

OBLICZENIOWY MODEL MÓZGU

Hubert Wątorczyk, Maciej Zawisz, Szymon Żylski

SPIS TREŚCI

Maszyna Turinga

Sztuczna inteligencja

Klasyczna obliczeniowa teoria umysłu

Sieci neuronowe

Relacja pomiędzy sieciami neuronowymi
a klasycznymi obliczeniami

Koneksjonizm

Neuronauka obliczeniowa

Alternatywne koncepcje obliczeń

Argumenty przeciwko obliczeniom

WPROWADZENIE

Czy maszyna może myśleć?

Czy umysł sam w sobie może być myślącą maszyną?

Rewolucja komputerowa otworzyła dyskusje na ten temat oferując perspektywy dla maszyn, które naśladują m.in. rozumowanie, podejmowanie decyzji, rozwiązywanie problemów, percepcję i inne procesy umysłowe.

Postęp technologiczny potwierdził, że umysł jest systemem obliczeniowym, nazywamy go obliczeniowym modelem mózgu.



MASZYNA TOURINGA

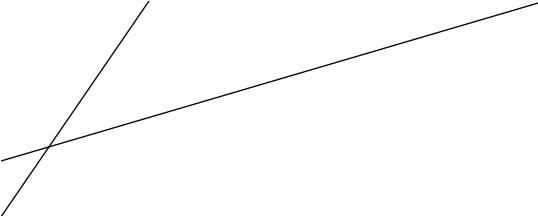
Maszyna Turinga nie jest maszyną w zwykłym sensie, ale raczej wyidealizowanym modelem matematycznym, który redukuje logiczną strukturę dowolnego urządzenia obliczeniowego do jej esencji.

Zgodnie z założeniami Turinga, maszyna wykonuje swoje funkcje w sekwencji dyskretnych kroków i w każdej chwili przyjmuje tylko jeden ze skończonej listy stanów wewnętrznych. Sama maszyna składa się z nieskończenie rozciągliwej taśmy, głowicy, która może wykonywać różne operacje na taśmie, oraz modyfikowalnego mechanizmu sterującego w głowicy, który może przechowywać wskazówki ze skończonego zbioru instrukcji.

Taśma jest podzielona na kwadraty, z których każdy jest albo pusty, albo ma nadrukowany jeden ze skończonej liczby symboli. Głowica taśmy ma możliwość przechodzenia do, odczytu, zapisu i kasowania dowolnego pojedynczego kwadratu, a także może w każdej chwili zmienić stan wewnętrzny na inny. Każdy taki akt jest uwarunkowany stanem wewnętrznym maszyny i stanem skanowanego kwadratu w danej chwili.



