

Czy istnieją nauki komputerowe?

Włodzisław Duch (duch@phys.uni.torun.pl)

Katedra Metod Komputerowych UMK,

ul Grudziądzka 5, 87-100 Toruń

1. Czym zajmuje się informatyka?
2. Po co nam nauki komputerowe?
3. Komputery.
4. Przykłady: matematyka, fizyka, chemia, biologia, systemy złożone, nauki humanistyczne.
5. Nauki komputerowe.
6. Plany kształcenia – inicjatywy amerykańskie.
7. Integrująca rola ośrodków superkomputerowych.
8. Nasze inicjatywy.

1 Czym zajmuje się informatyka?

Instytyty informatyki i computer science powstawały masowo w latach 60-tych. Nazwa „informatyka” powstała w 1968, przyjęła się w RFN, Francji i reszcie Europy, w USA stosowana jest nazwa „Computer Science”, w Kanadzie spotyka się „Computational Science”, Nauki Obliczeniowe. Z założenia są to nauki o komputerze, ukierunkowane na komputery, ich celem jest badanie architektur, sposobów programowania komputerów, rozważania nad obliczalnością różnych problemów.

Definicja encyklopedyczna (PWN 1974): informatyka zajmuje się całokształtem przechowywania, przesyłania, przetwarzania i interpretowania informacji. Wyróżnia się w niej 2 działy, dotyczące sprzętu i oprogramowania.

Definicja ta z jednej strony jest bardzo ogólna, z drugiej strony większość tego, co robi się w ośrodkach superkomputerowych, do tej definicji nie pasuje. Popatrzmy na kategorie, według których zorganizowane są *Computing Reviews*.

- A. General Literature
- B. Hardware
- C. Computer Systems Organization (architectures, networks, implementations)
- D. Software (programming, operating systems)
- E. Data (structures, representation, encryption, information theory)
- F. Theory of Computation (abstract computation, complexity of algorithms, logics)
- G. Mathematics of Computing (numerical analysis, discrete mathematics, statistics)
- H. Information Systems (databases)
- I. Computing Methodologies (artificial intelligence, graphics, simulation theory)
- J. Computer Applications (administrative, physical sciences, life and medical, CAE)
- K. Computing Milleux (history, industry, education).

Przykład drugi: lista głównych kierunków badawczych Instytutu Informatyki UW (informator II UW 1988):

Analiza algorytmów (złożoność obliczeniowa, metody konstrukcji, kombinatoryka, teoria automatów i języków formalnych), analiza numeryczna, języki programowania, metody informatyki (specyfikacji, projektowania), podstawy informatyki, systemy operacyjne, teoria programów (logika algorytmiczna, specyfikacja aksjomatyczna).

Są to ważne i ciekawe zagadnienia, ale zupełnie się nie pokrywają z zainteresowaniami „komputerowców”, którzy szukają rozwiązań konkretnych problemów naukowych.

2 Po co nam nauki komputerowe?

Podział między naukami jest czysto umowny! Dyskusje na temat: „czy fizyka matematyczna to jeszcze fizyka czy już matematyka” są bezpłodne. Nie chodzi o nazwy, lecz o pewne poczucie tożsamości dziedziny, metodologię, środowisko, język a co za tym wszystkim idzie sposób podejścia do rozwiązania problemu. Wspólny język trudno jest znaleźć nawet w obrębie jednej dziedziny. Ma to swoje ujemne, ale i dodatnie strony – dzięki temu nasze spojrzenie na rzeczywistość jest bogatsze.

Nauki zwykło się dzielić na teoretyczne i doświadczalne. Pojawienie się komputerów 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. Komputery są prawdziwie uniwersalnymi maszynami.

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. M. Gell Mann (Nobel 1969 za teorię kwarków), przemawiając w czasie „Complex Systems Summer School” w Santa Fe, powiedział (cyt. za G. Fox, Physical computation, C^3P – 928 technical report):

The transformation of society by the scientific revolution of the 19th and 20th centuries is about to be overshadowed by even more sweeping changes arising from growing ability to understand the complex mechanisms which are central to human concerns. The technology base of the new revolution will be provided by almost unimaginably powerful computers together with the mathematical and experimental tools and associated software which are essential to achieving an understanding of complexity... Examples of adaptive, complex systems include biological evolution, learning and neural processes, intelligent computers, protein chemistry, much of pathology and medicine, human behaviour and economics.

