

Stefan WĘGRZYN

Instytut Informatyki Teoretycznej i Stosowanej, PAN

## INFORMATYKA JAKO DYSCYPLINA NAUKOWA

**Streszczenie.** W artykule przedstawiono specyfikę informatyki jako dyscypliny naukowej oraz porównano biologiczne i techniczne systemy informatyki.

## INFORMATICS AS A KNOWLEDGE DISCIPLINE

**Summary.** The article presents the specific features of informatic understood as a knowledge discipline and compares biological and technical informatic systems.

### 1. Wprowadzenie

Informatyka jest dyscypliną naukową o własnych pojęciach podstawowych, badającą prawa rządzące procesami kodowania, przechowywania i przetwarzania danych. Badania te są prowadzone zarówno dla wewnętrznych potrzeb całej nauki, dla naszych potrzeb cywilizacyjnych, jak i dla stworzenia i rozwoju podstaw naukowych metod, urządzeń i systemów z tego zakresu potrzebnych ludziom.

Specyfika informatyki jako dyscypliny naukowej polega na tym, że dotychczas nauki techniczne i ścisłe, takie jak mechanika, fizyka czy chemia, badały prawa rządzące procesami ruchu i przetwarzania mas i energii. Informatyka natomiast, po raz pierwszy w historii nauki, rozpoczęła badanie praw rządzących ruchem i przetwarzaniem informacji. Celem badań naukowych w naukach, takich jak np. fizyka, jest poznanie praw rządzących ruchem i przetwarzaniem mas i energii, zarówno dla wewnętrznych potrzeb całej nauki, dla potrzeb cywilizacyjnych, jak i jako podstawy projektowania i budowy różnych potrzebnych ludziom urządzeń i systemów opartych na prawach fizyki.



Rozumiemy więc potrzebę i aprobujemy w fizyce np. sens i celowość badań nad teorią „wielkiego wybuchu”, które uzasadnione są naszymi potrzebami cywilizacyjnymi, potrzebami całej nauki, jak i badania np. nad promieniowaniem laserowym jako podstawą budowy i projektowania potrzebnych ludziom urządzeń i systemów optyki.

Identyczna sytuacja jak w fizyce jest i w informatyce jako dyscyplinie naukowej. I tu istnieją podstawowe potrzeby cywilizacyjne. Zwróćmy bowiem uwagę na to, że wszędzie tam, gdzie kończy się materia nieożywiona, a zaczyna materia ożywiona, pojawiają się zawsze pewne symbole kodowe, zapisy kodowe, można powiedzieć — zapisy informatyczne jako nieodłączne atrybuty pojawiania się życia.

Jak one powstawały, jak się rozwijały, dochodząc do tych istniejących obecnie na Ziemi systemów informatyki w organizmach żywych?

Te badania są potrzebne całej nauce, i to bardzo potrzebne, bo można powiedzieć, że to właśnie informatyka jest tą dyscypliną naukową, która może bardziej niż inne dyscypliny naukowe zbliżyć nas do tego, co by można nazwać granicą poznania.

Badanie praw rządzących procesami kodowania, przechowywania i przetwarzania danych są też nam również bardzo potrzebne jako podstawa projektowania i konstrukcji komputerów, urządzeń i systemów informatyki technicznej i ich oprogramowania, które wchodzi dziś do chyba wszystkich dziedzin działalności ludzkiej.