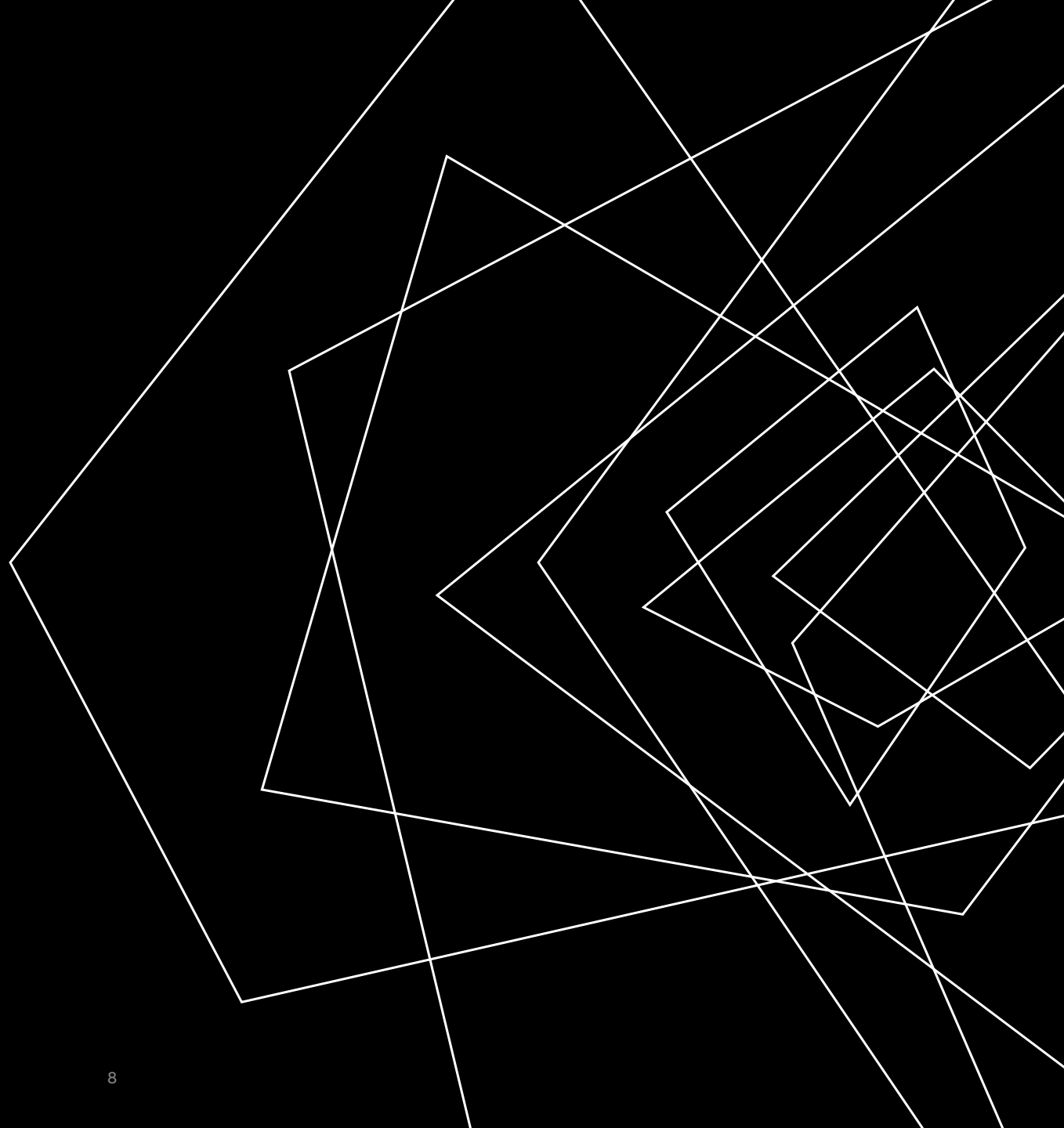
SZTUCZNA INTELIGENCJA

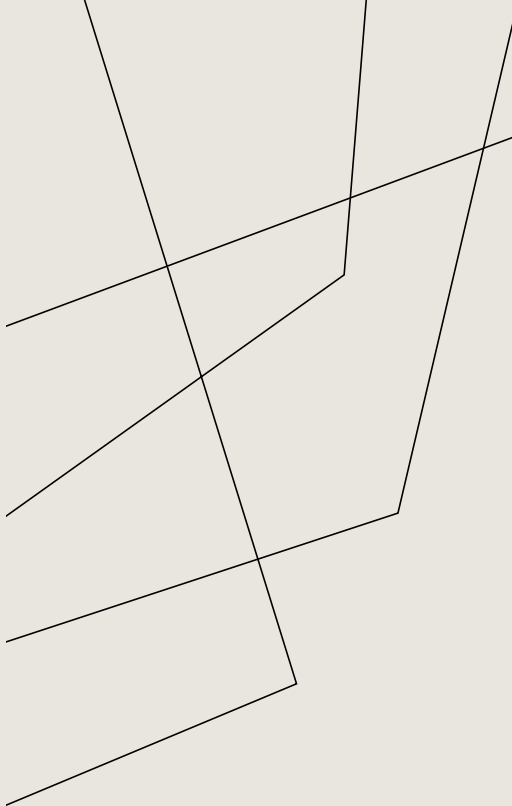


Szybki postęp w informatyce skłonił wielu, w tym Turinga, do zastanowienia się, czy można zbudować komputer zdolny do myślenia. Sztuczna inteligencja (AI) ma na celu skonstruowanie "myślących maszyn". Dokładniej, chodzi o skonstruowanie maszyn obliczeniowych, które wykonują podstawowe zadania umysłowe, takie jak rozumowanie, podejmowanie decyzji, rozwiązywanie problemów itd.

Jednym z problemów, z którymi borykały się wczesne prace nad SI, jest niepewność. Prawie każde rozumowanie i podejmowanie decyzji odbywa się w warunkach niepewności. Na przykład trzeba zdecydować, czy wybrać się na piknik, nie mając pewności, czy będzie padać. Bayesowska teoria decyzji jest standardowym matematycznym modelem wnioskowania i podejmowania decyzji w warunkach niepewności. Niepewność jest skodyfikowana poprzez prawdopodobieństwo. Precyzyjne reguły dyktują jak aktualizować prawdopodobieństwa w świetle nowych dowodów i jak wybierać działania w świetle prawdopodobieństw i użyteczności.

KLASYCZNA OBLICZENIOWA TEORIA UMYSŁU





Warren McCulloch i Walter Pitts po raz pierwszy w 1943r. zasugerowali, że coś przypominającego maszynę Turinga może stanowić dobry model umysłu. W latach 60. obliczenia Turinga stały się centralnym elementem powstającej interdyscyplinarnej inicjatywy kognitywistyki, która bada umysł, czerpiąc z psychologii, informatyki (zwłaszcza AI), lingwistyki, czy filozofii.

Według klasycznej obliczeniowej teorii umysłu, jest on systemem obliczeniowym podobnym pod istotnymi względami do maszyny Turinga, np. podstawowe procesy umysłowe (np. rozumowanie, podejmowanie decyzji i rozwiązywanie problemów) są obliczeniami podobnymi do obliczeń wykonywanych przez maszynę Turinga.

Powszechnym jest opisywanie obliczeniowej teorii mózgu jako "komputera", lepiej natomiast opisać umysł jako "system obliczeniowy". Jest to spowodowane między innymi silną sugestią, że komputer jest maszyną programowalną. Umysł natomiast podobnie jak większość maszyn Turinga nie są programowalne.