Tego typu zastosowania znajdują się na peryferium zastosowań informatyków, chociaż jest to gałąź, która powoli zaczyna przerastać całe drzewo... Zanim zastanowimy się dokładniej nad naukami komputerowymi przyjrzyjmy się najpierw narzędziu, na którym oparta będzie ta „nowa rewolucja”, próbując zrozumieć, na czym polega jego unikalność.

3 Komputery

Komputery zrobione są z „hardwaru” czyli żelastwa.

Wiemy, jak wielki postęp nastąpił w ostatniej dekadzie. Czy i w tej spodziewać się możemy równie wielkiego postępu? Rynek i wartość inwestycji komputerowych wzrosły bardzo znacznie (jest już ponad 70 milionów komputerów) można więc spodziewać się dalszego zwiększenia tempa innowacji. Komputery zbudowane są obecnie z półprzewodników. Porównajmy obecne możliwości technologiczne z tym, co mamy w komputerach.

Szybkość zegara	Czas cyklu	System
10 MHz	100 ns	PC XT
100 MHz	10 ns	Cray 1, najszybsze 486, P5, Alpha
250 MHz	4 ns	Cray 2
500 MHz	2 ns	Cray 3 (1992, miał być w 1989!)
100 GHz	10 ps	bramki Si (1989)
200 GHz	5 ps	GaAs (1989) izolowane tranzystory
2000 GHz	0.5 ps	bramka nadprzewodzącego złącza Hitachi (1989)

Im szybszy procesor tym musi być mniejszy! Jednostka centralna Craya-3 to podkowa 1 m średnicy i 15 cm wysokości.

Przedstawiciele IBM na konferencji Super! 1990: tempo rozwoju się nie zmniejszy, do fizycznych ograniczeń szybkości komputerów jest jeszcze daleko. Opracowuje się już 256 Mb kości pamięci, mówi o 45 GB danych na 2 calach kwadratowych.

W ramach USA Strategic Supercomputer Initiative postawiono sobie za cel budowę komputera o szybkości biliona operacji na sekundę (Teraflop Supercomputer). W 1992 roku przyznano 2.9 mld. \$ w ramach Federal High Performance Computing and Communication Initiative. Taka szybkość obliczeń ma być osiągalna już w 1995 w niektórych zastosowaniach. Mówi się o 4 maszynach, które mają największe szanse osiągnąć takie szybkości: GigaCube (Parsytec, RFN) opartej na sześciianach zawierających po 68 transputerów T9000, osiągających prędkość rzędu Gflops – można łączyć ze sobą 16000 takich sześciianów; CM5 (Thinking Machines), zbudowaną z procesorów SPARC z wektorowymi koprocesorami, których będzie również około 16000; NC3 (NCUBE) mający pojawić się w 1994 roku i przy 10000 procesorów nC3 osiągnąć szybkość teraflopa; CS-2 Performance Computer Industries (Meiko, Parsys -UK, Telmat-F) zamierza zbudować swój teraflopowy komputer już w 1993 roku w oparciu o SuperSPRAC, Fujitsu VP2000 IC i nowe kości komunikacyjne, używając 1000–2000 węzłów.

W ramach tej inicjatywy powstać mają interdyscyplinarne zespoły do rozwiązywania wielkich problemów obliczeniowych (takich jak symulacja pogody, obliczanie własności molekuł i materiałów, farmaceutyków, polimerów, półprzewodników, symulacji inżynierskich, modelowania złóż ropy, modelowania ekonometrycznego, biologii molekularnej i inżynierii genetycznej).

Jednocześnie w ramach tej inicjatywy projektuje się GIGAMAC, sieć o szybkościach przekazywania danych rzędu Gbitów/sekundę, jest to projekt rządu USA z firmą IBM i Bell Lab. Jednym z pierwszych zastosowań sieci GIGAMAC będzie analiza danych onkologicznych w 3 wymiarach. Inne zastosowania to rozproszone obliczanie na wielu superkomputerach, przekazywanie danych z eksperymentów, np. SSC (Superconducting SuperCollider) ma dawać 400 TB danych rocznie – dla porównania, w bibliotece kongresu USA jest tylko 20 TB danych.

Tranzystory kwantowe: obecnie półprzewodniki VLSI, ale już nadchodzi nowa technologia wykorzystująca efekty tunelowe w supersieciach, nie tylko studnie kwantowe (ograniczenie w 1 wymiarze) ale i kwantowe druty i kropki (w 3 wymiarach) mające być podstawą nanoobwodów: rozwój tej technologii potrwa 5-10 lat zanim wpłynie znacząco na rynek.

