

6

Czy powstaną maszyny autonomiczne?

Zdrowy rozsądek podpowiada, że wszelkie maszyny – w tym informatyczne i skomputeryzowane – są tylko i wyłącznie *narzędziami* w ręku ludzi realizujących za ich pomocą własne cele. Wyobraźnia z kolei roztacza przed rozumem fantastyczne wizje maszyn *autonomicznych*, które byłyby w stanie konkurować ze swoimi pomysłodawcami i konstruktorami.

Który pogląd – ten ostrożniejszy, czy ten śmielszy – ma lepsze uzasadnienie w bieżących dokonaniach informatyki? Oto pytanie, którego treść rozwiniemy w niniejszym szkicu.

§1. Komputer jako narzędzie

1.1. Jeśli spojrzeć na historię ludzkiej cywilizacji przez pryzmat rozwoju wynalazków, to niewątpliwym (choć, jak zawsze, tymczasowym) ukoronowaniem tej historii jest wynalazek komputera. Wynalazek to przedziwny. Wyjątkowość jego oddaje dobrze znane z reklamy hasło: „Wszystko w jednym”. „Wszystko” – bo chodzi o wszelkie inne maszyny informatyczne, „w jednym” – bo jeden jedyny komputer pozwala, za sprawą różnych programów, *naśladować* wszelkie inne maszyny do przetwarzania informacji (w pewnym sensie też potrafi naśladować maszyny inne, np. samochody, dając możliwość śledzenia ich pracy w wirtualnym środowisku, np. na ekranie). Za słuszością przytoczonego hasła nie trzeba specjalnie agitować – wszak nawet zwykły pecet potrafi zastąpić urządzenia tak różnorodne, jak maszyna do pisania, kalkulator czy magnetofon; mało tego, nie w tym tkwi jego siła, że urządzenia te zastępuje, lecz w tym, że poszerza znacznie ich tradycyjne możliwości.

Rzeczona wszechstronność maszyn informatycznych przekłada się gładko na kolejną ich cechę, która jeszcze silniej przydaje im blasku wyjątkowości. Chodzi o możliwość *sterowania* innymi maszynami. Skoro bowiem komputer potrafi maszyny te naśladować, to nadaje się równie dobrze do kierowania ich pracą. O tym też wiadomo z codziennej praktyki: telewizory, lodówki, pralki,

samochody – dziś niemal wszystkie z tych urządzeń kryją w sobie mikroprocesory, a te kierują w sposób pożądaný pewnymi ich funkcjami.

Wyjašnjmy drobnym drukiem wazną sprawę terminologiczną. Otóż terminu „komputer” będziemy używać w sensie maksymalnie szerokim, abstrahując od stosowanych w poszczególnych przypadkach technologii przetwarzania danych (niekiedy też, ze względu na uniwersalność komputerów w świecie maszyn informatycznych, będziemy określać komputer jako „maszynę informatyczną”).

Przeczy to wprawdzie pewnej tradycji, wywodzącej się z języka z angielskiego, zgodnie z którą komputery rozumie się wyłącznie jako komputery cyfrowe, ale oddaje dobrze istniejącą w świecie maszyn informatycznych (a więc komputerów) różnorodność. Wszak oprócz maszyn cyfrowych konstruuje się maszyny analogowe i równoległe, a myśli się o komputerach kwantowych i wykorzystujących żywą materię. Zaznaczmy też, że nawet w języku angielskim występuje ważny precedens terminologiczny, „neurocomputer”, który sugeruje, że zakres terminu „komputer” powinien być szerszy niż zakres terminu „maszyna cyfrowa” (skoro do komputerów zalicza się również oparte na sztucznych sieciach neuronowych neurokomputery).

1.2. Wiedząc o wyjątkowym statusie komputera jako narzędzia uniwersalnego i nadrzędnego wobec maszyn tradycyjnych, możemy zbliżyć się do kwestii kolejnej, dotyczącej tym razem istoty narzędzi. Otóż każde narzędzie – zarówno proste, jak i złożone – służy człowiekowi, a służąc mu rozwija jego przyrodzone możliwości.

Zachodzi zatem pytanie: jakiego rodzaju naturalne zdolności ludzkie doznają *wzmocnienia* za sprawą komputerów? Czy są to sprawności manualne, ruchowe, wzrokowe, komunikacyjne...? Z uwagi na wszechstronność komputerów narzuca się zrazu odpowiedź maksymalnie szeroka: maszyny informatyczne potrafią wzmóc wszelkie sprawności ludzkie. Wydaje się jednak, że nie trafia ona w sedno. Same komputery bowiem – choć można oczywiście sprzęgać je z innymi maszynami i można za ich pomocą symulować działanie innych maszyn – w największej mierze wzmocniają ludzkie zdolności *intelektualne*, szerzej zaś, umysłowe.