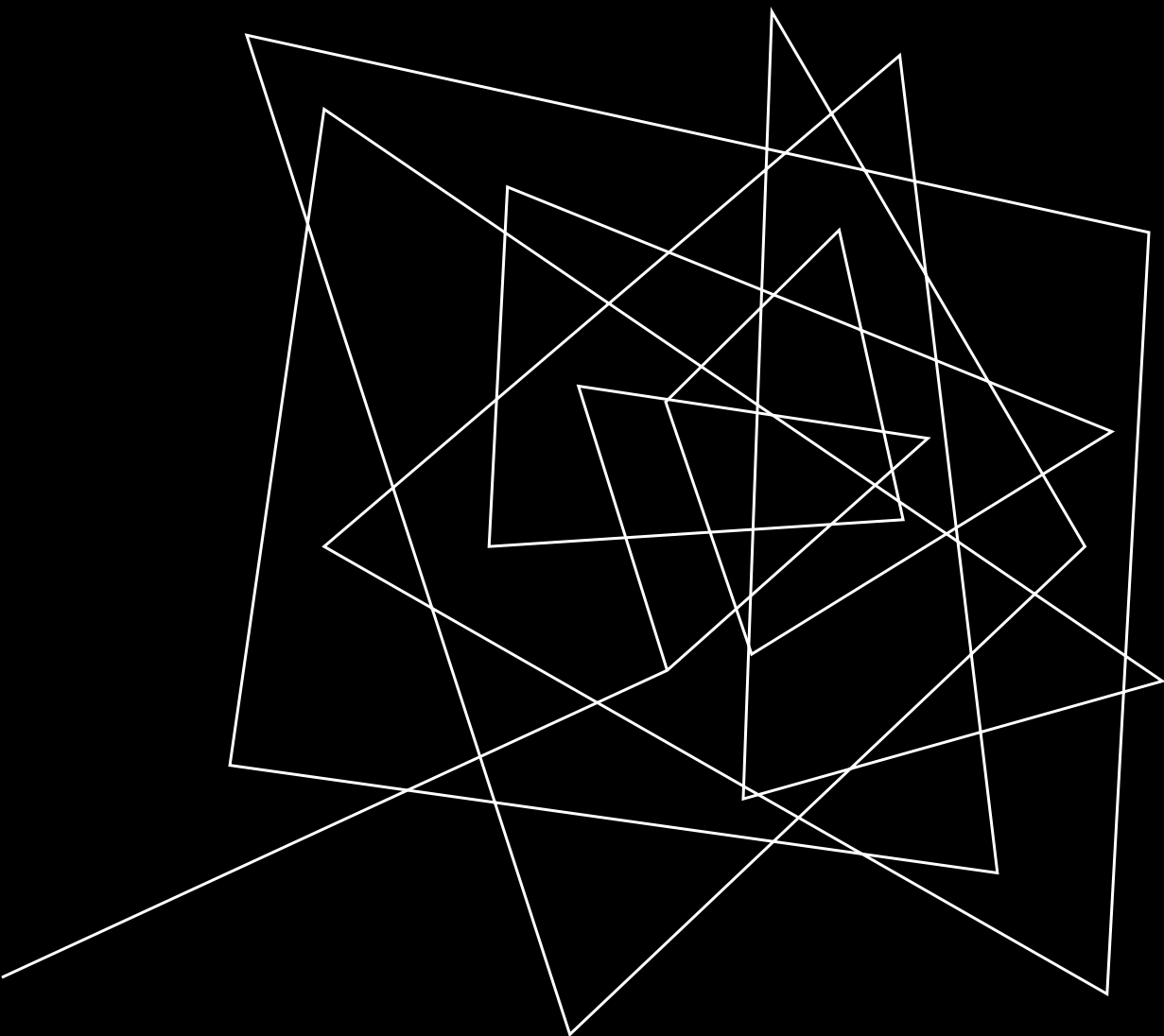
Maszyny Turinga wykonują czyste obliczenia symboliczne. Wejścia i wyjścia są symbolami wpisanymi w miejsca pamięci. W przeciwieństwie do nich umysł otrzymuje dane sensoryczne (np. bodźce z siatkówki) i wytwarza dane motoryczne (np. aktywacje mięśni).




SIECI NEURONOWE

W latach 80. koneksjonizm wyłonił się jako wybitny rywal klasycznego komputacjonizmu. Koneksjoniści czerpią inspirację raczej z neurofizjologii niż logiki i informatyki. Stosują modele obliczeniowe, sieci neuronowe, które znacznie różnią się od modeli w stylu Turinga. Sieć neuronowa jest zbiorem połączonych ze sobą węzłów. Węzły dzielą się na trzy kategorie: węzły wejściowe, węzły wyjściowe i węzły ukryte (które pośredniczą między węzłami wejściowymi i wyjściowymi).

W latach 2010. dość popularna stała się klasa modeli obliczeniowych znanych jako głębokie sieci neuronowe. Modele te są sieciami neuronowymi z wieloma warstwami węzłów ukrytych (czasem setkami takich warstw). Są one trenowane na dużych zbiorach danych za pomocą jednego lub innego algorytmu uczenia (zwykle wstecznej propagacji) - odniosły wielki sukces w wielu obszarach AI, w tym w rozpoznawaniu obiektów i strategicznych grach. Głębokie sieci neuronowe są obecnie szeroko stosowane w aplikacjach komercyjnych i są przedmiotem intensywnych badań zarówno w środowisku akademickim, jak i w przemyśle. Naukowcy zaczęli je również wykorzystywać do modelowania umysłu.