Komputery Optyczne i neurokomputery optyczne: całe numery Applied Optics poświęca się temu zagadnieniu. Praktyczne komputery pojawiają się około 2000 roku, pierwsze zastosowania komercyjne w połowie tej dekady, duży postęp nastąpił w 1990 roku, grupa AT&T Bell Lab, kierowana przez Allana Huanga zrobiła S-SEED (symmetric self electro optic effect), bramki zmieniające z częstotliwością 1 GHz współczynnik odbicia z 10 na 60%. Ich celem jest zrobienie 1000 kanałów wejścia/wyjścia, o szybkości 1 Gbit/s każdy, w jednym urządzeniu. Są plany budowy płaskiego procesora optycznego używając fotolitografii na kwarcu i holograficznej dyfrakcji światła. Duży program budowy komputerów optycznych ma również Japonia. Zaletą optycznych procesorów jest upakowanie trójwymiarowe – światło krzyżując się nie przeszkadza sobie a połączenia elektryczne przeszkadzają.

W dalszej perspektywie mamy *komputery molekularne (nanokomputery)*. W John Hopkins University, Applied Physics Laboratory, łączy się atomy miedzi z organicznymi cząsteczkami by otrzymać MED (molecular electronic device). Cząsteczki TCNQ używać można jako przełączniki: światło laserowe zmienia przewodność cienkich warstw takich materiałów. W Center for Molecular Electronic, Syracuse University, mówi się o bakteriorodopsynie jako pamięci RAM (oceniana gęstość informacji 1 GB na cm^3), i do konstrukcji bramki NAND o cyklu 3 ps.

Wpływu tych technologii spodziewać się można po roku 2000. Do tego czasu np. Intel planuje co roku nową wersję mikroprocesora i860, w roku 2000 ma ona wykonywać około 2 miliardów instrukcji w ciągu sekundy. Cena rynkowa mikroprocesora 80486 ma spaść do 1\$! Już w 1992 roku procesor P5 Intela wykonywać ma ponad 100 MIPS. Obecne ceny stacji roboczych dzielone przez ich numeryczną wydajność są w pobliżu 1.000\$ za Mflop podczas gdy jeszcze kilka lat temu 10.000\$ za Mflop było marzeniem.

Tanie procesory RISC i CISC używane są obecnie w komputerach o bardzo różnorodnej architekturze. Są wśród nich: architektury wektorowe, równoległe (współbieżne), VLIW (very long instruction word), rozproszone (distributed).

Przykłady komputerów współbieżnych: NCUBE 2 ma do 8192 procesorów, 27 Gflps peak, do 64 MB/procesor, hiperszescian, architekturę MIMD; Thinking Machines CM2, pierwotnie 4096 lub 8192 procesory za 1 M\$, architektura SIMD, 2 Gflop, obecnie CM5 jest maszyną typu MIMD i może mieć do 2^{16} procesorów. Intel produkuje iPSC Hypercube i860 mający do 128 procesorów, na tych samych procesorach działa seria FX firmy Alliant.

Przyszłość obliczeń to przede wszystkim nowe architektury, np. Advanced Computing Research Facility w Argonne ma wszystkie nowe typy komputerów, by się z nimi „pobawić”. Obserwować można 3 główne nurty rozwojowe: współbieżność, systemy rozproszone i grafika. Przez ostatnie 15 lat dominowały komputery wektorowe, teraz coraz więcej jest egzotycznych architektur. Maszyny typu SISD (Single Instruction Single Data) są najdroższe w przeliczeniu na liczbę operacji, SIMD (Single Instruction Multiple Data, np. wektorowe) są tańsze, a wieloprocessorowe jeszcze tańsze, wreszcie maszyny o bardzo wielu prostych procesorach (multikomputery lub maszyny MIMD, Multiple Instruction Multiple Data) są najtańsze. Grafika telewizyjnej jakości wymaga przesyłania danych z szybkością rzędu Gbita na sekundę! Lokalnie da się to zrobić, wysyłanie danych na większą odległość tak szybko wymaga superszybkich sieci komputerowych.

Za 10 lat zostaną chyba tylko 3 typy maszyn: stacje graficzne, pełniące rolę PC; superkomputery; serwery baz danych. Obecnie połowiczny czas życia dużych maszyn to 3 lata. Widoczny jest powolny ruch w stronę języków programowania 4 generacji i OOP (zorientowanych obiektowo). Oprogramowanie znacznie wolniej się rozwija niż hardware. Powszechna dostępność komputerów poprzez sieć powoduje problemy z systemami operacyjnymi, które mogą zostać zaatakowane przez wirusa na całym kontynencie w ciągu kilku sekund!

