat znaleźć można pod adresem <http://gal4.ge.uiuc.edu/>

1.2.5 Nauki o poznaniu (cognitive sciences)

W ostatnich dziesięcioleciach rozwinęła się z połączenia psychologii, lingwistyki, filozofii, neurobiologii i sztucznej inteligencji nowa gałąź nauki, określana jako „nauki o poznaniu” lub „nauki kognitywne”, stawiająca sobie za cel zrozumienie, w jaki sposób człowiek postrzega i poznaje świat, w jaki sposób reprezentowana jest w naszym umyśle informacja kształtująca nasz obraz świata. Nauki o poznaniu stosują metody psychologii eksperymentalnej, korzystają z wyników badań nad mózgiem, ale ich jądro to modele oparte na symulacjach komputerowych. Jest bowiem rzeczą coraz bardziej oczywistą, że esencją działania mózgu jest wykonywanie obliczeń i tworzenie wewnętrznego modelu środowiska¹¹.

Wiele prac prowadzonych w dziedzinie nauk o poznaniu związana jest z modelowaniem komputerowym różnych funkcji poznawczych. Powstały bardzo szczegółowe modele fragmentaryczne i kilka bardzo ambitnych modeli komputerowych mających za zadanie odpowiedzieć na wiele pytań empirycznych dotyczących natury ludzkiego umysłu, szybkości świadomych reakcji, skojarzeń i sposobu rozumowania. Wydaje się obecnie, że realizacja „zdrowego rozsądku” przy pomocy systemu opartego na regułach i heurystykach wymaga bardzo wielkiej bazy wiedzy. Grupa informatyków amerykańskich pod kierownictwem Douglasa Lenata, już od 1984 roku pracuje w firmie MCC nad takim właśnie projektem, znanym pod kryptonimem **CYC** (jest to fragment słowa *enCYClopedia*). Według jego oceny uniwersalny system ekspertowy, z którym można będzie o prawie wszystkim porozmawiać, wymaga bazy wiedzy zawierającej ogromną liczbę reguł, rzędu 100 milionów! W systemie CYC znalazło się już wiele milionów reguł. Tak ogromna baza wiedzy pozwala uniknąć konieczności „głębokiego” rozumowania, to jest kolejnego stosowania wielu różnych reguł, powodującego zagubienie się programu w zbyt wielu możliwych wnioskach. Program posiadający szeroką wiedzę albo coś wie albo nie, wyprowadzając tylko proste wnioski. Człowiek posiada bardzo szeroką wiedzę, trudną do zwerbalizowania i zapisania w postaci reguł.

¹¹ <http://www.cog.brown.edu/pointers/cognitive.html> zawiera doskonale zorganizowane informacje o dostępnych w Internecie zasobach w dziedzinie nauk o poznaniu.



Powstają już instytuty zaawansowanej technologii dla humanistów. Jeden z nich ma swój serwer WWW pod adresem <http://jefferson.village.virginia.edu/home.html>

Projekt CYC jest bardzo kosztowny i ryzykowny, gdyż nikt nie wie, czy tak wielki system będzie działał stabilnie i czy będzie naprawdę przydatny. Douglas Lenat ma nadzieję, że tworzona przez niego baza wiedzy będzie niezbędną częścią wyposażonego w zdrowy rozsądek komputera XXI wieku. Rozpoczynając ten projekt autorzy ocenili szansę na dużą przydatność swojej bazy wiedzy w inteligentnych programach zaledwie na 5%, jednak po pięciu latach pracy dają sobie już 60% szans. Między innymi taka baza wiedzy powinna umożliwić swobodny dialog z komputerem w języku naturalnym i w takich systemach już jest wykorzystywana. Niestety, przystosowanie jej do potrzeb języka polskiego nie będzie bynajmniej proste.

1.2.6 Ekonomia komputerowa

Za modelowanie matematyczne zagadnień ekonomicznych przyznano już kilka nagród Nobla, pierwsze modele pochodzą jeszcze sprzed ery komputerów, potrzebne są jednak realistyczne modele w makroskali lub w skali całego globu. Niektóre modele ekonometryczne pozwalają na dość dokładne przewidywania sytuacji ekonomicznej w wybranych dziedzinach na rok z góry. Bogate kraje zyskują na możliwości przewidywania różnych tendencji, chociaż takie czynniki zewnętrzne jak np. pogoda czy konflikty regionalne są nie do przewidzenia.

Jednym z pierwszych rezultatów komputerowego modelowania ekonomii był apokaliptyczny raport Klubu Rzymskiego „Ograniczenia Wzrostu” z 1972 roku. Był to model skrajnie prosty, uwzględniający oddziaływanie pomiędzy zaledwie 5 zmiennymi: zaludnieniem, wzrostem ekonomicznym, produkcją żywności, skażeniami i zasobami naturalnymi. Obecnie rozwiązuje się ponad 10 000 sprzężonych równań różniczkowych a do globalnych symulacji ekonomii świata potrzeba dokładniejszych modeli o ponad 100 000 równań - wymaga to ogromnych pamięci i szybkości komputerów (rzędu gigabajtów



Na mapie pod adresem <http://www.uib.no/nsd/diverse/utenland.htm> umieszczone zostały połączenia z archiwami przechowującymi informacje interesujące dla historyków i socjologów.

i gigaflopów). Modelowanie rynku finansowego, przewidywanie ruchów cen na giełdzie, to najszybciej rozwijające się zastosowania superkomputerów. Są to zagadnienia bardzo trudne ale możliwe zyski rozpalają również wyobraźnię naukowców. Amatorom pozostają prostsze, ale nie całkiem prymitywne, modele ekonomiczne i symulacje globalnej ekonomii świata, działające na komputerach osobistych i sprzedawane razem z książką rozważającą możliwe scenariusze rozwoju świata. Niektóre z symulacyjnych gier komputerowych działają w oparciu o proste modele ekonometryczne.