Ujmując rzecz obrazowo: podobnie jak samochód rozwija naszą naturalną umiejętność przemieszczania się, a teleskop poszerza znacznie zakres tego, co możemy zobaczyć, tak komputer wzmocnia nasz wrodzony potencjał intelektualny. Dla przykładu: wraz z wynalezieniem elektronicznych maszyn liczących (czyli pierwszych komputerów) mógł nastąpić kolosalny postęp nauk ścisłych (niewątpliwego wytworu naszych intelektów), a wiązał się ów postęp z możliwością niemal natychmiastowego przeprowadzania żmudnych obliczeń, które bez komputerów musiałyby zająć dziesiątki lat!

Obsadziwszy systemy informatyczne w typowej dla nich roli wzmocniaczy intelektu, możemy wprowadzić na scenę niezmiernie ważną w tym kontekście postać, a mianowicie *sztuczną inteligencję*. Postać tę niektórzy uznają za realną, aktualnie bądź potencjalnie; inni za całkowicie i nieodwołalnie fikcyjną (▷ zob. esej 3).

Na przekór drugiemu z tych poglądów trzeba odnotować, że we współczesnej informatyce toczą się jak najbardziej realne badania nad sztuczną inteligencją. Ich charakter oddaje bliżej następująca formuła: „badania nad automatyzacją procesu *rozwiązywania problemów* za pomocą technik informatycznych”. Ich cel zaś rysuje się dwoiście: po pierwsze, jako dążenie do wytworzenia użytecznych narzędzi, które wspomagałyby inteligencję człowieka, po drugie zaś, jako dążenie do zbudowania maszyn, które byłyby w stanie skutecznie i samodzielnie rozwiązywać problemy.

Dwoistość ta przekłada się wprost na dylemat niniejszego szkicu. Podejście pierwsze bowiem to nic innego jak wyraz przekonania w narzędziową wyłącznie potencję informatyki, a podejście drugie to objaw żarliwej wiary w możliwość powstania maszyn autonomicznych.

Oczywiście zabieg czysto językowy, czyli wybór sugestywnej (złośliwi powiedzą: chwytliwej) nazwy „sztuczna inteligencja” o niczym nie przesądza. Nie stanowi w szczególności żadnego argumentu na rzecz możliwości zaistnienia sztucznych intelektów. Dopowiedzmy nadto, że obecnie upowszechniają się nazwy konkurencyjne, jak „inteligencja maszynowa” czy „inteligencja obliczeniowa”, które to określenia „odrywają” nieco ludzką inteligencję od rzekomej inteligencji maszyn. Nie sugerują bowiem, że inteligencja ludzka mogłaby zostać sztucznie zrealizowana za pomocą inteligencji maszynowej.

1.3. Do spraw tych powrócimy już niebawem, tymczasem zaś pora wznović porzucony wątek i nazwać wyraźnie typy narzędzi, za jakie wolno uznać komputery. Liczby mnogiej nie użyliśmy przypadkiem. Sądzymy bowiem, że trzeba rozróżnić co najmniej trzy typy informatycznych narzędzi.

Na pierwszy plan wybija się sfera praktyki. W jej obrębie mieści się znakomicie formuła komputera jako *narzędzia praktycznego*, a więc takiego, które pozwala nam osiągać pewne konkretne cele, np. przechowywać dane, sterować ruchami robotów czy rozwiązywać skomplikowane równania matematyczne. To właśnie ten kierunek poszukiwań ma największe umocowanie w praktyce, a o jego przydatności i kolosalnej przyszłości rzadko kto wątpi.

Na planie drugim rysuje się dość wyraźnie domena ludzkiego poznania. Komputery bowiem pozwalają lepiej poznać to, czego doskonaleniu zdają się służyć, a więc ludzki umysł. Wolno zatem potraktować je jako pewnego rodzaju *narzędzia poznawcze*. Co za tym przemawia? Po pierwsze fakt, że komputer może pełnić w stosunku do innych maszyn podobne funkcje jak umysł (a wężiej: mózg) wobec ludzkich organów, to znaczy potrafi koordynować ich pracę. Po wtóre, budowa komputera odzwierciedla w pewnym stopniu strukturę ludzkiego umysłu – na przykład i tu, i tu występują moduły pamięci, moduły przetwarzania czy moduły wejścia/wyjścia; warto też zwrócić uwagę na neuro-komputery, których architektura jest wzorowana wprost na strukturze mózgu. Po trzecie wreszcie, co wiąże przekonująco sferę praktyki ze sferą poznania, komputery wykazują coraz większą skuteczność w podejmowaniu zadań zarezerwowanych dotychczas dla ludzi (np. logiczno-matematycznych,

ale nie tylko); to właśnie ich realizacji mają służyć wspomniane wyżej systemy inteligentne. Sumując zatem, można powiedzieć, że poznawcze, a nie tylko praktyczne, funkcje komputerów przejawiają się w tym, że mogą stanowić one adekwatny lub przynajmniej najlepszy z dotychczasowych, *model ludzkiego umysłu* (▷ zob. esej 1, §3.1; oraz esej 10, §3).