RELACJA POMIĘDZY SIECIAMI NEURONOWYMI A KLASYCZNYMI OBLICZENIAMI

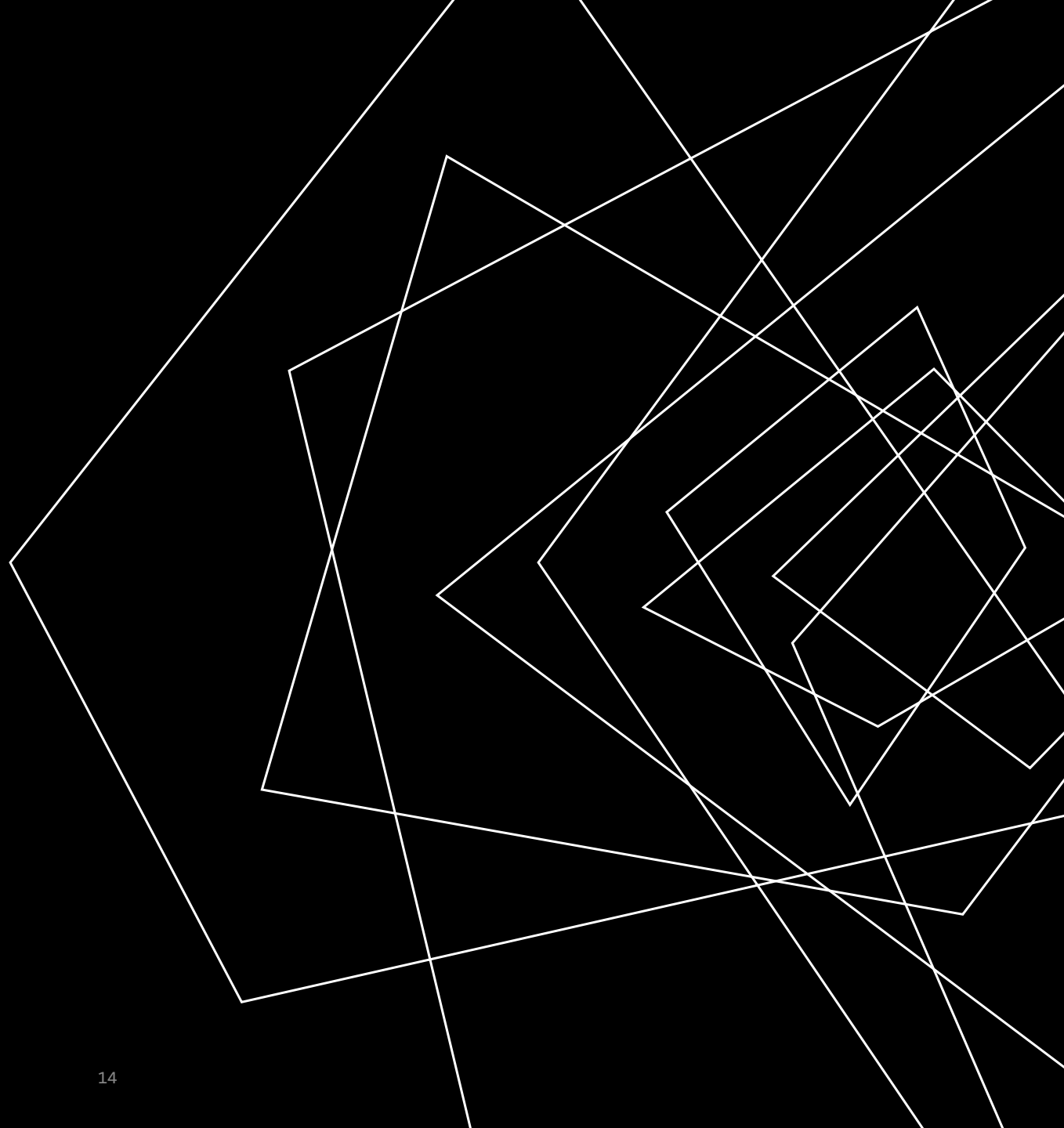


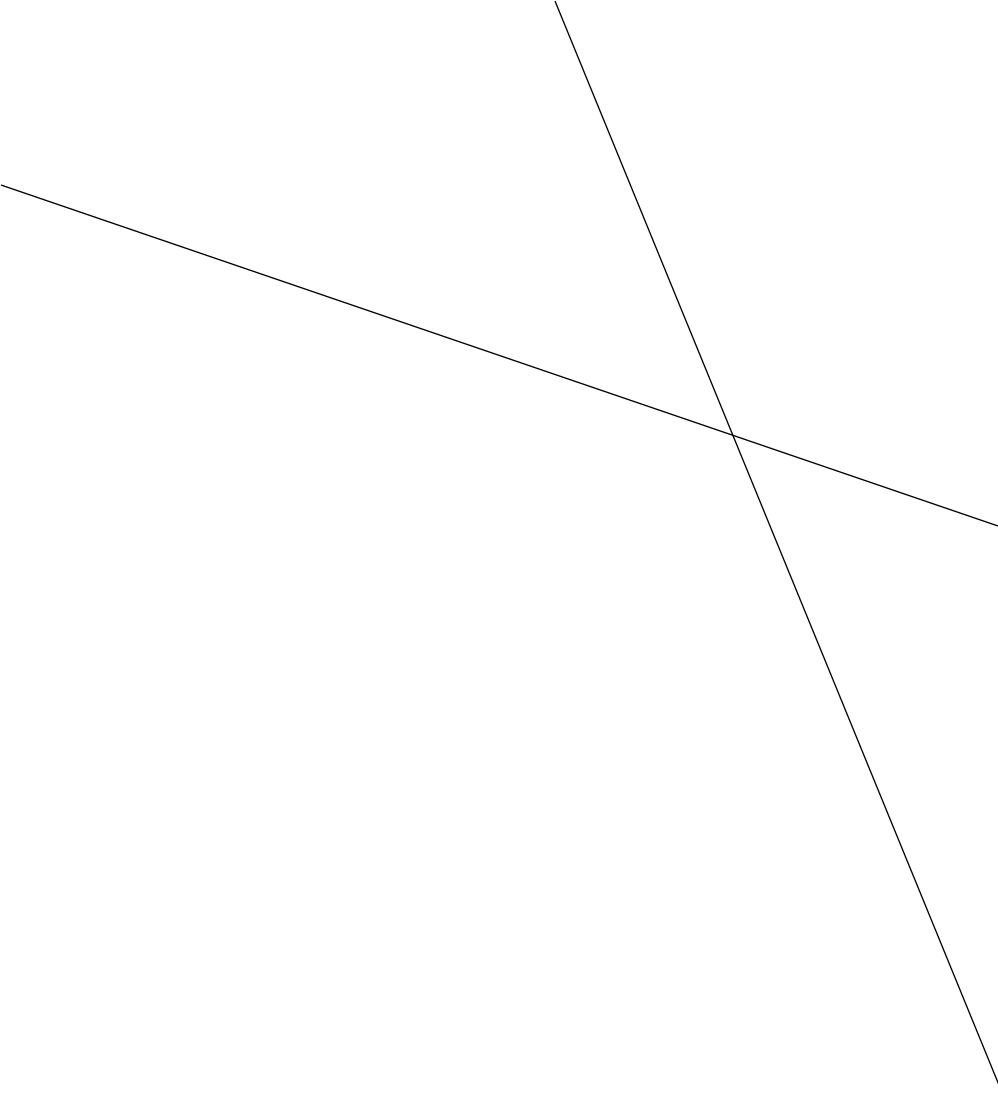
Sieci neuronowe mają zupełnie inne "odczucia" niż klasyczne modele (tzn. w stylu Turinga). Jednak klasyczne obliczanie i obliczanie sieci neuronowych nie wykluczają się wzajemnie:

Można zaimplementować sieć neuronową w modelu klasycznym. W rzeczy samej, każda sieć neuronowa, jaka kiedykolwiek została fizycznie skonstruowana, została zaimplementowana na komputerze cyfrowym. Można zaimplementować klasyczny model w sieci neuronowej. Nowoczesne komputery cyfrowe implementują obliczenia w stylu Turinga w sieciach bramek logicznych.

Filozofowie często mówią, że klasyczne obliczenia obejmują "rządzoną regułami manipulację symbolami", podczas gdy obliczenia w sieciach neuronowych są niesymboliczne. Intuicyjny obraz jest taki, że "informacja" w sieciach neuronowych jest globalnie rozłożona na wagi i aktywacje, a nie skoncentrowana w zlokalizowanych symbolach.

KONEKSJONIZM

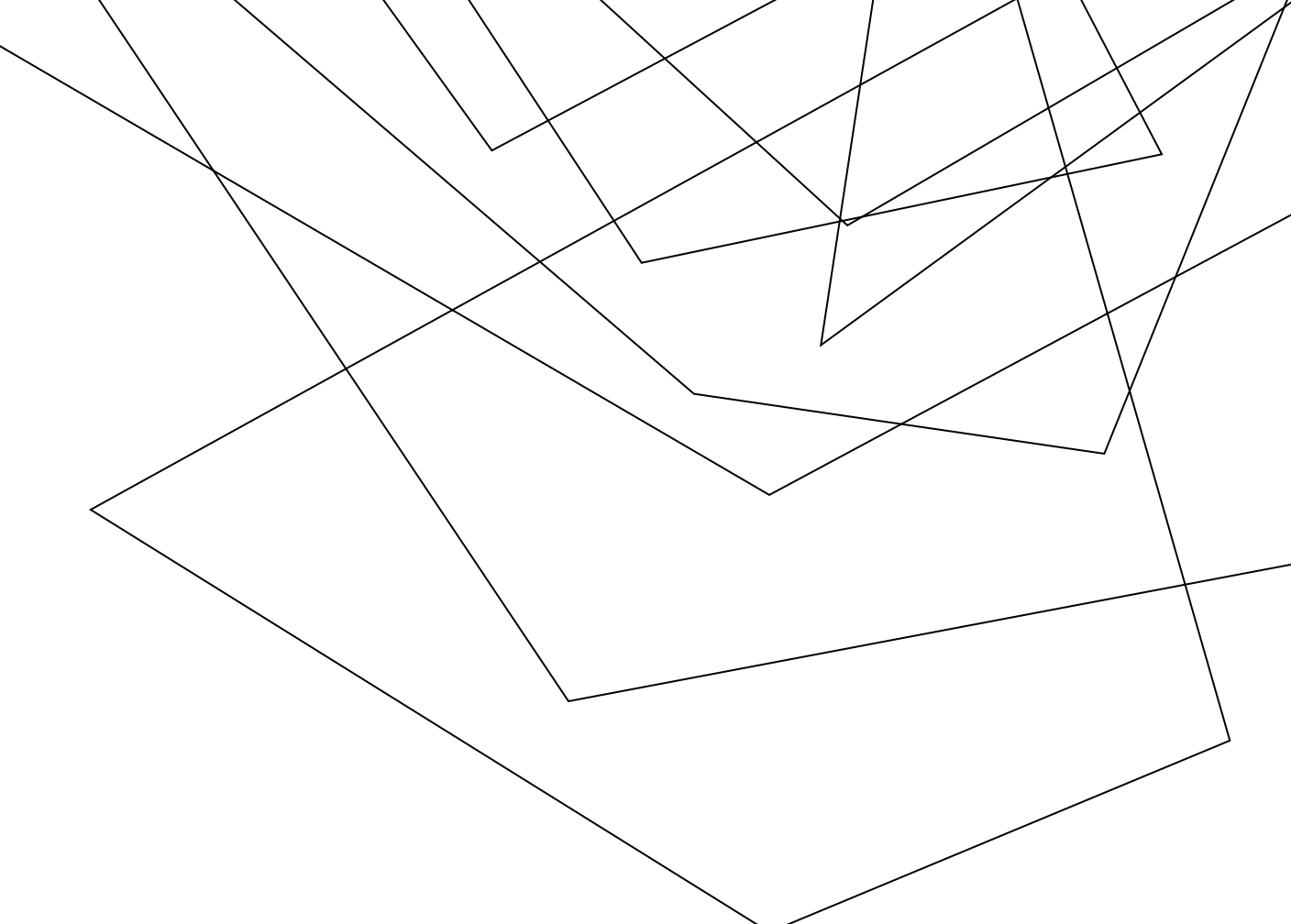