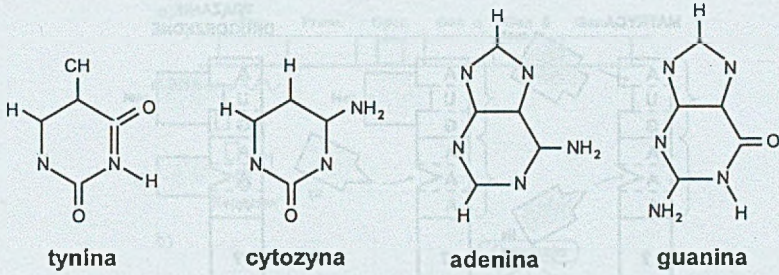
Prawa rządzące ruchem i przetwarzaniem mas i energii są niezmiennie. Podobnie prawa rządzące kodowaniem, ruchem i przetwarzaniem danych są też niezmiennie, i są na nich oparte zarówno struktura i działanie systemów informatyki, istniejących w organizmach biologicznych, jak i struktura i działanie systemów informatyki, które są obecnie projektowane i budowane przez ludzi.

Warto więc te rozwiązania porównywać i analizować.

## 2. Techniczne i biologiczne systemy informatyki

Zarówno biologiczne, jak i techniczne systemy informatyki oparte są na tych samych, następujących pojęciach podstawowych: algorytm, program i automatyczny proces realizacji programu. To są elementy wspólne. Różnice polegają natomiast na tym, że w układach informatyki technicznej przyjmuje się, że informacja elementarna, czyli bit, może reprezentować dwie różne wartości logiczne oznaczone symbolami 0,1, a w informatyce biologicznej informacja elementarna, czyli tak zwany nukleotyd, może reprezentować cztery różne wartości logiczne oznaczone symbolami T, C, A, G.





Rys. 1. Molekuły zasad azotowych dających cztery wartości logiczne kodu DNA  
 Fig. 1. Molecules of nitrogenous bases representing four logical values of DNA code

W technicznych układach informatyki bity są reprezentowane przez makroukłady elektroniczne lub magnetyczne, a w biologicznych układach informatyki nukleotydy są reprezentowane przez pojedyncze molekuły zasad azotowych tyminy, cytozyny, adeniny, guaniny (rys. 1).

Zauważmy, że koncepcje biologicznych i technicznych systemów informatyki są takie same, reprezentowane przez tę samą koncepcję podstawową: algorytm, program, a następnie system informatyczny, zawierający część urządzeniową, program i system operacyjny, nadzorujący realizację programu.

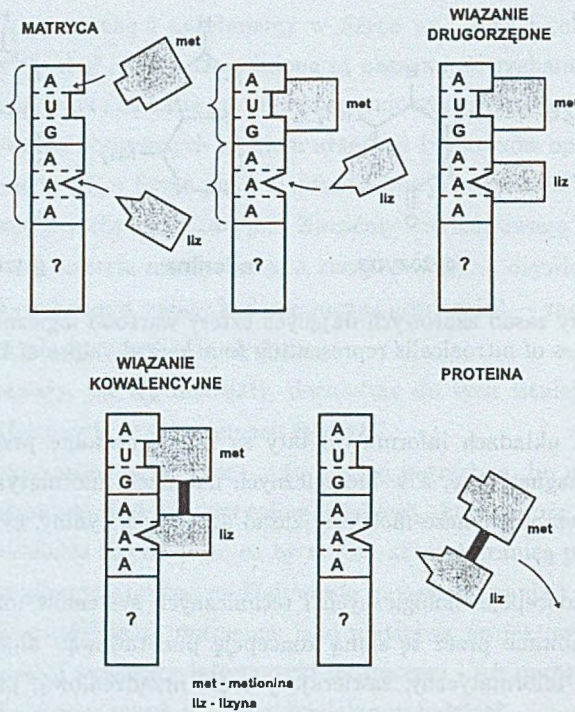
Skąd więc te różnice w sposobie realizacji tej samej koncepcji, a więc leżąca u podstaw technicznych układów informatyki, logika dwuwartościowa i leżąca u podstaw biologicznych systemów informatyki, logika czterowartościowa. Czy są to różnice wynikające z jakichś głębszych powodów, czy są tylko przypadkowe?