Gdzieś pomiędzy sferą praktyki a sferą poznania rozciąga się rozległy obszar doświadczenia. Jego znaczenie oddaje dobrze metafora poligonu. Podobnie bowiem jak żołnierz ćwiczy na poligonie umiejętności, które później może wykorzystać na polu walki, tak użytkownik komputera może wyćwiczyć za pomocą programów komputerowych różne umiejętności praktyczne, a ponadto – odwołajmy się tutaj do sfery poznania – może zrozumieć lepiej to, co uchwycone intuicyjnie. Stąd też jak najbardziej zasadne wydaje się postrzeganie komputerów jako *narzędzi treningowych*. I nie chodzi tylko o to, że za pomocą systemów informatycznych można symulować świat zewnętrzny oraz działające w nim urządzenia, tym samym zaś przeprowadzać za ich pomocą wirtualne treningi (co zapewniają między innymi symulatory lotów, używane powszechnie w szkoleniu pilotów). Idzie również – co być może umyka praktykom – o ludzki umysł.

Aby rzecz uwypuklić, ponownie wywołajmy na scenę sztuczną inteligencję. Próbując spojrzeć na jej zadania z nieco innej niż praktyczna perspektywy, zadajmy dwa, retoryczne chyba, pytania. Czy nie jest tak, że usiłując konstruować inteligentne systemy informatyczne (a więc wnioskujące, rozpoznające mowę czy uczące się), nie pogłębiamy – na zasadzie treningu – własnej wiedzy o czynnościach poznawczych? I czy nie jest to prawdą tym bardziej, im bardziej naszym wysiłkom konstruktorskim towarzyszy zamiar stworzenia maszyn na swój obraz i podobieństwo?

Podsumujmy zatem. Służebna rola komputera wobec ludzkiego intelektu przejawia się w trzech warstwach: praktycznej – bo komputer pozwala zautomatyzować rutynowe czynności poznawcze i usprawnić nierutynowe; poznawczej – bo znajomość komputera przybliży umysł do zrozumienia swojej własnej intelektualnej zagadki; treningowej – bo używając komputerów i projektując nowe systemy komputerowe, ćwiczymy swoje umiejętności poznawcze i praktyczne.

§2. Od kalkulatorów do maszyn interaktywnych

2.1. Zbadawszy bliżej stanowisko przypisujące komputerom funkcje wyłącznie narzędziowe, pora przyjrzeć się pogładowi przeciwnemu, zgodnie z którym współczesne maszyny informatyczne są czymś więcej, a mianowicie dobrym kandydatem na maszyny autonomiczne. Do poglądu tego zbliżymy się drogą okrężną, tj. historyczną. Postaramy się wskazać w *historii informatyki* trop,

który zdaje się wieść ku coraz większej niezależności maszyn względem ludzi i świata (▷ por. eseje 12 i 13).

Poszukiwany trop nie jest, rzecz jasna, wyraźny. Aby uczynić go widocznym, musimy już na samym wstępie odróżnić dwa odmienne, choć splecione ze sobą, czynniki postępu w informatyce. Czynnikiem pierwszy to nieustanny *rozwój sprzętu*, czyli maszynerii przetwarzania danych. Jego dynamikę znaczą kolejne przełomy technologiczne, które za każdym razem umożliwiały zwielokrotnienie szybkości przetwarzania i jeszcze dalej idącą miniaturyzację elementów przetwarzających. Przypomnijmy: w latach 40-tych XX wieku miał miejsce pierwszy przełom związany z zastosowaniem lamp próżniowych – w oparciu o nie działał m.in. ENIAC, wykonujący obliczenia numeryczne komputer-kolos o wadze 30 ton; z początkiem lat 60-tych sięgnięto do wynalazku tranzystora – dzięki niemu mogły powstać o wiele mniejsze maszyny elektroniczne (a nie, tak jak wcześniej, elektryczne); w drugiej połowie lat 60-tych wprowadzono na rynek układy scalone – za ich sprawą mogły zaistnieć mini, a potem mikro-komputery. Owoce ostatniego wynalazku, wciąż doskonalonego i miniaturyzowanego, zbieramy właściwie do dziś.