Koneksjonizm ekscytuje wielu badaczy ze względu na analogię między sieciami neuronowymi a mózgiem. Węzły przypominają neurony, a połączenia między węzłami - synapsy. Zdaniem koneksjonistów, mózg jest komputerem i bardzo szybko interpretuje on nieprecyzyjną informację sensoryczną. Tym samym odróżnia on np. szept w hałaśliwym pomieszczeniu. Koneksjonizm dąży do zastąpienia metafory komputerowej metaforą mózgową. Samo funkcjonowanie mózgu ma przebieg statyczny, dlatego też koneksjoniści uważają, iż w stanie czynnościowym procesów dane magazynują się na bardzo krótki czas.

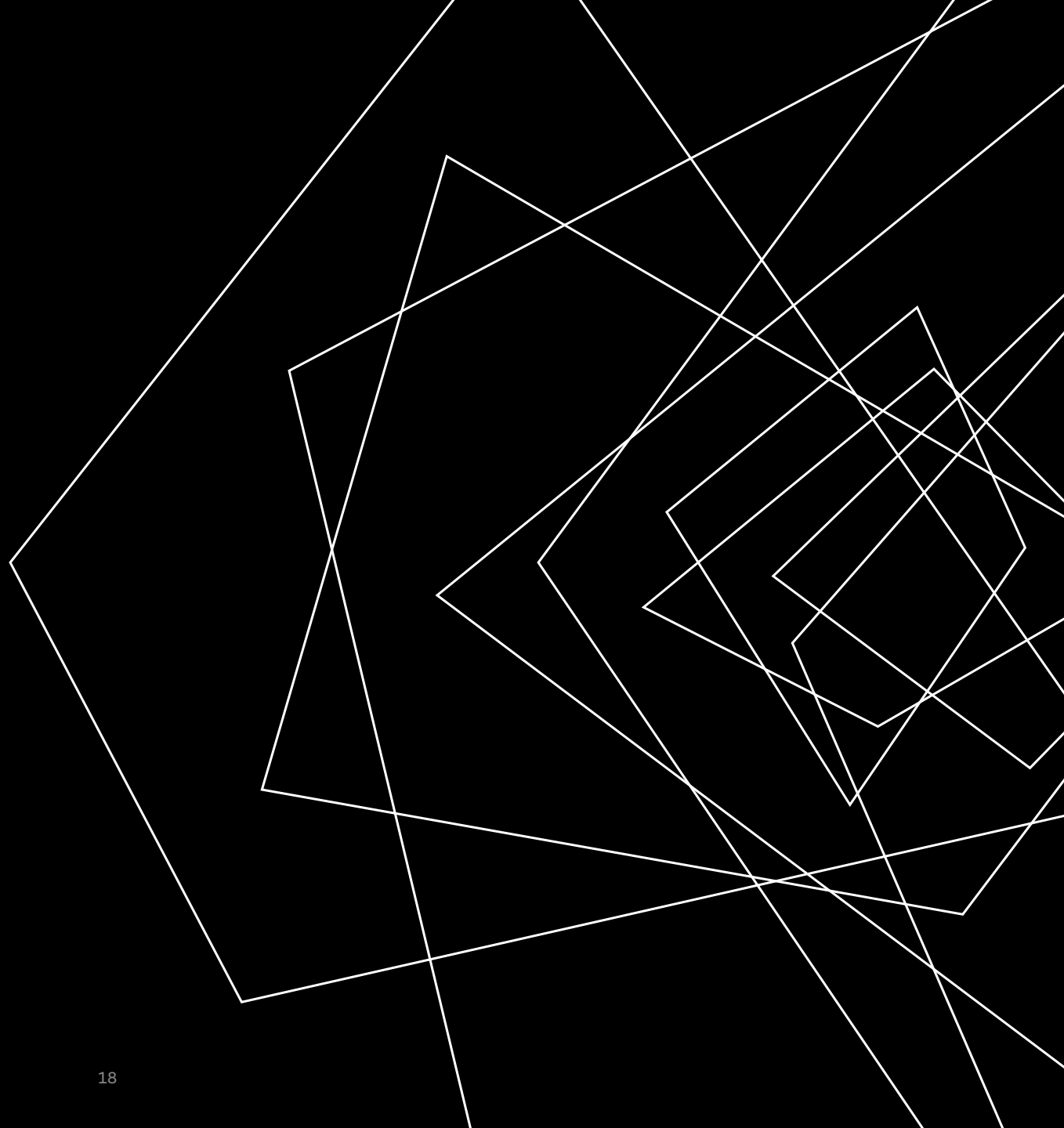


NEURONAUKA OBLICZENIOWA

The top-left portion of the slide features a complex, abstract pattern of thin black lines. These lines intersect to form various irregular polygons and shapes, creating a sense of dynamic movement and geometric complexity. The lines are scattered across the upper-left quadrant, with some extending towards the center of the slide.