Przerażenie budzą wypowiedzi polityków na tematy ekonomiczne. W tej dziedzinie każdemu wydaje się, że ma receptę uzdrawiającą gospodarkę. Symulacje ekonomiczne nie zawsze dają poprawne odpowiedzi, gdyż zbyt wiele jest czynników chaotycznych, takich jak wydarzenia polityczne na świecie, wojny czy wielkie klęski żywiołowe, których nie można uwzględnić w modelach. Jeśli jednak pominąć takie czynniki szczegółowe symulacje w tej dziedzinie są możliwe i nietrudno przewidzieć, jaki będzie efekt dodrukowania dużej ilości pieniędzy, jak to zalecają partie populistyczne.

Informacje dotyczące zagadnień finansowych zebrane są na stronie WWW:

<http://www.cob.ohio-state.edu/dept/fin/overview.htm>

a informacje ogólne dotyczące ekonomii na stronach:

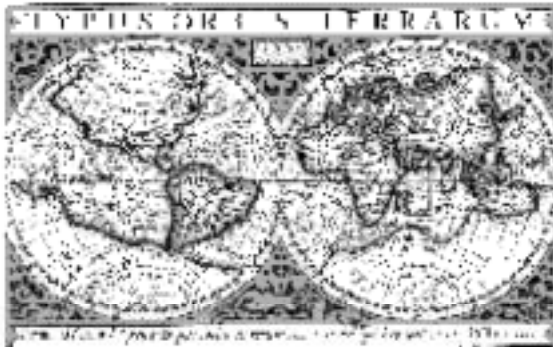
<http://www.yahoo.com/Business/>

<http://www.ame.cornell.edu/econinfo.html>

1.2.7 Nauki humanistyczne

Do czego przydają się komputery w naukach humanistycznych i czy jest sens mówić o ich komputerowych wersjach? Komputery nadają się oczywiście do prac pomocniczych, pisania, przechowywania i wyszukiwania danych, mogą służyć jako podręczne notatniki w pracy archeologa, ale nie wystarcza to jeszcze do określenia „komputerowa humanistyka”. Lingwistyka komputerowa jest częścią informatyki, podobnie próby tworzenia systemów prowadzących dialog w języku naturalnym czy tłumaczących z jednego języka na drugi. Lingwiści mogą się oczywiście wciągnąć w takie zagadnienia¹². Wiele ambitnych zastosowań i modeli komputerowych tworzy się dla potrzeb socjologii. Pojawiają się jednak inne zastosowania, bliższe temu, co można określić jako „komputerowa humanistyka.”

Klio, muza historii, udzieliła swojego imienia nauce o nazwie **kliometria**, zajmującej się ilościowymi (statystycznymi) metodami w historii. Przodują w tej dziedzinie Amerykanie: już w połowie lat siedemdziesiątych komputerowa baza danych, zwana „Archiwum historyczne”, zawierała prawie gigabajt informacji (miliard znaków), włączając w to wszystkie dane dotyczące wyborów na różnych szczeblach od 1790 roku. Dane, zawarte w tej bazie, pozwoliły na nowe spojrzenie na wiele aspektów amerykańskiej historii, np. na „radikalną reinterpretację sprawy niewolnictwa” dzięki przebadaniu aspektów



ekonomicznych utrzymywania tego systemu (do końca lat 60-tych panował pogląd, że system ten przed wojną domową był bliski załamania). Amerykańscy studenci mają do dyspozycji stworzony w 1987 roku program o nazwie „Wielka Maszyna Amerykańskiej Historii”, przedstawiający różne aspekty demograficzne i ekonomiczne w zadanym okresie czasu bezpośrednio na mapach

kartograficznych i wykresach dla podanych obszarów. Można na przykład porównać ze sobą sytuację ludności pochodzenia polskiego i irlandzkiego w 1870 roku w stanie Illinois. Wkrótce doczekamy czasów, gdy wszystkie informacje historyczne będą natychmiast dostępne badaczom i zamiast szperać po starych dokumentach będą spędzali całe dnie przed monitorem. The Association for History and Computing (AHC), czyli stowarzyszenie do spraw historii i komputerów, ma swoją stronę informacyjną pod adresem:

¹² Association for Computational Linguistics, czyli Stowarzyszenie Lingwistyki Komputerowej, udostępnia informacje z tej dziedziny pod adresem <http://www.cs.columbia.edu/~acl/HOME.HTML>

<http://grid.let.rug.nl/ahc/welcome.html>

Stowarzyszenie to wydaje od 1989 roku pismo „History and Computing”. Również w Polsce zainteresowanie tworzeniem elektronicznych archiwów dla potrzeb historyków jest bardzo duże. Polskie Towarzystwo Historyczne powołało Komisję Metod Komputerowych, która jest jednocześnie Polskim Oddziałem AHC. Choć programy ułatwiające dostęp i wspomagające analizę dokumentów historycznych można uznać za „komputerową historię” nie są to zastosowania komputerów wymagające szczególnie wyrafinowanych metod. Tworzenie oprogramowania do wspomaganie określania dat historycznych i dat wydarzeń astronomicznych to również stosunkowo prosta sprawa. Na tym jednak zastosowanie komputerów w historii się nie kończy.