W związku z przytoczoną serią wynalazków mówi się o tzw. generacjach komputerów. W oparciu o lampy próżniowe działały komputery 1-ej generacji (lata 1940-1959); w oparciu o tranzystory pracowały maszyny 2-ej generacji (lata 1959-1964); pierwsze układy scalone wykorzystywano w maszynach 3-ej generacji (lata 1965-1970); układy scalone o wysokiej i bardzo wysokiej skali integracji stanowią podstawę przetwarzania danych w komputerach 4-tej generacji (od roku 1971). Więcej o tych sprawach można przeczytać, na przykład, w książce W. Ducha, „*Fascynujący świat komputerów*”. Polecamy ją w ramce na końcu tekstu.

2.2. Drugi z czynników stymulujących rozwój maszyn informatycznych ma z naszego punktu widzenia o wiele bardziej doniosłe znaczenie, a dotyczy nie sprzętu, lecz *oprogramowania*. Chodzi o ciąg idei, które pozwoliły – mówiąc na razie bardzo zgrubnie – przekształcić mechaniczne kalkulatory w interaktywne maszyny do przetwarzania symboli reprezentujących świat.

Idea pierwsza stanowi matematyczny fundament informatyki, a jej przesłanki są tak stare jak sama matematyka. Chodzi o myśl następującą: „obiekty matematyczne, w tym najprostsze z nich czyli liczby, nie są ważne same dla siebie (czy też dla wielbiących ich piękno matematyków), lecz nadają się dobrze do opisu świata”. Ujmując rzecz z drugiej strony: „świat poddaje się matematycznemu opisom”, a wyrażając to samo w konwencji informatycznej: „interesujące nas fragmenty świata możemy odzwierciedlić w pewnym kodzie liczbowym”.

I w tym właśnie, ostatnim, sformułowaniu zawiera się kluczowa dla informatyki, a obecna od dawna w matematyce stosowanej, idea *kodowania liczbowego*. Komputer bowiem, zgodnie z używaną niegdyś nazwą „maszyna matematyczna”, to nic innego jak urządzenie operujące na liczbach – obiekty

przezeń przetwarzane (o bojętnie czy chodzi o teksty, dźwięki czy obrazy) są kodowane jako liczby, a operacje na tych obiektach (np. odtwarzanie sekwencji dźwięków) są realizowane wewnątrz komputera jako obliczenia (▷zob. też esej 10, §3; oraz esej 17, §3). Tak przedstawia się sprawa od strony koncepcyjnej. Od strony technicznej natomiast koncepcja kodowania cyfrowego musiała zostać uzupełniona o pewne czysto fizyczne techniki odzwierciedlenia liczb (np. 0 i 1) w postaci pewnych stanów układów fizycznych, takich jak lampy próżniowe czy tranzystory (o czym wspominaliśmy wyżej).

Uważny czytelnik mógłby zażądać od autora pewnego uszczegółowienia, a mianowicie zastąpienia ogólnej idei kodowania liczbowego bardziej konkretną (i częściej wykorzystywaną w informatyce) ideą kodowania binarnego. Wszak wiadomo, że współczesne komputery operują na bitach (zerach i jedynkach) i w takiej właśnie binarnej postaci przetwarzają liczby reprezentujące zewnątrz-komputerowy świat. Sugerowanego uszczegółowienia nie przeprowadziliśmy jednak celowo. Nie chcieliśmy bowiem przeczyć faktu, że prócz najpopularniejszych dziś maszyn cyfrowych, operujących na kodzie binarnym, konstruuje się również maszyny analogowe (u zarania informatyki były one nawet bardziej powszechne), a te mogą operować – przynajmniej teoretycznie – na kodzie rzeczywisto-liczbowym.

Idea kolejna, związana tym razem wprost z programowaniem, a nie tylko z dającym mu początek kodowaniem, to idea *programu składowanego w pamięci*. Zawdzięczamy ją zespołowi wynalazców pracujących w latach 40-tych XX wieku pod kierunkiem wspomnianego już wielokrotnie Johna von Neumanna (▷zob. esej 12, §3). Idea ta sprawiła, że komputer stał się urządzeniem w pełni *automatycznym*, a nie tak jak jego pierwsze realizacje – ręcznie konfigurowanym. Chodziło o to, by zarówno kod danych wejściowych, jak i kod sterujący kolejnością obliczeń, czyli komputerowy program, mógł być zapisywany (a także modyfikowany i kasowany) w tym samym module elektronicznym, podzielonym na komórki o ustalonych adresach. Moduł ten nazwano pamięcią.

Idea, o której mowa, była obecna już w poprzedzającej narodziny maszyn elektronicznych koncepcji uniwersalnej maszyny Turinga. Taśma takiej maszyny mogła zawierać bowiem i dane wejściowe, i kod tabeli rozkazów symulowanej w danej chwili, konkretnej maszyny Turinga (▷zob. esej 10, §2.1). Rzecz jednak w tym że dopiero zespół von Neumanna dał owej przed-inżynierskiej intuicji realny kształt.