Kryzys oprogramowania: nie mamy czasu na liczenie, trzeba ciągle dostosowywać programy bo języki, architektury komputerów i systemy operacyjne tak szybko się zmieniają. Szczególnie superkomputery wymagają często „dostrojenia” programów do ich architektury. Testy szybkości przy pomocy 13 du-

żych pakietów programów pokazały, że na najszybszych komputerach (NEC SX-2, CRAY 2S/4128, Y-MP/832 itp) wykonywały one od 5 do 500 milionów zmiennoprzecinkowych operacji na sekundę, więc na kilkuprocessorowych wektorowych maszynach otrzymać można różnice o czynnik 100! Jeszcze większe różnice wystąpić mogą na komputerach współbieżnych o dużej liczbie procesorów.

Powszechna jest opinia, że nawet w komercyjnym oprogramowaniu znacznie sprzęt wyprzedził oprogramowanie. Z okazji 15-lecia pisma BYTE specjaliści pytani: *jaka jest największa przeszkoda dla rozwoju zastosowań komputerów*, zgodnie odpowiadali: oprogramowanie. Jest kosztowne, np. Japonia wydaje 50 M\$ rocznie na rozwój interakcyjnej grafiki 3D. Tymczasem brakuje nam lobby popierającego rozwój dziedzin, związanych z oprogramowaniem – znacznie łatwiej uzyskać dobrą ocenę projektów związanych ze sprzętem, niż z oprogramowaniem.

Metodologia programowania: każdy buduje swój własny dom, odkrywa koło, brak jest cegieł. Takimi prymitywnymi cegiełkami są BLAS, LINPACK, EISPACK, NAG, ostatnio LAPACK na maszyny współbieżne, 3 lata, 1.2 M\$. Pewną nadzieją dla dużych projektów, pisanych od podstaw, jest CASE (Computer Aided Software Engineering), ale kto w nauce ma pieniądze na finansowanie dużych projektów związanych z oprogramowaniem? Większość dużych programów wywodzi się z lat 60-tych i jest zbieraniną fragmentów trudną do modyfikacji. Bardzo popularne staje się programowanie zorientowane obiektowo: ma to umożliwić masową produkcję oprogramowania przez zmianę filozofii jego budowy na filozofię „klocków Lego”. Mamy już rozszerzenia obiektowe do C, Pascala, Fortranu 90. Niestety, brak pieniędzy na pisanie dużych systemów od nowa każe wątpić w możliwość pojawienia się dobrych pakietów programów naukowych w bliskiej przyszłości.

Trudności w rozwijaniu dużych systemów oprogramowania doprowadziły Kennetha Wilsona do wysunięcia projektu o nazwie Gibbs, którym kieruje obecnie matematyk, David Gries. Grupa 8 ludzi pracuje nad obiektowo zorientowanym językiem, mającym uprościć pisanie dużych systemów, integrując 3 aspekty: struktury programu, struktury danych i dokumentację. Projekt Molix G. Dierckse na Monachium jest nieco skromniejszy, bo usiłuje narzucić pewne standardy w Fortranie 77 by osiągnąć to samo.

Rozwój algorytmów dla komputerów współbieżnych wymaga wielkiej pracy, jest to jedna z najbardziej aktywnych dziedzin związanych z oprogramowaniem na świecie.

4 Nauki komputerowe: przykłady

Trudno jest zdefiniować, co to np. jest fizyka, gdzieś kiedyś padło takie określenie: fizyka to to, co robią fizycy po nocach. Przyjmijmy roboczą definicję: każdy, kto spędza większość swojego czasu pracy z komputerem, zmuszając go do nietrywialnych obliczeń celem rozwiązania jakiegoś naukowego problemu, uprawia nauki komputerowe. Przedstawicielem nauk komputerowych nie jest więc programista w banku, spędzający całe dni nad rozwojem systemu oprogramowania, ani programista systemów operacyjnych ani informatyk, dowodzący w ścisły sposób poprawności programu sortującego. Nauki komputerowe uprawia natomiast biolog, który spędza całe lata symulując ewolucję lub związanie się białka, fizyk, usiłujący rozwiązać równania QCD czy matematyk, spędzający lata rozwijając, testując i używając programy sprawdzające wszystkie możliwe przypadki jakiegoś twierdzenia. Złożoność takich zagadnień i programów, służących do ich rozwiązania jest tak duża, że praca w tej dziedzinie przypomina bardziej pracę doświadczalną niż teoretyczną.

Zanim pozwolę sobie na ogólne uwagi dotyczące nauk komputerowych popatrzmy na ich przykłady, zwracając szczególnie uwagę na duże projekty softwarowe.