Minęło niedawno 500-lecie odkrycia Ameryki. 11 października 1492 roku, o godzinie 10 wieczorem, ze statku Krzysztofa Kolumba dostrzeżono przy świetle księżycowym zarys brzegu. Fala była duża i trzeba było czekać do drugiej w nocy zanim brzeg nie pojawił się wyraźniej. Po 33 dniach podróży statki Kolumba dotarły do lądu. Ale jakiego lądu? Wiadomo, że była to wyspa należąca do archipelagu wysp Bahama. Kolumb nazwał ją San Salvador. Niestety, jej dokładne położenie pozostało przez wieki nie znane. Oryginalny dziennik okrętowy Kolumba zaginął, pozostała jedynie kopia, pełna błędów i niedomówień. Historycy, usiłujący na tej podstawie odtworzyć miejsce lądowania Kolumba w Nowym Świecie proponowali ni mniej ni więcej tylko 10 różnych wysp! Po latach studiów i analiz Samuel Morison, biograf Kolumba, ogłosił w 1942 roku, że była to z pewnością Wyspa Watlinga. Niektóre atlasy zaznaczają nawet przy nazwie tej wyspy, że na niej właśnie wylądował Kolumb. Ale czy na pewno?

Dyskusja nad miejscem lądowania Kolumba rozgorzała na nowo w 1980 roku, po powtórnym opublikowaniu pracy Pietera Verhooga argumentującego, że była to Wyspa Caicos. Dyskusja przyciągnęła uwagę Josepha Judge'a, jednego z wydawców National Geographic. Przeczytał on uważnie kopię dziennika Kolumba i doszedł do wniosku, że dane, dotyczące odległości, kierunków rejsu, wiatru i prądów morskich powinny wystarczyć do rozstrzygnięcia sporu. Dane te nigdy nie były przez historyków brane pod uwagę, gdyż trudno jest wykonać na papierze wszystkie obliczenia celem odtworzenia całej trasy, zwłaszcza, że dane zawierały błędy. Od czego jednak są komputery? Judge poprosił swojego przyjaciela, Louisa Mardena, znanego żeglarza, o wykreślenie trasy podróży odkrywców. Marden, do spółki ze swoją żoną, wykonał mnóstwo obliczeń, biorąc pod uwagę nie tylko dane z dziennika statku *Santa Maria*, lecz również prądy oceaniczne i dryf statku. Na końcu trasy znalazł wyspę Samana Cay, niewielką wysepkę położoną 100 kilometrów na południe od Wyspy Watlinga.

Kolejnym krokiem było stworzenie komputerowego modelu brzegów wysp Bahama. Potrzebny był do tego doświadczony programista, znający się nie tylko na grafice

komputerowej ale i na kartografii. Po dłuższych poszukiwaniach Judge natrafił na Roberta Lillestranda, wiceprezydenta znanej firmy komputerowej CDC, weterana dwóch ekspedycji na biegun północny. Lillestrand używał poprzednio komputera do odtworzenia trasy wyprawy z 1909 roku Roberta Peary na biegun północny. Tym razem sprawa była nieco bardziej skomplikowana.

W ten sposób w 1986 roku Lillestrand i współpracujący z nim programiści zabrali się za ostateczne wyjaśnienie sprawy. Opracowali oni program komputerowy, znany pod nazwą CRT (Columbus Research Tool), będący narzędziem, pozwalającym na elektroniczne odtworzenie wyprawy Kolumba. Do bazy danych wprowadzono informacje o kształtach brzegów 342 wysp, o płycznach i o niebezpiecznych wodach w okolicy wysp Bahama. Celem było nie tylko zbadanie, na której wyspie wylądował po raz pierwszy Kolumb, ale i odtworzenie pełnej trasy jego podróży. Uwzględniono przy tym takie informacje, jak wysokość drzew, porastających brzegi wysp. W zapiskach Kolumba można było bowiem odczytać, że Indianie używali łodzi, w których mieściło się 40 ludzi, a stąd wyliczyć można wysokość drzew, z których pni łodzie zrobiono. Gniazdo bocianie na *Santa Maria* znajdowało się na wysokości 20 metrów, można więc było obliczyć, z jakiej odległości widoczne będą drzewa na wyspach. Okazało się, że Marden miał rację: Kolumb dobił do brzegu Samana Cay.

Posługując się informacjami o geografii regionu można było, krok po kroku, badać wszystkie prawdopodobne trasy jego podróży i porównywać obraz, widziany na ekranie komputera z dziennikiem Kolumba. Dzięki temu odtworzono pełną trasę tej historycznej podróży.

Nauką, w której komputery przydają się nie tylko do zbierania i przechowywania danych, jest również **archeologia**. Rekonstrukcje znanych zabytków, np. kompleksu Borobodur na Jawie, nie byłyby możliwe bez możliwości wyszukiwania pasujących do siebie fragmentów kamiennych przy pomocy komputera. W archeologii często wykorzystuje się komputery przenośne, nadające się do prac w terenie. Tworzy się również graficzne bazy danych obiektów archeologicznych rozproszonych po wielu muzeach. Wydanie CD-ROMu jest znacznie mniej kosztowne niż wydanie wydawnictwa książkowego typu albumowego, a ilość informacji na nim zawarta może dorównywać 20-tomowej encyklopedii.

Gigantyczne kamienne posągi na Wyspie Wielkanocnej do dziś dnia stanowią zagadkę. Skąd się tam wzięły? Po sporządzeniu trójwymiarowego modelu wyspy grupa badaczy z Uniwersytetu Kalifornijskiego w Los Angeles przeprowadziła szereg obliczeń szukając takiej drogi, która wymagała najmniejszego nakładu wysiłku przy transporcie kamieni z gór na wybrzeże. Okazało się, że najłatwiejsza droga biegnie ze zboczy wygasłego wulkanu Rano Raraku. Posągi przewożono na drewnianych wałkach, w pozycji horyzontalnej. Grupa około stu robotników potrzebowała około tygodnia na transport. Te

śmiało przypuszczenia udało się potwierdzić przez tradycyjne badania archeologiczne - wzdłuż wytyczonej dzięki symulacjom komputerowym drogi znaleziono nie zauważone przez pokolenia badaczy fragmenty posągów.