W celu efektywnego wykorzystania pamięci wprowadzono dwa rewolucyjne rozwiązania. Przede wszystkim, podzielono maszynę na cztery moduły o odrębnych funkcjach: 1) *urządzenia wejściowe* – za pomocą których użytkownik mógł wprowadzać dane i zakodowane instrukcje programów, 2) *moduł pamięci* – w którym dane i instrukcje były przechowywane, 3) *urządzenia wyjściowe* – za pośrednictwem których użytkownik mógł zapoznać się z cząstkowymi i ostatecznymi wynikami obliczeń, oraz 4) *jednostkę centralną* – która koordynowała wykonywanie programu. Podział ten zwie

się dziś *architekturą von Neumanna*. Po wtóre, ponieważ w pamięci były przechowywane i dane, i rozkazy, opracowano specjalną organizację odwołań do pamięci. Wprowadzono mianowicie dwa cykle pracy: w cyklu czytania instrukcji jednostka centralna interpretowała dane pobierane z pamięci jako instrukcje i przygotowywała się do ich wykonania, w cyklu wykonywania instrukcji – zawartość pamięci była interpretowana jako dane i do danych tych stosowano uprzednio wczytane instrukcje.

2.3. Automatyzacja przetwarzania danych za pomocą programów składowanych w pamięci stanowiła pierwszy krok na drodze od kalkulatorów, czyli maszyn realizujących „suche” obliczenia, do maszyn bardziej uniwersalnych, czyli przetwarzających symbole reprezentowane przez liczby. Krok drugi wiązał się z arcyważnym spostrzeżeniem. Otóż siła komputera, jako maszyny użytkowanej przez człowieka, musi zależeć nie tylko od sprawności jednostki centralnej, nie tylko od pojemności pamięci, lecz także (a może: przede wszystkim) od działania urządzeń wejścia/wyjścia, czyli od jakości komunikacji z użytkownikiem. Chcąc przekuć to spostrzeżenie w inżynierski czyn, powołano do życia ideę *interfejsu*.

Interfejsy, czyli specjalne programy ułatwiające komunikację na linii człowiek-maszyna, sprawiły, że komputery stały się dostępne nie tylko dla uprzywilejowanej kasty programistów, lecz także dla zwykłych ludzi (gdyby ktoś – wobec normalnej dziś łatwości kontaktów z maszyną – miał trudność z rozszyfrowaniem znaczenia nieco technicznego słowa *interfejs*, spieszmy donieść, że interfejsem jest, na przykład, układ okienek graficznych, poprzez które przemawia do nas system Windows). Owa dostępność zaś, czy wręcz przyjazność komputerów, spowodowała, że stały się one w jeszcze większym stopniu maszynami do *przetwarzania symboli* (obiektów) niż maszynami do obliczeń. Mówiąc technicznie, właściwa pierwszym komputerom funkcja obliczania stała się teraz funkcją pomocniczą, służącą w największej mierze do przetwarzania danych symbolicznych – kodowanych wprawdzie jako liczby, lecz symbolizujących obiekty uchwytnie zmysłowo (jak obrazy i dźwięki). Ów nowy, niematematyczny image maszyny informatycznej musiał przypaść do gustu szerokiemu ogółowi; ogół bowiem ma do czystych rachunków zamiłowanie raczej średnie.

Bycie maszyną w pełni interaktywną wymaga, rzecz jasna, czegoś więcej niż działającego automatycznie, łatwego w obsłudze i ergonomicznego interfejsu. Wymaga jeszcze czegoś, co nazywa się fachowo *sprzężeniem zwrotnym*, a zapewnia maszynie zdolność reagowania na zmieniające się oczekiwania użytkownika (szerzej zaś: środowiska). Wyobraźmy sobie taką oto sytuację. Używamy nowoczesnego interfejsu, który potrafi rozpoznawać polecenia głosowe i odpowiednio na nie reagować, np. wypowiedzenie zdania „Zamknij system” powoduje wyłączenie komputera. Pewnego razu siadamy jednak do komputera z lekkim przeziębieniem, któremu to przeziębieniu towarzyszy chryпка zniekształcająca głos. W odpowiedzi na taką zmianę nie-

zawodny dotychczas komputer raportuje „nieoczekiwany problem” i nie wykonuje wypowiedzianego przez nas polecenia. Czy mamy prawo uznać go za interaktywny? Oczywiście nie. Oczekujemy raczej, że maszyna dostosuje się jakoś do nowej sytuacji – poprosi, na przykład, o wpisanie polecenia z klawiatury i kilkakrotne wypowiedzenie go zmienionym głosem. Spodziewamy się też, że w efekcie tych czynności maszyna nauczy się nowego brzmienia rozkazu i w przyszłości będzie na nie reagować poprawnie. Innymi słowy oczekujemy, że komputer sam się przeprogramuje.