4.1 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 znacznie mierze 19-wieczna, daje się uprawiać przy pomocy programów komputerowych służących do manipulacji symbolicznej.

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 Cray-u w Harwell 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!

Żadna z tych prac nie kwalifikuje się jako matematyka komputerowa, gdyż komputery grały w nich 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, doprowadziło do znacznego rozwoju takich gałęzi matematyki jak teoria grafów i doczekało się dowodu, którego nikt nie może sprawdzić i który zawiera kilka błędów! Problem sprowadzono do zbadania bardzo dużej liczby grafów: w 1970 roku oceniano, że potrzeba na to 11 lat ciągłej pracy szybkiego komputera. Wydawało 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 3 komputerów, w latach 1970-76, zużywając ponad 1000 godzin CPU i prowadząc w istocie dialog z programem. Ponieważ w używanych komputerach pojawiały się czasem przekłamania istnieje pewne, choć bardzo małe, prawdopodobieństwo, że dowód jest błędny (vide późniejsze kontrowersje). Komentując swoje wyniki K. Appel i W. Haken stwierdzili (cyt za: Zagadnienie czterech barw, w: Matematyka współczesna, 12 esejów, WNT 1983)

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.

W ostatnich latach tego typu dowody zdarzają się często, powstał nawet *Geometry Supercomputer Project*, 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 ciekawa (cyt. za *The Mathematical Intelligencer* 12 (1990) 8):

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 o takich superkomputerach marzą. 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 (Mandelbrot od dawna nie uprawia już „klasycznej” matematyki), symulacją automatów komórkowych czy próbą wyobrażenia sobie 3-wymiarowych przekrojów wielowymiarowych obiektów, możliwą 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 potrafi w ciągu sekundy zbadać miliony przypadkowych możliwości.

4.2 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”

S. Wolfram (nazywany przez kolegów w Princeton następcą Einsteina), zajmował się cząstkami elementarnymi, automatami komórkowymi, stworzył system *Mathematica*: (cyt. za *Comp. in Physics* 7/89)

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.

Geoffrey Fox, CalTech, twórca programu C^3P (Caltech Concurrent Computation Program) tak to określił (cyt. za *Comp. in Physics* 1/1990):

Fizyka komputerowa jest trzecią siłą, ale nadal słabą. Dopiero stacje robocze o szybkościach Giga-flopów i superkomputery o szybkościach Tera-flopów to zmieniają. Wielkim wyzwaniem dla uniwersytetów będzie kształcenie ludzi, którzy z takiego 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 Tera-flopowego komputera.

Najbardziej popularnym podejściem (w znacznie mierze wywodzącym się z prac K. Wilsona) do rozwiązywania równań chromodynamiki kwantowej, podstawowej teorii materii, jest ich symulacja na sieciach czasoprzestrzennych (QCD on lattices), chociaż osiągnane 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 10^8 . Pewne użyteczne rezultaty, przy pomocy sieci 64^4 , osiągnąć można będzie „już” przy pomocy maszyn o szybkości tera-flopów – 40 fizyków z USA napisało wspólnie projekt budowy takiego komputera do celów QCD. Duża grupa naukowców zamierza zbudować szybki komputer do specjalnych obliczeń QCD (kilka podobnych prób, np. IBM GF-11, podjęto już w tej dziedzinie wcześniej). Czy taki sposób działania nie przypomina fizyki teoretycznej studiowanej metodami eksperymentalnymi?

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 danych (por. *Computers in Physics*, vol. 6, no.1 (1992)). Inni „pożeracze” mocy obliczeniowych to fizycy symulujący zachowanie plazmy w różnych warunkach eksperymentalnych, fizycy atmosfery symulujący zachowanie dziury ozonowej, zanieczyszczenia atmosferyczne, pogodę (dlaczego mamy kiepskie prognozy pogody? Nawet Indie zakupiły do tego celu superkomputer! Nasze prognozy są niedokładne bo nich ich dla nas specjalnie nie oblicza.), astrofizycy ewolucję gwiazd, ewolucję wszechświata i galaktyk a specjaliści od mechaniki kwantowej ciało stałe i własności pojedynczych cząsteczek. Fizyka i astronomia zużywa ponad połowę czasu superkomputerów w ośrodkach akademickich.

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.