W archeologii używa się również zaawansowanych technik komputerowej analizy obrazu, wykorzystując do tego celu zdjęcia lotnicze i zdjęcia z amerykańskiego satelity LANDSAT i francuskiego satelity SPOT. Systemy Informacji Geograficznej (GIS) gromadzą informacje z różnych źródeł, takich jak zdjęcia i wykopaliska archeologiczne. Badacze mają nadzieję, że można będzie na podstawie informacji zawartych w tym systemie określać potencjalne miejsca wykopalisk nie ruszając się sprzed monitora.

Rozwiniętą gałęzią nauki jest **geografia komputerowa**. Modele matematyczne w geografii społecznej stosuje się od dawna a liczba danych, które należy uwzględnić powoduje, że jedynie komputerowe opracowywanie danych ma sens. Ważnym zagadnieniem jest również wizualizacja danych zgromadzonych w systemach informacji geograficznej (GIS) w postaci map cyfrowych. Bez komputerów nie może się również obyć **kartografia**. Opracowanie nowej mapy czy planu na podstawie już istniejących map cyfrowych jest znacznie prostszym zagadnieniem niż opracowywanie mapy tradycyjnymi metodami. Komputer może przy tym przedstawiać te same dane w różnych rzutach kartograficznych i dowolnie przesuwać centrum mapy dzięki czemu możliwe stało się drukowanie dowolnych map na żądanie klienta.

Przed metodami komputerowymi nie obroni się nawet **filozofia**, nie tylko ze względu na konieczność filozoficznej refleksji nad obecnymi tendencjami w tej dziedzinie, a w szczególności dyskusją nad możliwością powstania sztucznej inteligencji czy rozumieniem świadomości. Możliwość dyskusji filozoficznych z partnerami z całego świata poprzez sieci komputerowe otwiera tu zupełnie nowe możliwości. W ciągu pierwszego miesiąca od założenia bazy danych, dostępnej przez sieci komputerowe, zawierających teksty najnowszych prac filozoficznych skorzystało z niej ponad 3000 osób. Powstają również projekty w których większa grupa specjalistów z różnych dziedzin filozofii stara się stworzyć spójną filozofię opartą na jednolitych zasadach. W realizacji projektu *Principia Cybernetica* bierze udział wielu filozofów wspólnie pracujących nad pewnym stanowiskiem ogólnym, opartym na podejściu cybernetycznym, ewolucyjnym, ujmującym w sposób systemowy spontaniczne pojawianie się różnych poziomów organizacji systemów złożonych. Informacja o tym projekcie dostępna jest pod adresem <http://pespmc1.vub.ac.be/>

1.2.8 Komputery w badaniach literackich

Komputery przydają się też w badaniach literackich. Przy końcu 13 wieku wynaleziono konkordancje - indeksy słów, podających miejsce ich występowania i fragmenty tekstu, w których występują. Jest to ogromnie pracochłonne. Początkowo jedynie Biblia była

dziełem godnym opracowania konkordancji, jednakże w 19 wieku praktyka analizowania tekstu przy pomocy konkordancji rozszerzyła się na dzieła literackie. Jednym z pierwszych zastosowań komputerów do tego typu zastosowań było stworzenie konkordancji dzieł Świętego Tomasza z Akwinu.

Włoski jezuita, Ojciec Roberto Busa, zgłębiał teologiczne subtelnosci słowa „obecność” w dziełach Św. Tomasza. Niestety, bardzo często słowo to zastąpione zostało po prostu skrótem „w”, np. „jest w czymś”, pomijając „obecny”. Po latach pracy nad tekstami, zawierającymi 10 milionów słów, Ojciec Busa doszedł do wniosku, że jedynym rozwiązaniem jest stworzenie pełnej konkordancji wszystkich tekstów. W 1949 roku Busa napisał do szefa i założyciela młodej firmy IBM, Thomasa Watsona. W poczekalni zauważył hasło firmy: „Rzeczy trudne załatwiamy od ręki; niemożliwe zajmują nam troszkę dłużej.” Firma IBM obiecała mu pomoc. W kościele w pobliżu Mediolanu zainstalowano kilka dziurkarek kart perforowanych, czytnik kart i drukarkę a później również komputer. Każda pomyłka w wpisywaniu danych wymagała drukowania kart od nowa. Trwało to przez 18 lat, do roku 1967! Przez następne 13 lat Ojciec Busa sortował i przygotowywał do druku konkordancję. W sumie projekt ten trwał ponad 30 lat, wymagał 1.8 miliona godzin pracy ludzkiej i 10 tysięcy godzin pracy komputera, a jego rezultatem było 60-tomowe dzieło, zawierające 70 tysięcy stron!

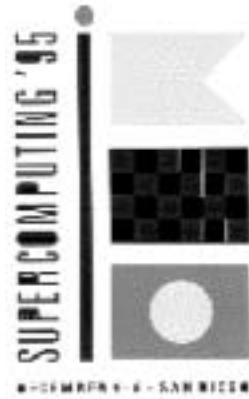
Dzisiaj komputer osobisty, wyposażony w CD ROM, pozwala w ciągu sekund odszukać kolejne pojawienie się danego słowa w tekście. Bez trudu zakupić można dzieła wszystkie Szekspira a nawet poezje zebrane, wydane w języku angielskim od początku świata. Tworzenie konkordancji dużych tekstów, jeśli już ktoś bardzo chce to zrobić, trwa minuty, a publikacja takich wyników mija się całkowicie z celem.