Podany przykład ma obrazować czwartą już ideę informatyczną, która zgodnie z naszym rozeznaniami przyczynia się do ewolucji komputerów. Mamy na myśli ideę *samoprogramowania*, czyli automatyzacji procesu doskonalenia komputerowych programów. Współcześnie trwają intensywne prace nad wcieleniem jej w życie, a powstające w wyniku tych prac programy są nazywane uczącymi się (▷ zob. esej 5). W przyszłości – o ile wspomniany koncept znajdzie efektywne rozwiązanie – samodoskonalące się programy będą musiały stanowić jądro systemów autonomicznych.

Na marginesie naszych poszukiwań – poszukiwań tropu wiodącego ku maszynom autonomicznym – trzeba wspomnieć o jednej jeszcze idei, która zrewolucjonizowała świat komputerów: o idei *przetwarzania sieciowego*. Jej dobrze znanym rozwinięciem w skali makro jest Internet.

W skali mikro natomiast rozwinięciem równie efektywnym są *sztuczne sieci neuronowe*. Przypomnijmy: są to układy przetwarzające dane równoległe, za pomocą sztucznych neuronów, przypominających rzeczywiste komórki w systemie nerwowym człowieka. Z uwagi na pewne ich szczególne własności – np. potencjalną analogowość i zdolność do samoorganizacji – upatruje się w nich realnej konkurencji dla komputerów cyfrowych. Tym samym zaś jawią się one jako poważny kandydat na systemy autonomiczne (▷ zob. koniecznie esej 4).

Jak dobrze wiadomo Internet, będąc dla komputerów czymś w rodzaju gigantycznej infrastruktury, uwolnił drzemiący w nich potencjał; zmienił też stosunek człowieka do informacji i przetwarzających ją maszyn. Napisano o tym tomy i w niniejszym esej nie ma potrzeby zagadnień tych rozwijać. W kontekście naszych poszukiwań warto jednak zwrócić uwagę na dwie kwestie. Po pierwsze, za sprawą powszechności Internetu doznała wzmocnienia i zwielokrotnienia idea maszyny interaktywnej – za pośrednictwem sieci bowiem, oferującej dostęp do tysięcy usług rozproszonych na tysiącach maszyn, wzrósł ludzki apetyt na informację udostępnianą interaktywnie. Po drugie wreszcie, środowisko sieciowe wprowadziło do świata samych komputerów, a nie tylko ich użytkowników, zupełnie nową jakość – komputery osadzone w globalnej multi-sieci tworzą bowiem coś na kształt elektronicznej populacji; w tej zaś mogą zachodzić zupełnie nowe zjawiska, na przykład elektroniczna ewolucja.

§3. W stronę komputerowej autonomii

3.1. Spróbujmy podsumować to, co powiedziano wyżej w kontekście historycznym. Oto równolegle z postępowaniem technologicznym, skutkującym maszynami coraz szybszymi i coraz bardziej zminiaturyzowanymi, dokonuje się wciąż postęp inny, wyznaczany przez pewne przełomowe idee dotyczące organizacji przetwarzania danych. Polega on na doskonaleniu metod automatycznej interakcji z użytkownikiem (a szerzej: ze środowiskiem zewnętrznym).

Trzeba przyznać, że jakkolwiek proces ten faktycznie ma miejsce, a w jego wyniku maszyny informatyczne stają się coraz bliższe ludziom i coraz ściślej „zespalone” ze środowiskiem, to wykazują one zasadniczą słabość. Posługując się pewną grą słów, powiedzielibyśmy, że choć nazywa się je interaktywnymi, to są one bardziej *reaktywne* niż aktywne. Mówiąc inaczej, w swoich kontaktach ze światem nie wykazują one samorzutnej aktywności, a jedynie bierne (tj. zaprogramowane) reakcje, których określone sekwencje mają przybliżyć do celu wyznaczonego przez programistę lub narzuconego przez środowisko.

I w tym właśnie miejscu dochodzimy do idei *maszyn autonomicznych*, a więc zgodnie ze słownikowymi znaczeniami terminu „autonomia” do idei maszyn niezależnych i niezawisłych; a wtapiając te znaczenia w tematykę eseju: do idei maszyn nie będących wyłącznie naszymi narzędziami. Choć maszyny takie jeszcze nie powstały, możemy wyobrazić sobie, na czym będzie polegać ich specyfika. Będzie ona wyrażała się w połączeniu automatyzacji działań ze zdolnością do samorzutnego oddziaływania na środowisko (którego elementem jest człowiek). Wracając zatem do gry słów z terminem „aktywne” w roli głównej, powiemy że:

Systemy autonomiczne winny być nie tylko reaktywne, tj. kształtowane przez środowisko, i nie tylko interaktywne, tj. zdolne do realizacji celów komunikowanych im przez użytkownika, lecz przede wszystkim aktywne, tj. zdolne generować własne cele działania i zgodnie z nimi kształtować środowisko.