Te różnice między technicznymi a biologicznymi systemami realizacji programów nie są przypadkowe. Powodem tych różnic są różnice między celami, które chcemy osiągnąć, tworząc algorytmy dla technicznych systemów informatyki, a celami, które chciała osiągnąć natura, tworząc swoje algorytmy dla biologicznych systemów informatyki.

Algorytmy techniczne realizowane w komputerach to zbiory określonych operacji matematycznych na wskazanych danych, w postaci liczb, czasem obrazów czy nawet tekstów, po wykonaniu których otrzymujemy rozwiązanie w postaci liczbowej, tekstowej czy obrazowej, dowolnego zadania z określonej klasy zadań. Natomiast algorytmy realizowane w systemach informatyki biologicznej to zbiory określonych operacji technologicznych na argumentach, którymi są elementarne cząstki materii, molekuły, operacji, takich że po ich wykonaniu otrzymujemy jako wynik kompletny obiekt materialny, organizm biologiczny.

W skrócie można więc powiedzieć, że danymi dla algorytmów technicznych są liczby,





Rys. 2. Ilustracja elementarnej operacji technologicznej tworzenia białka

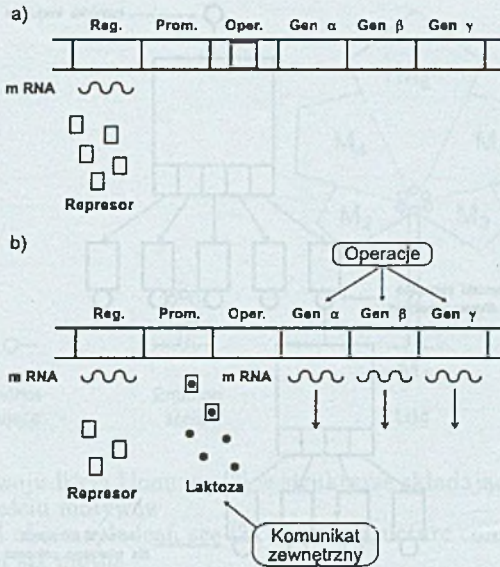
Fig. 2. Illustration of an elementary technological operation of forming a protein

operacjami są operacje matematyczne, a wynikami procesów ich realizacji są też liczby, natomiast dla algorytmów biologicznych danymi są cząstki materii (molekuły), operacjami są operacje technologiczne, a wynikiem procesów ich realizacji są obiekty (organizmy biologiczne).

### 3. Organizacja budowy obiektów biologicznych

Podstawowymi składnikami organizmów biologicznych są białka, które są łańcuchami molekuł aminokwasów. Dlatego podstawowymi składnikami algorytmów biologicznych są programy syntezy białek, zwane genami strukturalnymi, a operacjami elementarnymi tych algorytmów są operacje dołączenia molekuł aminokwasów, które trzeba w danym momencie dołączyć do tworzonego łańcucha białka, tak jak to przedstawiono obrazowo według [1] na rys. 2. Oprogramowanie natomiast to, podobnie jak w informatycznych





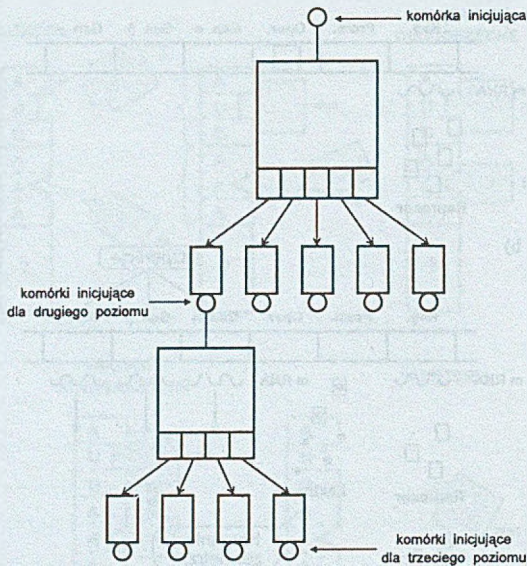
Rys. 3. Aktywizacja genów strukturalnych przesyłem komunikatów według Jakoba-Monoda — Nagroda Nobla