Kiedy w 1992 roku odwiedzałem Uniwersytet Oxfordzki wdałem się w czasie obiadu w rozmowę na temat komputerowych badań literackich z siedzącym obok mnie arabistą. Wspomniałem mu o pionierskich pracach w tej dziedzinie wykonanych w Oxfordzie i o heroicznej pracy ojca Busa. Mój rozmówca uśmiechnął się i powiedział: „A, tak, ojciec Busa dzwonił właśnie w ubiegłym tygodniu, proponował, byśmy zrobili konkordancję Koranu, ale to już dawno zrobiono”. Szybkość rozwoju technik komputerowych powoduje, że w wielu dziedzinach pracę całego życia daje się, lub da się wkrótce, powtórzyć w ciągu kilku godzin.

Stylometria, czyli badanie stylu literackiego metodami ilościowymi, powstała jeszcze przed epoką szerokich zastosowań komputerów, jednak dopiero dzięki wprowadzeniu nowych, komputerowych metod klasyfikacji i pełnej dostępności tekstów w formie elektronicznej dziedzina ta mogła się w pełni rozwinąć. Dzięki stylometrii można rozstrzygnąć sporne kwestie dotyczące autorstwa dzieł literackich, np. na początku lat 90-tych rozstrzygnięcia doczekała się kwestia autorstwa ostatniego z dzieł przypisywanych Szekspirowi (*The two noble kinsmen*), napisanego wspólnie z Johnem Fletcherem. Niektóre z rozdziałów (pierwszy i piąty) okazały się w całości być dziełem

Szekspira, rozdział drugi jest autorstwa Fletchera, a trzeci i czwarty w różnych proporcjach nosi cechy stylu obu autorów.

1.2.9 Nauki prawnicze



Istnieje informatyka prawnicza, zajmująca się głównie bazami danych dla potrzeb prawa. Wśród tysięcy aktów prawnych i dziesiątek tysięcy orzeczeń sądów dużo łatwiej jest odszukać te, które odnoszą się do danej sprawy w elektronicznej bazie danych niż w książkach i segregatorach. Również w Polsce powstały wyspecjalizowane biura prawnicze oferujące całe dzienniki ustaw lub wyciągi z przepisów dotyczących określonej dziedziny, np. prawa handlowego czy podatków, w elektronicznej formie. Do najbardziej znanych należy gdańska firma **Lex** oferująca takie usługi. Bardziej wyrafinowane możliwości stwarza program **Lex Polonica** (Lex Media Kraków). Na jednym dysku CD-ROM zawiera wybór przepisów z prawa administracyjnego, cywilnego, gospodarczego, finansowego i karnego, łącznie z bibliografią, orzecnictwem i wzorami typowych umów. Program można rozszerzać o własne zbiory

przepisów. Wyszukiwanie informacji, wykorzystujące wyrafinowane algorytmy firmy NeuroSoft, możliwe jest za pomocą pytań sformułowanych w potocznym języku polskim, szukanie skojarzeń z danym terminem lub na podstawie rdzenia danego słowa.

Bardziej ambitne zastosowania komputerów w naukach prawniczych zmierzają do wprowadzenia metod sztucznej inteligencji do wyszukiwania informacji na podstawie opisów, które wymagają głębszej analizy (rozumienia pytania, a nie tylko szukania słów kluczowych), do wyszukiwania sprzeczności wewnętrznych w danym zbiorze przepisów itp. Można sądzić, że wraz z rozwojem metod analizy języka naturalnego pojawią się komputerowi eksperci nauk prawniczych. Za najbardziej inteligentne oprogramowanie stworzone w ramach japońskiego projektu komputerów piątej generacji uznano system Helic-II, przeznaczony dla symulacji rozważań prawników, dyskusji między prokuratorem i adwokatem, interpretacji przepisów prawnych, intencji oskarżonych i symulacji rozumowania sędziego.

1.2.10 Nauki rolnicze

Nauki rolnicze (w odróżnieniu od rolnictwa) mamy nad wyraz rozwinięte. Czy są jakieś ambitne zastosowania komputerów w tych naukach? Jedną z nagrodzonych prac na

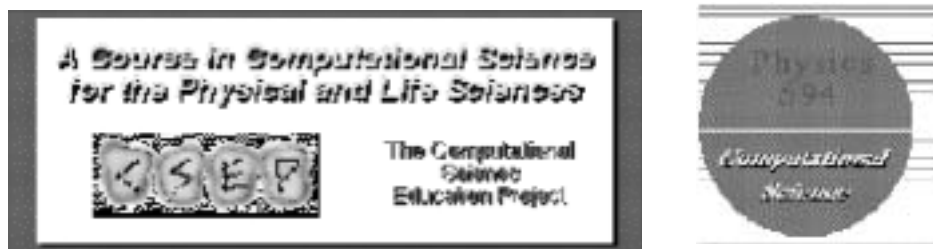
konferencji „Large scale analysis and modeling” (analiza i modelowanie komputerowe na dużą skalę), sponsorowanej przez IBM, dotyczyła oddzielenia informacji genetycznej od środowiskowej dla krów mlecznych, uwzględniając informację od wszystkich spokrewnionych krów - wymagało to nie tylko ogromnej bazy danych, ale rozwiązania układu równań o wymiarze 10 milionów! Taka praca ma duże praktyczne znaczenie i wymaga dużego doświadczenia w pracy z superkomputerem.

Wiele projektów dotyczących nauk rolniczych wymaga głębokiej wiedzy w zakresie gleboznawstwa czy zoologii oraz znajomości informatyki i analizy numerycznej. Znajomość natury zwierząt i możliwości sterowania komputerowego pomagają tworzyć bardziej naturalne warunki życia i opieki nad zwierzętami hodowlanymi. Komputery nie tylko przydają się do rozwiązywania zagadnień modelowych dla potrzeb rolnictwa, przewidywań dotyczących klimatu, lecz również są bardzo przydatne do zdalnego sterowania maszyn rolniczych a nawet całych ośrodków. W 1994 roku zwiedzałem największą farmę w Japonii specjalizującą się w hodowli krów. Było ich tam 4500 a pracujących przy nich ludzi zaledwie kilku. Krowy automatycznie dostawały odpowiednią ilość paszy i we właściwym czasie same udawały się do urządzeń dojących, również nawóz usuwany był w sposób automatyczny. Człowiek potrzebny jest przede wszystkim do spijania śmietanki...