Trzeba przyznać, że nazywanie maszyn „autonomicznymi”, na podobieństwo polityków (w domyśle: samodzielnych) i regionów politycznych (w domyśle: samorządnych), nie jest praktyką powszechną. Częściej stosuje się w tym kontekście określenia psychologizujące, takie jak „maszyna myśląca” czy „maszyna inteligentna”.

W naszym przekonaniu jednak apsycho logiczny termin „autonomia” trafia w sedno intencji oddzielenia od siebie dwóch kwestii: (1) kwestii podobieństwa maszyny i człowieka (maszyna myśląca ma być w swoim myśleniu podobna do człowieka), oraz (2) kwestii niezależności maszyny od człowieka (maszyna autonomiczna ma być po prostu niezależna od człowieka). Oddzielając te kwestie, uwalniamy się od poglądu, zgodnie z którym maszyny autonomiczne miałyby myśleć i działać podobnie do ludzi.

3.2. Czy maszyny autonomiczne kiedykolwiek powstaną? Zgodnie z deklarowaną intencją eseju, który miał służyć raczej rozwinięciu treści tego pytania niż jego rozstrzygnięciu, wyrazimy w tym miejscu jedynie nasze intuicyjne przypuszczenia.

Na początek możemy przyjąć śmiało, że pewne warunki konieczne autonomii zostały już spełnione, a rychłemu spełnieniu innych sprzyja dotychczasowy bieg historii informatyki. Przypomnijmy, że:

- (i) już współczesne maszyny są w stanie *odzwierciedlać* świat zewnętrzny i *oddziaływać* na niego – a to dzięki idei kodowania liczbowego;
- (ii) działają w pełni *automatycznie* – a to dzięki idei programu składowanego w pamięci;
- (iii) są częściowo *samoprogramowalne* – a to dzięki programom uczącym się;
- (iv) wykazują szeroki zakres *interakcji* ze światem zewnętrznym – a to dzięki coraz sprawniejszym interfejsom i (ponownie) programom uczącym się;
- (v) coraz sprawniej *komunikują się* ze sobą – a to dzięki środowisku sieciowemu;

Aby dokończyć naszą wyliczankę, a więc zapewnić komputerom możliwość samodzielnego formułowania własnych celów czy też samodzielnego stawiania nowych problemów, trzeba wskazać kolejną ideę, która uzupełniając i dopełniając poprzednie, wprowadziłaby do świata komputerów nową jakość. Nie trudno zgadnąć o jaką to jakość chodzi. Nowy wartościowy cel, nowy ciekawy pomysł, skuteczne uczenie się, skuteczna interakcja z otoczeniem – to wszystko wymaga pomysłowości, zwanej także *inwencją* (por. esej 19, §1.2).

Jaka jednak droga mogłaby wieść do inżynierskiej realizacji inwencji? Wydaje się, że istnieje tu jedna tylko odpowiedź, którą podpowiada obserwacja rezultatów ludzkiej pomysłowości. Otóż rezultaty te są tyleż użyteczne, co *nieprzewidywalne*. Nieprzewidywalność zaś, jeśli nie wynika wprost z naszej niewiedzy, ma swoje źródło w zdarzeniach losowych; począwszy od tak prostych jak rzut kostką do gry, a skończywszy na tak złożonych jak powstanie nowego genotypu. O zdarzeniach takich wiadomo, że... nie do końca wiadomo, jaki dadzą wynik; jednocześnie jednak istnieją w nauce wyrafinowane metody ich opisu i szacowania związanych z nimi prawdopodobieństw (czym zajmują się różne działy probabilistyki).

Gwoli ścisłości trzeba dodać, że chodzi o nieprzewidywalność kontrolowaną – kontrolowaną na poziomie społecznym. Znaczy to, że efekt czyjejs pomysłowości ma szansę przetrwać, tj. upowszechnić się, o ile zyska ogólną akceptację, nie wyłamuje się zbyt mocno poza ogólnie przyjęte normy kulturowe, nie zagraża bezpieczeństwu innych itp. Zauważyć też trzeba, że kontrola nad nieprzewidywalną inwencją komputerów byłaby w pewnym

sensie możliwa już dziś – a to dzięki zapewniającemu międzykomputerową komunikację środowisku sieciowemu.

3.3. Zakładając zatem, że u podłoża skutkującej nieprzewidywalnością inwencji leżą pewnego rodzaju zdarzenia losowe, dokładając do tego myśl, że komputery realizują na najgłębszym poziomie pewne obliczenia, możemy przypuścić, że za ich hipotetyczną pomysłowość będą odpowiadały *obliczenia inicjowane losowo*. Co ciekawe, pomysły związane z tego rodzaju obliczeniami już się realizuje. Wspomnijmy o dwóch.