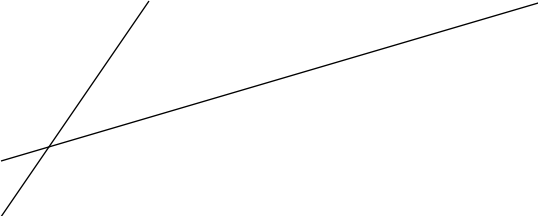
Neuronauka obliczeniowa opisuje układ nerwowy poprzez modele obliczeniowe. Chociaż ten program badawczy jest oparty na matematycznym modelowaniu pojedynczych neuronów, charakterystycznym punktem zainteresowania neuronauki obliczeniowej są systemy połączonych ze sobą neuronów. Neuronauka obliczeniowa zwykle modeluje te systemy jako sieci neuronowe. W tym sensie jest to odmiana, odrost lub potomek koneksjonizmu. Jednak większość neuronaukowców obliczeniowych nie identyfikuje się jako koneksjoniści.

ALTERNATYWNE KONCEPCJE OBLICZEŃ



Kognitywiści często opisują obliczenia jako przetwarzanie informacji. Claude Shannon wprowadził ważne dla nauki pojęcie "informacji". Intuicyjny pomysł polega na tym, że informacja mierzy redukcję niepewności, gdzie zmniejszona niepewność przejawia się jako zmieniony rozkład prawdopodobieństwa na możliwe stany. Shannon skodyfikował tę ideę w rygorystycznych ramach matematycznych, tworząc podstawy teorii informacji

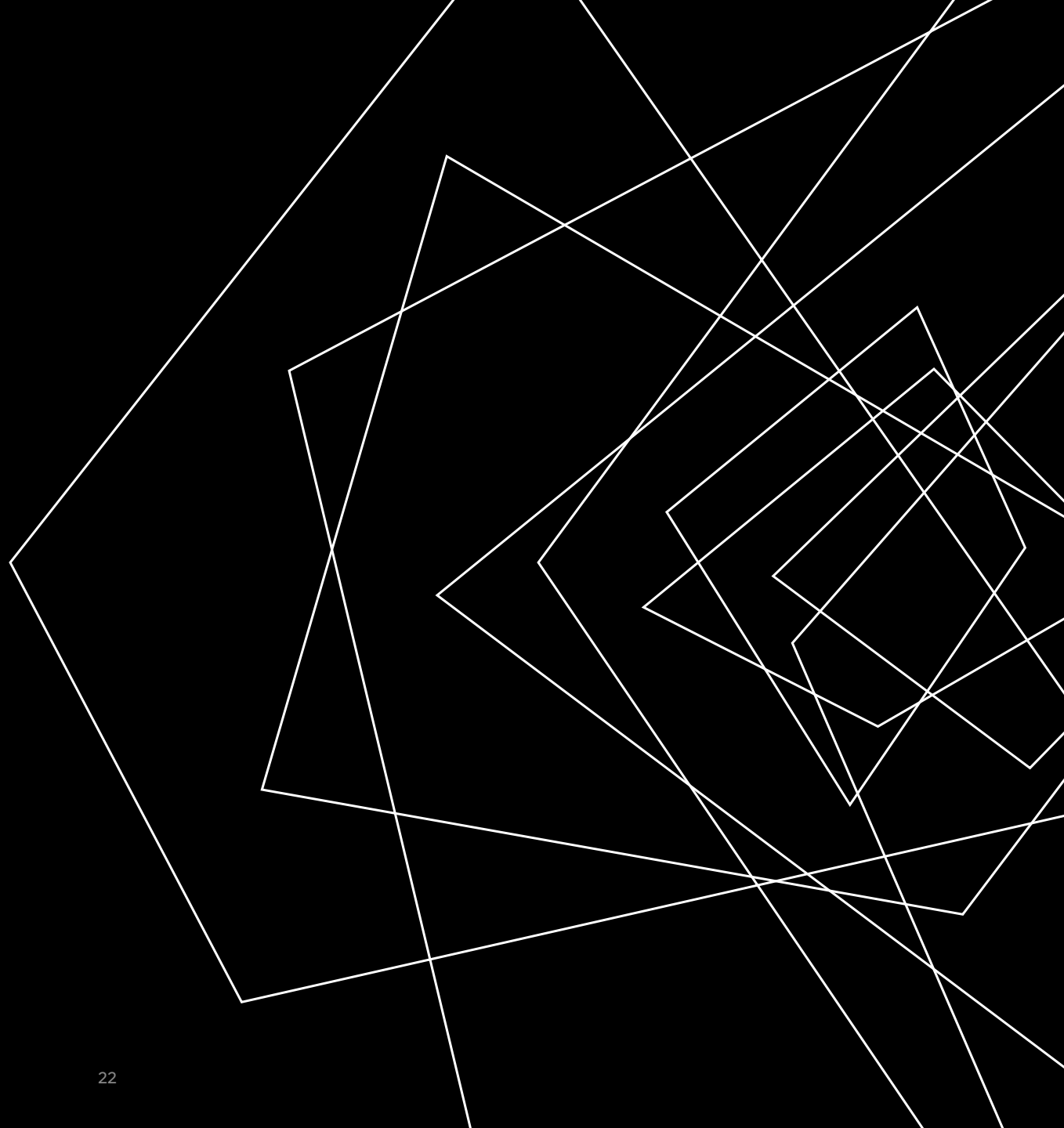
Wiele dyskusji filozoficznych ucieleśnia strukturalistyczną koncepcję obliczeń: model obliczeniowy opisuje abstrakcyjną strukturę przyczynową, nie biorąc pod uwagę konkretnych stanów fizycznych, które instancjonują tę strukturę.