1.2.11 Nauki komputerowe *per se*

Coraz więcej jest działów nauki, które wyłamują się prostej klasyfikacji. Można do owych „właściwych” nauk komputerowych zaliczyć część teorii systemów złożonych, można włączyć różnego rodzaju modelowanie matematyczne, np. metody komputerowego badania układów dynamicznych, najczęściej rozwijane przez fizyków, chociaż zagadnienia fizyki to zaledwie jedno z wielu możliwych zastosowań tych metod, można do nich zaliczyć teorię automatów komórkowych, również przez fizyków rozwijaną. Można tu również wliczyć programowanie naukowe (scientific programming); dziedzina ta doczekała się swojego pisma, wychodzącego od połowy 1992 roku.

Czy nauki komputerowe nie powinny pozostać działami poszczególnych nauk? I tak i nie. Tak, gdyż wymagają one kształcenia w zasadniczych kierunkach, z których wyrastają, w naturalny sposób wyodrębniając się z tych kierunków, jak to się stało z chemią komputerową. Nie, bo ludzie mają w nich wspólny język i można kształcić w znacznym stopniu tak samo ekonomistów, biologów, chemików i fizyków: programowanie, analiza numeryczna, techniki wizualizacji danych, symulacji, modelowania, nienumeryczne metody komputerowe. Często łatwiej porozumieć się chemikowi komputerowemu z fizykiem czy ekonomistą robiącym symulacje ekonometryczne niż z chemikiem analitykiem czy biochemikiem. Nauki komputerowe to



Coraz częściej studentom nauk przyrodniczych i fizyki oferowane są kursy z podstaw nauk komputerowych, nie związane z kursami programowania czy innymi zajęciami z informatyki.

nietrywialne, a więc wykraczające poza elementarne wiadomości, zastosowania komputerów w tych dziedzinach. Poza tym do nauk obliczeniowych należą podejścia interdyscyplinarne, które nigdzie nie pasują.

Czy nie są to działy informatyki? W pewnym sensie tak, jest to informatyka stosowana w różnych naukach, a cała informatyka jest gałęzią matematyki. Taki dział nie wyrasta jednak z badań w już istniejących instytutach informatyki. Programy nauczania informatyki nie są dostosowane do potrzeb w tych dziedzinach. Jeśli któraś z gałęzi zaczyna przerastać całe drzewo to czas już chyba na jej odłączenie... Nauki komputerowe to część matematyki stosowanej w takim samym sensie jak wszystkie teoretyczne nauki ścisłe, np. fizyka teoretyczna, są częścią matematyki stosowanej. Z punktu widzenia matematyki prawie nigdy nie jest to nic nowego, z punktu widzenia nauk przyrodniczych może to być wielki postęp. Nauki przyrodnicze próbują bowiem opisać świat w coraz bardziej szczegółowy sposób. Udało się zebrać sporo danych, np. zawartych w roczniku statystycznym, ale zrozumienie tych danych wymaga stworzenia modeli, a modele teoretyczne są zbyt proste by opisać rzeczywistość. Wielkie opisywanie świata dopiero się zaczęło a pojawienie się komputerów pozwala nam na dokonanie wielkiego postępu na każdym etapie tego procesu.

Do czego na przykład zaliczyć grafikę komputerową? Artyści korzystają jedynie z narzędzi i niepotrzebna im informatyka, ale jest też cała grupa ludzi - a potrzeba ich coraz więcej - posiadająca głęboką wiedzę o matematycznym modelowaniu rzeczywistości, psychologii percepcji, teksturach fraktalnych, znających się na animacji a jednocześnie obdarzonych artystycznym talentem. Im bardziej wyrafinowane narzędzia dają artystom do dyspozycji producenci oprogramowania tym głębszej trzeba wiedzy by je udoskonalić.

Co zrobiono, by ustalić tożsamość tych nowych gałęzi nauk? Mamy zatrzęsienie pism z „komputer” w tytule, konferencje na przeróżne komputerowe tematy, ale niewielu specjalistów skłonnych jest jeszcze przyznać, że jest to istotnie „trzecia siła” w nauce.

Nowych nauk nie można jednak zdefiniować, powoli nabierają one tożsamości w miarę jak pojawia się grupa specjalistów zaczynająca pracować nad podobnymi tematami używając podobnych metod, zakładając pisma i organizując konferencje.

1.3 Plany kształcenia w zakresie nauk komputerowych

Murray Gell-Mann w dalszej części cytowanej w tym rozdziale wypowiedzi na „Complex Systems Summer School” w Santa Fe, stwierdził:

Jest rzeczą coraz wyraźniej widoczną, że zrozumienie złożonych systemów wymagać będzie wspierających się nawzajem badań, prowadzonych przez specjalistów reprezentujących szerokie spektrum, od matematyki i nauk przyrodniczych do nauk humanistycznych. Społeczeństwo musi znaleźć sposoby, by pielęgnować to niezbędne zbliżenie się różnych dyscyplin naukowych i innych ważnych czynników. Istniejące obecnie instytucje akademickie nie są dobrze przygotowane by podołać tym nagłym potrzebom.

Pojawianie się symulacji komputerowych jako nowego sposobu rozwiązywania zagadnień wymaga dostosowania programów nauczania. Zmiany powinny sięgać bardzo głęboko. Geoffrey Fox porównuje obliczanie do tak fundamentalnych umiejętności jak czytanie, pisanie czy arytmetyka, umiejętności których posiadanie jest podstawą wszelkiego działania w świecie współczesnym. Powstaje więc problem, jak nauczać takiego podejścia do nauki.