4.3 Chemia komputerowa

Po etapie intensywnego rozwoju teorii i oprogramowania w latach 60 i 70-tych chemia kwantowa stała się w latach 80-tych przydatna chemikom doświadczalnikom i w tej chwili znacznie jest więcej użytkowników wielkich pakietów gotowych programów kwantowochemicznych niż ludzi znających się na metodach w tych pakietach wykorzystywanych. Osiągnane dokładności dla małych 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. Mamy nie tylko pisma z chemii kwantowej lecz również *Journal of Computational Chemistry*. W akademickiej sieci komputerowej mamy grupę dyskusyjną „computational chemistry group” (serwer znajduje się w ośrodku superkomputerowym w Ohio).

W powszechnym użytku jest kilkanaście uniwersalnych systemów programów, każdy z nich liczący setki tysięcy wierszy. Napisanie 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 E. Clementi’ego w firmie IBM (niestety, ostatnio projekt przerwano). Idea globalnej symulacji jest prosta: seria modeli opartych na fundamentalnych równaniach opisujących materię, od mikroskopowych do makroskopowych, korzystających z wyników pośrednich modeli na kolejnych poziomach powinna nam pozwolić wyprowadzić wszystkie własności materii (począwszy od atomów) z komputerowych obliczeń. MOTECC 90 składa się z 30 dużych pakietów programów: od mechaniki kwantowej przez mechanikę statystyczną do dynamiki ośrodków ciągłych.

Przykładem symulacji globalnej są obliczenia dla wody. Zaczynamy od atomów: 3 jądra, 10 elektronów, używamy procedury SCF, dodajemy korelację elektronową, obliczamy hiperpowierzchnie energetyczne, analizujemy vibracje cząsteczki wody, następnie badamy 2 cząsteczki (tj. dimer) wody, obliczamy ich potencjały oddziaływania, dodajemy więcej cząsteczek badając małe klasterki wody i poprawiając potencjały oddziaływania. Przechodzimy następnie do mechaniki statystycznej: używając dynamiki molekularnej lub prostego Monte Carlo, symulujemy $10^3 - 10^5$ cząsteczek w ustalonej objętości i temperaturze, generujemy miliony konfiguracji obliczając energię całości używając potencjałów wyliczonych poprzednio, liczymy średnie statystyczne i własności, takie jak funkcje korelacji par, przekroje na rozpraszanie X i neutronów, translacyjno-rotacyjne f. 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 daje się dość dokładnie obserwować turbulencję i dostać prawidłową liczbę Reynoldsa. Mając takie makroskopowe własności możemy rozwiązywać równania dynamiki płynów i badać np. ruch fal w zatoce.

Globalna symulacja wymaga dużych mocy obliczeniowych (obliczenia dla wody prowadzone były na specjalnie skonstruowanym systemie, składającym się z wielu komputerów), ale do końca tej dekady będą możliwe na biurku. Bardzo intensywnie rozwijaną częścią chemii obliczeniowej jest modelowanie molekularne, wymagające bardzo szybkich stacji graficznych i programów dynamiki molekularnej. Wkracza to już w chemię białek, należącą do biochemii.

4.4 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 próba określenia struktury przestrzennej białka na podstawie sekwencji aminokwasów. Najczęściej próbuje się stosować różne metody heurystyczne, statystyczne lub detektory regularności oparte na modelach sieci neuropodobnych.

Jednym z najciekawszych projektów modelowania ekologicznego opublikowano w 1986 roku (D.L. Cappock *et al.*, *Journal of Applied Ecology* 1986). Dotyczył on plemienia Turkana, nomadów zamieszkujących północno-zachodnią Kenię. Plemię to żyje większość czasu na krawędzi głodu. W początkach lat 80-tych National Resource Ecology Laboratory z Colorado State University 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 zwierzęta z powodu niedożywienia dają mało mleka upuszcza pije się ich krew, wymienia na zboże lub zjada. 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.

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 wyniesie 160 milionów par rocznie, a w 1999 roku wzrośnie do 1.6 miliarda (genom człowieka zawiera prawie 3 mld par nukleotydów, w ok. 100000 genów). Do tej pory nawet tak dobrze zbadany organizm jak *Escherichia coli* zmapowany został zaledwie w 16%. 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 projektem (Comm. of the ACM 11 (1991) 45) stwierdzono, że bardzo brakuje komputerowych biologów, w tej chwili potrzeba ich już ok. 4000.

Na sponsorowanej przez firmę IBM konferencji „Large Scale Analysis and Modelling” nagrodzono 2 prace biologiczne: pierwsza dotyczyła 3-wymiarowej rekonstrukcji graficznej ucha środkowego, a druga modelowania pracy serca i analizy elektrokardiogramu. Ważnym działem biologii komputerowej są symulacje działania komórek nerwowych jak i fragmentów układu nerwowego, a w szczególności mózgu. Przykładem wpływu biologii na metody komputerowego rozwiązywania problemów są algorytmy genetyczne, pozwalające na znalezienie optymalnych rozwiązań dla szerokiej klasy zagadnień, oparte na genetycznych mechanizmach selekcji naturalnej.