Fig. 3. Activation of structural genes by sending a message, following Jacob-Monod — the Nobel Prize

systemach technicznych, zbiór programów, których realizacja odbywa się pod nadzorem systemu operacyjnego. Pełny zbiór programów strukturalnych genów i sterujący nimi system operacyjny są zapisane w każdej komórce w tak zwanym łańcuchu DNA. Objętość zajmowana przez system operacyjny dochodzi do około 80% objętości całego łańcucha DNA.

Kompletne zbiory programów zakodowane w łańcuchach DNA znajdują się w każdej komórce organizmu biologicznego. Sposób realizacji programów informatycznych w organizmach biologicznych jest równoległy, i to na najniższym poziomie operacyjnym, i jest sterowany przesyłem komunikatów między komórkami. Ilustruje to rys. 3, na którym przedstawiono zasadę syntezy białka według teorii Jacoba i Monoda. Gdy brak jest laktozy, represor blokuje operator i transkrypcja genów strukturalnych nie następuje. W obecności laktozy represor staje się nieaktywny i zachodzi transkrypcja genów strukturalnych. Pojawienie się laktozy można więc traktować jako komunikat sterujący operacją syntezy białka. Koncepcję organizacji można natomiast określić jako organizację embrionalną. Organizacja embrionalna budowy obiektów polega na tym, że budowę całego obiektu, jak też dowolnego jego fragmentu, rozpoczyna się zawsze od tylko jednego





Rys. 4. Ilustracja embrionalnej strategii budowy obiektu

Fig. 4. An illustration of the embrion strategy for building an object

elementu, zwanego komórką matką, jeżeli chodzi o cały obiekt, a komórkami inicjującymi jeżeli chodzi o jego fragmenty.

Po rozpoczęciu procesu budowy system dąży do stworzenia największej możliwej liczby komórek inicjujących, od których rozpoczynają się równoległe budowy fragmentów, co można zapisać w następujący sposób:

**początek**

EMBRION  $(c_1, \dots, c_i, \dots, c_n)$

**początek**

dla  $i \in [1, n]$  równoległe

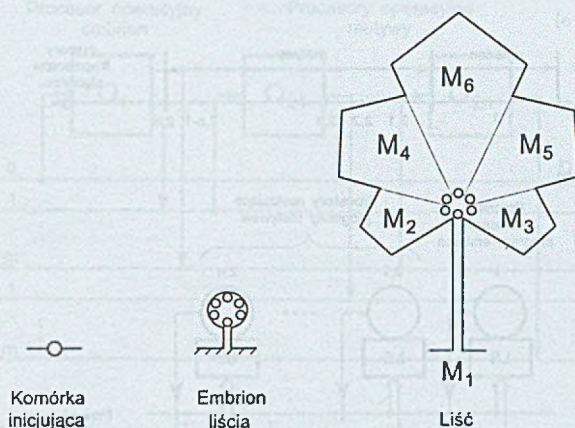
$x_i(c_i)$

**koniec**

**koniec**

Zasada organizacji embrionalnej obowiązuje, zarówno gdy chodzi o budowę całego obiektu, jak i budowę każdego poszczególnego jego fragmentu. Można wtedy mówić o wielopoziomowej embrionalnej organizacji procesu. Zapewnia to najkrótszy czas trwania budowy obiektu (rys. 4). Tę organizację budowy obiektów nazwiemy organizacją em-





Rys. 5. Model rozwoju liścia klonu wg [3] w strukturze składającej się z części embrionalnej i sześciu motywów

Fig. 5. The model of a maple leaf, see [3], having structure composed of the embrion part and of six motifs

brionalną. Ilustracją embrionalnej organizacji budowy organizmów biologicznych może być model rozwoju liścia klonu według [3] przedstawiony na rys. 5.