Myśl pierwsza ma charakter zdecydowanie sprzętowy, a nawiązuje do zjawisk losowych zachodzących w mikroświecie (np. na poziomie cząstek elementarnych), tj. *zjawisk kwantowych*. Zgodnie z nią zjawiska kwantowe można wykorzystać wewnątrz komputera, np. projektując mikro-bramki logiczne przetwarzające tzw. kubity – odpowiedniki zwykłych bitów, wykraczające jednak swoją zawartością poza klasyczne 0 i 1, a realizowane w praktyce za pomocą cząstek elementarnych. Istotą maszynierii przetwarzającej kubity (opartej np. na kwantowych bramkach logicznych) jest typowa dla mikroświata nieprzewidywalność wyniku pojedynczej operacji. Oto odnośne objaśnienie specjalisty: „*Pojedynczy wynik obliczeń komputera kwantowego będzie niepewny. Istotne staje się wykonanie całej serii obliczeń i dopiero ich średnia wartość z dużą dokładnością określi prawidłowy wynik – tym dokładniejszy, im więcej komputer dokona obliczeń (...)* Oznacza to, że uruchamiając ten sam program na komputerze kwantowym dwukrotnie, można by było otrzymać zupełnie różne wyniki ze względu na losowość procesu kwantowego pomiaru” [źródło: Wikipedia, hasło komputer kwantowy].

Ekstrapolując powyższe objaśnienie na interesujące nas tu zagadnienie komputerowej inwencji, moglibyśmy zaryzykować przypuszczenie, że choć kwantowe technologie informatyczne są jeszcze w powijakach, i choć dziś myśli się raczej o przyspieszeniu za ich pomocą tradycyjnych obliczeń (w domyśle: precyzyjnych), to w przyszłości mogą one przydać komputerom blasku oryginalności. Istotą tych technologii jest bowiem losowość, a wraz z nią nieprzewidywalność wyników wewnątrz-komputerowych mikrooperacji. A jeśli tak, to nic nie stoi na przeszkodzie, by tę właśnie ich naturalną właściwość wykorzystać do generowania przez maszynę pewnych przypadkowych mikro-wyników, których kumulacja mogłaby zaowocować jakimś nowym, zaskakującym pomysłem.

Myśl druga, niekoniecznie alternatywna w stosunku do pierwszej, tyczy się raczej programowania komputerów (nawet tych tradycyjnych), a pochodzi z biologii. Chodzi o tzw. *obliczenia ewolucyjne*. Obliczenia takie przebiegają zgodnie ze schematem inspirowanym przez biologiczną teorię ewolucji – teorię, która wyjaśnia naturalny rozwój organizmów za pomocą elementów stricte losowych, jak mutacje, przypadkowe krzyżówki i po części losowa selekcja.

Więcej informacji na ten temat zawierają eseje 5 i 8, „Czy roboty powinny się uczyć?” i „Czy maszyny przyszłości będą ewoluować?”. Powołując się na ich treść, możemy stwierdzić, że obliczenia ewolucyjne pozwalają wprowadzić do programów komputerowych *ukierunkowany indeterminizm*. Znaczący to, że pewne ich mikro-elementy, czyli pewne obliczenia podstawowe, są nieprzewidywalne (np. wskutek mutacji); ale mimo to, ogólnie i statystycznie rzecz biorąc, całość obliczeń zmierza we właściwym kierunku – kierunku określonym przez wymogi realizowanego zadania.

Oczywiście droga od programów rozwiązujących konkretne problemy (a takimi są dzisiejsze programy ewolucyjne) do super-programu, który kierowałby pracą automatu pokonującego wszelkie trudności, nie jest ani prosta, ani jasna. Istotne jednak, że informatykom udało się wynaleźć sprytną strategię rozwiązywania problemów w sposób, z jednej strony, nieprzewidywalny, a z drugiej, ukierunkowany. Czy strategia ta będzie kolejnym krokiem ku maszynowej autonomii. Tego nie wiemy. Obiecująco jednak przedstawia się właściwe tej metody połączenie informatyki z biologią, a więc tą dziedziną, która bada jedyny znane nam byty autonomiczne, tj. organizmy żywe.



WARTO PRZECZYTAĆ

Duch W., *Fascynujący świat komputerów*, Wydawnictwo NAKOM, 1997.

Rheingold H., *Narzędzia ułatwiające myślenie. Historia i przyszłość metod poszerzania możliwości umysłu*, Wydawnictwa Naukowo-Techniczne, Warszawa 2003.

Penrose R., *Cienie umysłu. Poszukiwanie naukowej teorii świadomości*, Zysk i S-ka Wydawnictwo, Poznań 2000