4.5 Nauki o poznaniu (cognitive sciences)

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 nad takim właśnie projektem, znanym pod kryptonimem CYC: według ich ocen uniwersalny system ekspertowy, z którym można będzie o prawie wszystkim porozmawiać, wymaga bazy wiedzy zawierającej ogromną liczbę reguł, rzędu 100 milionów! Do końca 1990 roku w systemie CYC znalazło się około 2 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. Jest to projekt bardzo kosztowny i ryzykowny, gdyż nikt nie wie, czy tak wielki system będzie stabilny i naprawdę przydatny. Lenat ma nadzieję, że tworzona przez niego baza wiedzy będzie niezbędną częścią wyposażonego w zdrowy rozsądek komputera 21 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 dialog z komputerem w języku naturalnym.

4.6 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 pamięci rzędu Gigabajtów i szybkości rzędu Gigaflopów. Modelowanie rynku finansowego, przewidywanie ruchów cen na giełdzie, to najszybciej rozwijające się zastosowania superkomputerów.

4.7 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 zagadnienia. 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 „radykalną 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. Chociaż można to określić jako „komputerową historię” nie są to zastosowania komputerów wymagające szczególnie wyrafinowanych metod.

Mija właśnie 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 znaim 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 nieznane. 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 Marden, znanego żeglarsza, o wykreślenie trasy podróży odkrywcy. 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 poraz 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 używa się również zaawansowanych technik komputerowej analizy obrazu, wykorzystując do tego celu zdjęcia lotnicze i zdjęcia z amerykańskiego satelity LANDSAT i francuskiego satelity SPOT. System Informacji Geograficznej (GIS) zawiera 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.

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 subtelności słowa „obecność” w dziełach Św. Tomasza. Niestety, bardzo często słowo to zastępywane zostało po prostu skrótem „w”, np. „jest w czymś”, pozostawiając „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 trochę 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.

Szybkość rozwoju technik komputerowych powoduje, że w wielu dziedzinach pracę całego życia daje się, lub da się wkrótce, powtórzyć w ciągu godziny.

4.8 Nauki prawnicze

Istnieje informatyka prawnicza, zajmująca się głównie bazami danych dla potrzeb prawa. 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 pojawiają się komputerowi eksperci nauk prawniczych.

4.9 Nauki rolnicze

Nauki rolnicze (w odróżnieniu od rolnictwa) mamy nad wyraz rozwinięte. Czy są jakieś ambitne zastosowania komputerów w tych naukach? Jedna z nagrodzonych prac na konferencji „Large scale analysis and modelling”, 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.

4.10 Nauki komputerowe *per se*:

Coraz więcej jest działów nauki, które wylamują 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, czy teoria automatów komórkowych, również przez fizyków rozwijana.

Można tu również wliczyć programowanie naukowe (scientific programming); dziedzina ta doczekała

się swojego pisma (będzie wychodzić od połowy 1992 roku).

5 Nauki komputerowe

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 obliczeniową. 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 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 jest to dział informatyki? W pewnym sensie tak: zastosowania komputerów w różnych naukach. Nie wyrasta on jednak z istniejących instytutów informatyki, ich programy nauczania 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 teoretyczne nauki ścisłe są częścią matematyki stosowanej - z punktu widzenia matematyki nie jest to nic nowego.

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

6 Plany kształcenia – inicjatywy amerykańskie.

M. Gell Mann w dalszej części swojej 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 podjąć tym nagłym potrzebom.

Pojawianie się symulacji komputerowych jako nowego sposobu rozwiązywania zagadnień wymaga dostosowania programów nauczania. Zmiany powinny sięgać bardzo głęboko. For porównuje obliczanie do tak fundamentalnych umiejętności jak czytanie, pisanie czy arytmetyka, których zrozumienie i używanie jest podstawą wszelkiego działania w świecie współczesnym. Powstaje więc problem, jak nauczać takiego podejścia do nauki? Przede wszystkim konieczny jest sprzęt (w Polsce nie mamy prawie żadnych doświadczeń w programowaniu współbieżnym).

Jak wyglądają propozycje amerykańskie? 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. Znam tylko

2 uczelnie, które mają już takie programy (na pewno jest 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. Kanaadyjskie uniwersyteckie instytuty informatyczne mają zwykle w nazwie „Computing Science” i spotyka się na nich grupy zajmujące się robotyką, tomografią komputerową (zastosowaniami medycznymi) czy epidemiologią.