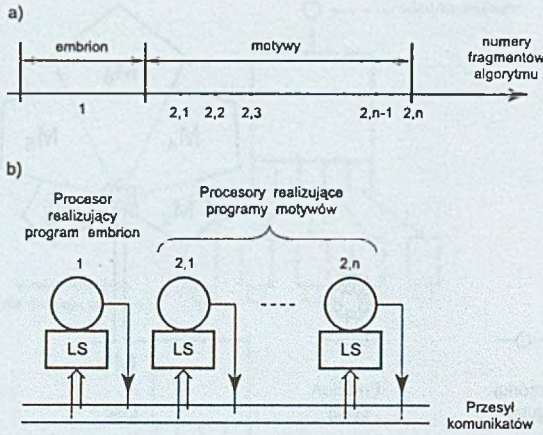
#### 4. Embrionalna organizacja procesów obliczeniowych

Przeniesienie tej, stosowanej w systemach informatyki w biologii, koncepcji realizacji złożonego obiektu na sposób realizacji złożonych programów w systemach informatyki technicznej przedstawiono na rys. 6. Ilustracją embrionalnej organizacji procesu obliczeniowego w przypadku rozwiązywania zadań o charakterze matematycznym może być np. organizacja procesu znajdowania pierwiastków całkowitych równania (1)

$$x^n + \sum_{i=n-1}^0 a_i x^i = 0 \quad (1)$$

gdzie  $a_0, a_1, \dots, a_{n-1}$  - liczby całkowite.

Odpowiedni algorytm podający rozwiązanie tego problemu można przedstawić w postaci:



Rys. 6. Ilustracje: a) embrionalnej struktury algorytmu, b) technicznej struktury systemu i jego realizacji, opartego na zasadzie sterowania przesyłem komunikatów  
 Fig. 6. Illustrations of: a) embrion algorithm structure, b) technical structure of the system and of its realization, based on the principle of the control by message transfers

- (1) operacja  $0_1$ , którą można uważać za operację embrionalną polegającą na wydzieleniu dzielników  $d_1, d_2 \dots, d_m$  liczby  $a_0$ ,
- (2)  $m$  operacji  $0_{21}, 0_{22}, \dots, 0_{2m}$ , które można uważać za motywy, których zadaniem jest sprawdzenie, które dzielniki są pierwiastkami  $p_2$  równania (1), co można zapisać:

$$p_{j \in \{1, m\}} = \begin{cases} d_j & \text{jeżeli } d_j^n + \sum_{r=n-1}^n a_r d_j^r = 0 \\ \phi & \text{jeżeli } d_j^n + \sum_{r=n-1}^0 a_r d_j^r \neq 0 \end{cases}$$

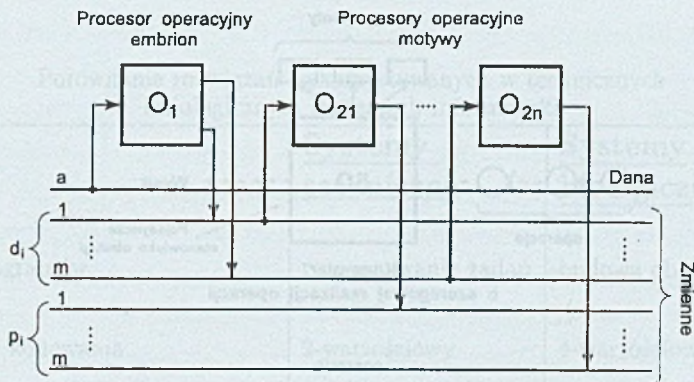
Struktura informatycznego systemu realizującego ten algorytm i sterowanego przesyłem komunikatów jest przedstawiona na rys. 7.