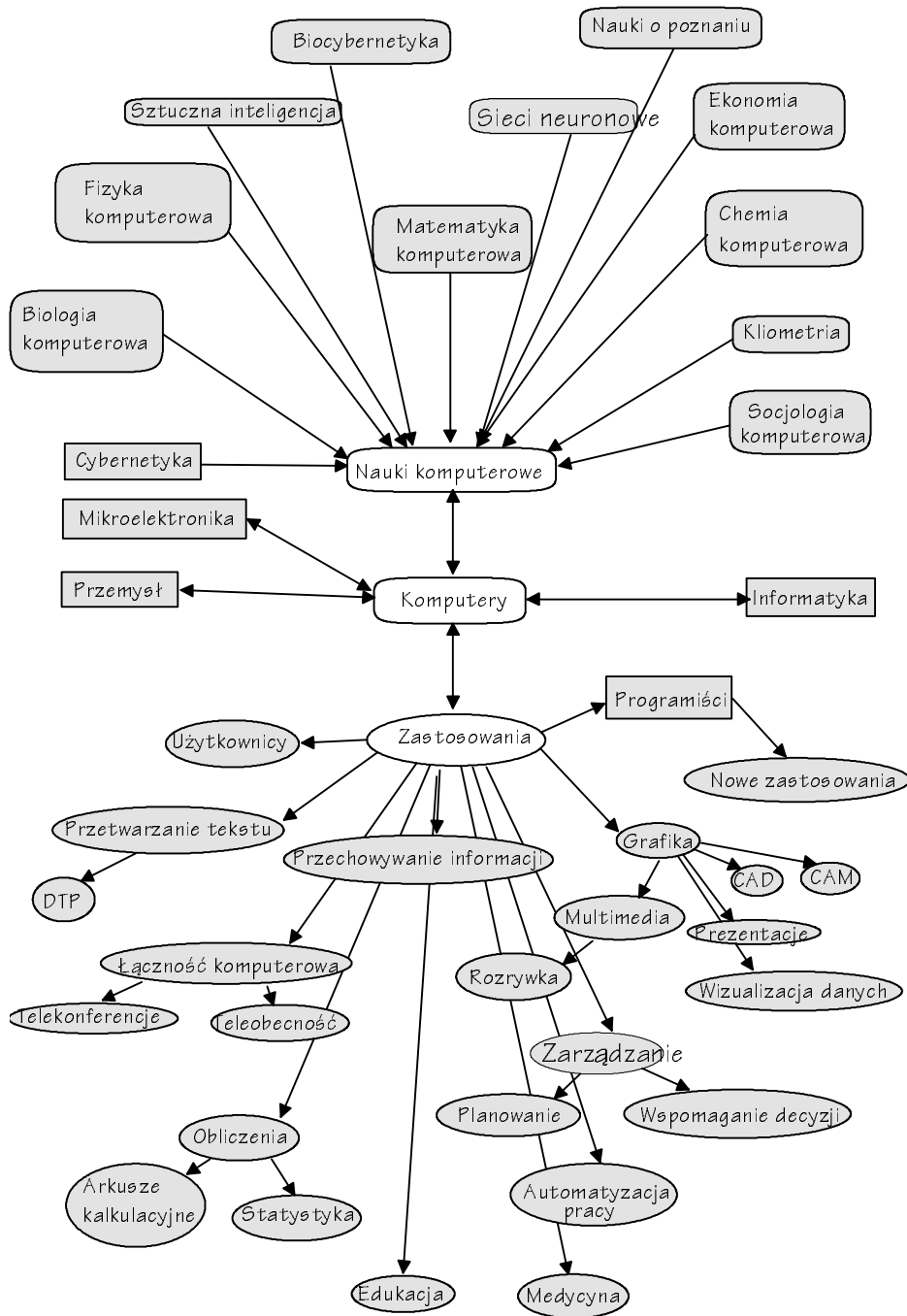
Nie ma wątpliwości, że w tej dziedzinie USA znacznie wyprzedza resztę świata. Jak więc wyglądają propozycje amerykańskie w zakresie kształcenia? Przy programach interdyscyplinarnych istnieje duże niebezpieczeństwo spłylenia wszystkiego - studenci nie osiągną poziomu magisterskiego w żadnej dziedzinie. Brakuje zarówno programów nauczania, jak i podręczników nauk komputerowych. W 1990 roku były tylko dwie uczelnie, które miały takie programy (teraz jest ich znacznie więcej). *California Institute of Technology* ma program doktorancki „Physical Computation and Complex Systems” i *Syracuse University*, gdzie mieści się „The Syracuse Center for Computational Science”, oferuje podobny program zarówno w naukach ścisłych jak i w informatyce, której część ma być nastawiona w kierunku zastosowań. Kanadyjskie uniwersyteckie instytuty informatyczne mają zwykle w nazwie „Computing Science” i spotyka się na nich grupy zajmujące się robotyką, tomografią komputerową (zastosowaniami medycznymi) czy epidemiologią, zagadnieniami bardzo dalekimi od tematów uprawianych przez informatyków w naszym kraju.

Matematyka stosowana bardzo rzadko zbliża się na tyle do nauk przyrodniczych, by przyczynić się w nich do postępu w rozwiązywaniu specyficznych problemów numerycznych dla danej dziedziny. Np. metody diagonalizacji dużych, rzadkich macierzy o stosunkowo dużych elementach diagonalnych, potrzebne w niektórych działach fizyki, zostały rozwinięte przez samych fizyków, podobnie dzieje się w przypadku niektórych typów równań różniczkowych. Interdyscyplinarne programy edukacyjne stwarzają szansę lepszej współpracy pomiędzy matematykami, informatykami i specjalistami z danej dziedziny.