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 równań różniczkowych. Takie programy edukacyjne stwarzają szansę lepszej współpracy.

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 będzie rozszerzyć w kierunku nauk komputerowych. Programy nauczania większości uczelni w Polsce nie sprzyjają studiom interdyscyplinarnym i niewiele się pod tym względem dzieje. Warto jednak wspomnieć o studiach „przyrodniczych” na UW.

Czy mamy w Polsce jakieś instytuty, które świadomie dążą w kierunku nauk komputerowych? Wiem tylko o „Katedrze neurologii i nauk komputerowych” na AM w Krakowie i naszej „Katedrze Metod Komputerowych” na UMK. Proponujemy utworzenie *Zaawansowanego Laboratorium Metod Komputerowych* dla potrzeb fizyki i chemii komputerowej, jako laboratorium dydaktycznego dla magistrantów i doktorantów na UMK (staramy się o pieniądze z programu TEMPUS i z KBN).

7 Integrująca rola ośrodków superkomputerowych

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 (1981, obecnie mają wiele superkomputerów, m. innymi Cray-2 z 4 GB pamięci), ma w swojej radzie programowej („institute fellows”) 10 osób ze specjalnością „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 neuropodobnych. Instytut organizuje kilkudniowe sympozja i wydaje raporty techniczne.

CERFACS, czyli European Center for Research and Advanced Training in Scientific Computation, otworzono w 1987 roku w Tuluzie. Ośrodek ten posiada zarówno Cray-2, XMP jak i maszyny współbieżne, jest silnie związany z przemysłem lotniczym i kosmicznym, 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. 4 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 „workstations” lub tanie systemy współbieżne (np. sieć nowych transputerów T9000) ośrodki superkomputerowe spełniają bardzo ważną rolę jednoczącą środowisko.

Potrzebne jest lobby w Polsce, popierające rozwój takiego ośrodka!

Mamy w kraju silne lobby mikroelektroniczne, chociaż nasze szanse na dorównanie „azjatyckim tygrysom” w tej dziedzinie nie są wielkie. Nie mamy niestety żadnego lobby popierającego rozwój oprogramowania i nauk „software’owych” w odróżnieniu od „hardware’owych”. Jest w tej dziedzinie pewna szansa dotrzymania kroku najlepiej rozwiniętym krajom, chociaż czas pracuje na naszą niekorzyść – coraz trudniej jest opanować i właściwie wykorzystać, czy choćby tylko się orientować, możliwości superkomputerów. Dużych systemów oprogramowania nie można już tworzyć po amatorsku, nasi programiści nie mają zazwyczaj odpowiednich narzędzi i przygotowania do takiej pracy.

Interesująca jest historia tych ośrodków w USA. Początkowo mało kto był przekonany, że będą one przydatne i środowiska akademickie nie miały do superkomputerów dostępu do połowy lat 80-tych, potem nastąpił gwałtowny wzrost liczby takich ośrodków, dzisiaj jest ich kilkanaście i wszystkie pracują bez przerwy. Jest to związane również z rozwojem sieci komputerowych, przypadającym na połowę lat 80-tych. Do końca tej dekady przewiduje się, że w USA będzie 995 instalacji superkomputerowych, w Japonii 768, w Europie 345, w pozostałych krajach 65, więc razem ponad 2100 ośrodków!

Plany Katedry Metod Komputerowych UMK:

Fizyka molekularna - obliczenia i inteligentne systemy, chemia kwantowa, biologia molekularna, teoria i zastosowanie modeli neuropodobnych, lingwistyka komputerowa, teoria układów złożonych. Bliższe informatyki: projekty współbieżnych programów numerycznych, metodologia programowania (abstrakcyjne typy danych). Edukacja: współpraca z innymi dyscyplinami nauki nad interakcyjnymi podręcznikami.

Rynek pracy: oceny firm przemysłowych Wielkiej Brytanii przydatności poszczególnych dziedzin studiów (w nawiasie % firm, uznających za bardzo użyteczne przygotowanie w danej dziedzinie): elektronika anal/cyfrowa (87), czujniki (83), fizyka komputerowa (82), półprzewodniki i mikroelektronika (79), DSP (79), optyka i lasery (70), termodynamika (65), fizyka atomowo-molekularna (42), mechanika statystyczna (42), mechanika kwantowa (28), biofizyka (22), fizyka cząstek elementarnych (18), astronomia i astrofizyka (7).