W systemach informatyki istniejących w organizmach biologicznych komórki, które można przyrównać do mikrokomputerów realizujących programy zapisane w tych łańcuchach DNA, nie są stanowiskami izolowanymi, autonomicznymi, lecz węzłami ogólnorganizmowej sieci komórkowych mikrokomputerów, w której przesyłane są komunikaty zewnętrzne (międzykomórkowe) i wewnętrzne, komórkowe w każdej pojedynczej komórce.

Komunikaty zewnętrzne przesyłane są w oparciu o media hormonalne (duże odległości) i o media neuronalne (małe odległości).

Komunikaty wewnętrzne oparte są na działaniu membran, które można przyrównać do urządzeniowej realizacji protokołów komunikacyjnych w technicznych sieciach komputerowych.





Rys. 7. Struktura systemu realizacji algorytmu (2) sterowanego przepływem komunikatów

Fig. 7. The structure of the algorithm (2) execution, controlled by the message transfers

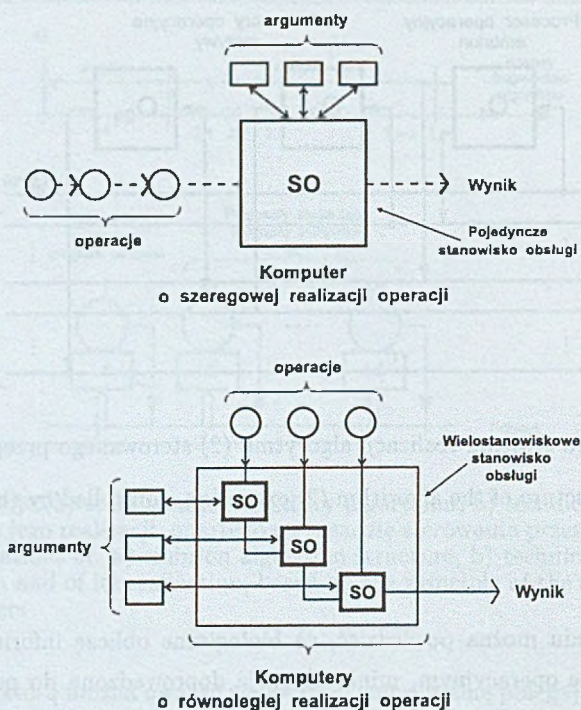
W podsumowaniu można powiedzieć, że biologiczne oblicze informatyki charakteryzuje, na poziomie operacyjnym, miniaturyzacja doprowadzona do poziomu molekuł, maksymalnie rozwinięta równoległość i sterowanie przesyłem komunikatów, a na poziomie koncepcyjnym - organizacja embrionalna.

## 5. Kierunki rozwojowe technicznych systemów informatyki

Przejdźmy teraz do charakterystyki technicznych systemów informatyki. Początek to, pochodzące jeszcze od matematyków arabskich, pojęcie algorytmu jako uporządkowanego zbioru operacji na określonych danych, takich że po ich wykonaniu otrzymuje się rozwiązanie dowolnego zadania z określonej klasy zadań. Chodziło wtedy o zadania matematyczne. Przyjmowano, że pojęcie algorytmu umożliwia podział wszystkich możliwych zadań matematycznych na trzy części. Pierwsza to zadania, dla których udało się udowodnić istnienie jednoznacznego rozwiązania i podać algorytm jego znajdowania, druga część to zadania, dla których udowodniono nieistnienie takiego algorytmu i trzecia to zadania, dla których nie udało się znaleźć ani odpowiedniego algorytmu, ani udowodnić jego nieistnienia.