Nieco mniej ambitne programy studiów, zmierzające w kierunku studiów interdyscyplinarnych, mają różne uniwersytety europejskie. Np. Uniwersytet w Odense, w Danii, prowadzi studia mieszane „informatyka z fizyką”, „informatyka z chemią”, i „informatyka z biologią”. Programy te łatwo jest rozszerzyć w kierunku nauk komputerowych. Programy nauczania większości uczelni w Polsce nie sprzyjają studiom interdyscyplinarnym i niewiele się pod tym względem dzieje. Nie można robić kariery naukowej w nowej dziedzinie, bo z braku specjalistów nie znajdzie się rada naukowa, przed którą taką pracę można by obronić. Specjalistów nie będzie jednak, dopóki nie będzie można robić stopni naukowych w danej dziedzinie. I tak mamy problem: czy najpierw była kura czy jajko?

Warto wspomnieć o „międzywydziałowych studiach przyrodniczych” na Uniwersytecie Warszawskim, prowadzonych od początku lat 90-tych. W połowie lat 90-tych wiele polskich uniwersytetów prowadziło intensywne prace nad reformą studiów można więc mieć nadzieję, że powoli skończy się kształcenie ściśle związane z kierunkiem a znacznie zdobywanie przydatnej wiedzy. Z drugiej strony organizacja szkolnictwa wyższego w Polsce nie sprzyja takiej reformie. Tradycjoniści wbili sobie do głowy, że kierunki studiów to geografia, matematyka czy biologia i uzyskanie stopnia magisterskiego w tych dziedzinach wymaga zgłębienia wszystkiego, co w nich najważniejsze. Tymczasem coraz więcej będzie trzeba specjalistów o różnorodnych umiejętnościach, np. w zakresie zarządzania i administracji, z pewnym rozeznaniem w zagadnieniach geografii społecznej, migracji ludności, ekonomii, informatyki, tworzenia i wykorzystywania map cyfrowych.

Programy edukacyjne na Syracuse University nie byłyby możliwe bez współpracy z Northeast Parallel Architectures Center (NPAC). Wiele ośrodków superkomputerowych prowadzi swoją własną działalność naukową, rozwijając oprogramowanie do celów naukowych i oferując (za darmo) czas swoich komputerów badaczom, ocenianym przez rady naukowe tych ośrodków. Np. Minnesota Supercomputer Institute (MSI), pierwszy amerykański uniwersytet, który zakupił superkomputer (jeszcze w 1981, obecnie mają wiele superkomputerów najwyższej klasy), ma w swojej radzie programowej („institute fellows”) 10 osób ze specjalnością „obliczenia naukowe” (scientific computation), ma też chemików, fizyków, informatyków, matematyków, astronomów, materiałoznawców, inżynierów różnych specjalności, biochemików, farmakologów, geologów i neurologów.



MSI organizuje liczne seminaria i sympozja, ma własne wydawnictwa a liczba prac w nim wykonanych sięga kilkuset w ciągu roku.

Ośrodki japońskie zorganizowane są na podobnej zasadzie. W 1989 roku było już ponad 120 ośrodków badawczych (połowa to ośrodki przemysłowe) posiadających superkomputery. Przykładem może być Institute for Supercomputing Research (ISR), położony w centrum Tokyo, w którym pracuje 16 osób z 5 krajów, zajmujących się przede wszystkim porównywaniem różnych architektur komputerów współbieżnych jak również obliczeniami z dynamiki płynów, mechaniki kwantowej, struktur półprzewodnikowych, fuzji termojądrowej, automatów komórkowych, sieci neuronowych. Instytut organizuje kilkudniowe sympozja i wydaje raporty techniczne na wysokim poziomie.

CERFACS, czyli European Center for Research and Advanced Training in Scientific Computation, otworzono w 1987 roku w Tuluzie. Ośrodek ten posiada najlepsze superkomputery, jest silnie związany z przemysłem lotniczym i kosmicznym, i powstał jako przeciwwaga dla amerykańskich ośrodków tego typu ściągających francuskich ekspertów. Ośrodek jako jedno ze swoich głównych zadań traktuje działalność edukacyjną, ucząc metod i zastosowań symulacji numerycznych. Cztery główne grupy badawcze to grupa algorytmów współbieżnych, mechaniki płynów, symulacji aerodynamicznych i hipersonicznych i grupa badająca niestabilności i turbulencje.

Chociaż z punktu widzenia kosztów obliczeń opłaca się inwestować w sprzęt typu roboczych stacji ośrodki superkomputerowe spełniają bardzo ważną rolę jednocząc środowisko badaczy z różnych dziedzin. Również w Polsce w 1993 roku powstały pierwsze ośrodki superkomputerowe: w Warszawie, Krakowie i Poznaniu. Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego (ICM) Uniwersytetu Warszawskiego wzorowane jest na najlepszych ośrodkach amerykańskich tego typu.

Literatura

- David Harel, *Rzecz o istocie informatyki* (Wyd. Naukowo-Techniczne, Warszawa 1992)
Bardzo dobry wstęp do „prawdziwej” informatyki.
- H.W. Roetzheim, *Laboratorium złożoności* (Intersofland, Warszawa 1994)
Eksperymenty komputerowe z fraktalami, galaktykami, genetyk
- Ellen Thro, *Sztuczne życie, zestaw narzędzi badacza* (SAMS Publishing, 1994)
Eksperymenty komputerowe z genetyką, ekosystemami i sztucznymi żyjątkami.
- Richard Dawkins, *Ślepy Zegarmistrz* (PIW, W-wa 1994)
Opisuje eksperymenty komputerowe z modelowaniem ewolucji genetycznej.