Autorzy pierwszych algorytmów przyjmowali też, że realizatorem podanych w nich operacji będzie człowiek, który będzie je realizował „po ludzku”, jedną po drugiej, stąd ich





Rys. 8. Ilustracja szeregowej i równoległej koncepcji realizacji algorytmów  
 Fig. 8. Illustration of sequential and parallel realization of algorithms

zakładany szeregowy, sekwencyjny tryb realizacji reprezentowany również w idei pierwszego, abstrakcyjnego komputera, zwanego maszyną Turinga. Stąd też sekwencyjny, szeregowy tryb realizacji operacji w pierwszych komputerach elektronicznych. Dalsza ewolucja tych konstrukcji była wynikiem krytycznego spojrzenia na tę część definicji algorytmu, w której mówi się o uporządkowanym zbiorze operacji na określonych danych (argumentach). Dlaczego zbiór operacji ma być uporządkowany? Jeżeli po wykonaniu operacji na określonych argumentach otrzymujemy rozwiązanie, to po co narzucać warunek uporządkowania tych operacji. Wystarczy, aby były wykonane na wskazanych argumentach, wtedy gdy te argumenty stają się dostępne. Prowadzi to do koncepcji równoległej realizacji operacji sterowanej przesyłem komunikatów, stosowanej w maszynach równoległych, sieciach i metakomputerach. Warunek uzyskania minimalnego czasu realizacji programów złożonych prowadzi natomiast do embrionalnej organizacji procesu realizacji, tak jak w biologicznych systemach informatyki.

Trendy współczesnych kierunków rozwojowych informatyki technicznej można tak



Tabela 1

Porównanie rozwiązań wykorzystywanych w technicznych i biologicznych systemach informatyki

	<u>Systemy techniczne</u>	<u>Systemy biologiczne</u>
Cele programów	rozwiązywanie zadań	budowa obiektów
Systemy kodowania	2-wartościowy	4-wartościowy
Fizyczny zapis symboli kodowych	makromolekularny, symbole 0,1	molekularny, symbole A, G, C, T
Operacje w algorytmach	operacje matematyczne	operacje technologiczne
Wyniki realizacji programów	liczby, obrazy, teksty	obiekty materialne
Systemy operacyjne	sterowanie przesyłem operacji, sterowanie przesyłem komunikatów	sterowanie przesyłem komunikatów na najniższym poziomie równoległości
Strategia realizacji programów	różne, w tym i embrionalne	embrionalne dające największe przyspieszenie budowy obiektów
Komunikacja	kanały i protokoły komunikacji	kanały i membrany komunikacji



określić: *równoległość i miniaturyzacja, usieciowienie, multimedialność, inżynieria oprogramowania.*

Równoległość (rys. 8) znajduje swój wyraz w coraz powszechniejszym jej eksponowaniu w opracowywanych algorytmach i w architekturach projektowanych komputerów i organizacji sieci komputerowych. Duży nacisk kładzie się też na miniaturyzację.

Usieciowienie znajduje swój wyraz w przechodzeniu, w coraz większym stopniu, z układów reprezentowanych przez pojedyncze komputery na wysoko wydajne sieci komputerowe szkieletowe, federacyjne, hierarchiczne, multimedialne o różnych topologiach, mediach komunikacyjnych, regułach dostępu i bezpieczeństwie danych przesyłanych i przechowywanych w rozmieszczonych w sieci bazach danych. Podstawowymi parametrami sieci komputerowych są media komunikacyjne (elektryczne, optyczne, fale radiowe), topologie (sieć strukturalna, sieci lokalne, metropolitalne i rozległe) i metody dostępu (protokoły komunikacyjne).

Multimedialność znajduje swój wyraz w coraz większej różnorodności form, w których jest zawarta i przesyłana informacja (tekst alfanumeryczny, obraz, dźwięk, multimedialne bazy danych) i w coraz większej różnorodności branż pod uwagę technologii sprzętowych (elektronika, optyka, bionika).

Inżynieria oprogramowania znajduje swój wyraz w tym, że w zakresie programowania z indywidualnego często amatorskiego i heurystycznego opracowywania programów przez poszczególnych ludzi w coraz większym stopniu przechodzi się na prace wieloosobowych, profesjonalnych zespołów nad złożonymi programami i zbiorami programów, oparte na profesjonalnych technologiach inżynierii programowania (specyfikacja, segmentacja, scalenie, testowanie i kontrola). Trwają też prace nad rozszerzeniem charakteru danych mogących być wykorzystywanymi w programowaniu komputerów.

## 6. Wnioski

Porównanie rozwiązań wykorzystywanych w systemach informatyki technicznej i w biologicznych systemach informatyki przedstawiono w tabeli 1. Wskazuje ono na jedność podstaw naukowych, na których oparte są te oba rozwiązania, a w szczególności podstawowej koncepcji: algorytm, program, zakodowanie programu w systemie urządzeniowym, a następnie proces realizacji programu pod nadzorem systemu operacyjnego.

Warto też odnotować, że kierunki trendów rozwojowych technicznych systemów informatyki dążą do podobnych rozwiązań jak te, które cechują biologiczne systemy infor-



matyki, a mianowicie miniaturyzacja, sterowanie przesyłem komunikatów, równoległość i embrionalna organizacja procesu realizacji programu. Dążą do tych rozwiązań i to nie tylko w zakresie metod, ale i celów. Myśli się bowiem już też o konstrukcji takich technicznych systemów informatyki, w których możliwa byłaby laboratoryjna realizacja programów, w wyniku których otrzymywałoby się gotowe produkty i obiekty syntezy elementów molekularnych. Ten sposób wytwarzania produktów i budowy obiektów można by nazwać technologią informatyczną, w odróżnieniu od technologii energetycznych, zużywających bardzo duże ilości surowców i energii dla otrzymania niewielkiej ilości produktu końcowego.

W tej chwili otrzymuje się już takie produkty laboratoryjne na bazie programów „skradzionych” genom strukturalnym systemów informatyki biologicznej.

Znajomość i rozwój podstaw naukowych informatyki i kultura informatyczna dały więc nie tylko podstawy budowy komputerów, ale stają się coraz bardziej istotnym warunkiem pracy w wielu dziedzinach działalności ludzkiej, wydawałoby się czasem nic z informatyką nie mających wspólnego, jak np. biologia molekularna, gdzie dają początek takim nowym działom informatyki, jak np. bioinformatyka.

## LITERATURA

1. Watson J.D.: *Biologie moleculaire du gène*. Inter Editions, Paris 1978.
2. Węgrzyn S., Gille J.C., Vidal P.: *Genetyka procesów rozwoju*. Wydawnictwo Politechniki Śląskiej, Informatyka, z.12, 1988.
3. Węgrzyn S., Gille J.C., Vidal P.: *Developmental Systems*. Springer Verlag 1990.
4. Węgrzyn S., Kozielski S.: *Message-driven Distributed Systems for the Execution of Parallel Algorithms*. *Archiwum Informatyki Teoretycznej i Stosowanej*, 1996, tom 8, z.3-4, pp. 247-258.
5. Węgrzyn S.: *Dwa oblicza informatyki - biologiczne i techniczne*. *Pro Dialog* 8, Wyd. NAKOM Poznań 1999, pp.1-9.

Recenzent: Prof. dr hab. inż. Andrzej Grzywak

Wpłynęło do Redakcji 29 kwietnia 1999 r.



## Abstract

The article compares some solutions applied in systems of technical and biological informatics. It indicates the unity of scientific bases of the both implementations; in particular of the principle: the algorithm, the program, the coding of the program, the process of automatic execution of the program. The occurring differences, as e.g. the four-value logic, characteristic for biological systems and two-value logic proper for technical systems, arise from different goals to be achieved by the biological and technical information systems.

It is worth to note that the development trends of technical information systems are convergent to the similar features of implementations in biological systems, namely miniaturization, the control with the use of message transfers, parallelism and embryonal organization of program execution. This convergence is visible not only in methods but also in goals of the two kinds of systems.