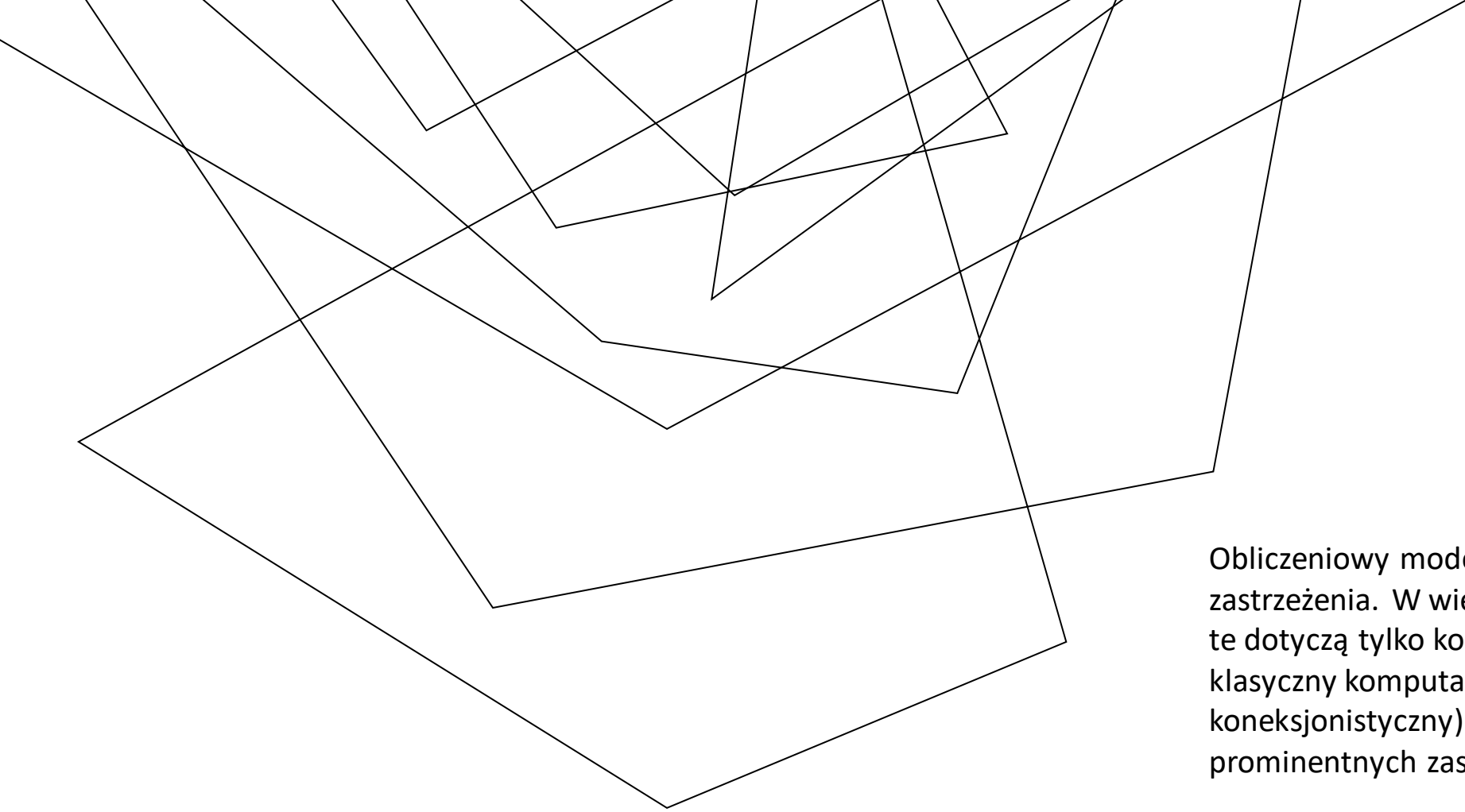


Mechanistyczna natura obliczeń jest powracającym tematem w logice, filozofii i kognitywistyce. Gualtiero Piccinini (2007, 2012, 2015) i Marcin Milkowski (2013) rozwijają ten temat w mechanistyczną teorię systemów obliczeniowych. Mechanizm funkcjonalny to system połączonych ze sobą komponentów, gdzie każdy z nich pełni jakąś funkcję w ramach całego systemu. Wyjaśnienie mechanistyczne przebiega poprzez rozłożenie systemu na części, opisanie jak części są zorganizowane w większy system i wyodrębnienie funkcji wykonywanej przez każdą część. System komputerowy jest funkcjonalnym mechanizmem szczególnego rodzaju.

Przeanalizowaliśmy różne kontrastujące, a czasem pokrywające się koncepcje obliczeń: obliczenia klasyczne, obliczenia koneksjonistyczne, obliczenia neuronowe, obliczenia formalno-syntaktyczne, obliczenia angażujące treść, obliczenia przetwarzające informacje, obliczenia funkcjonalne, obliczenia strukturalistyczne i obliczenia mechanistyczne. Każda koncepcja daje inną formę obliczeniowości. Każda koncepcja ma swoje mocne i słabe strony. Można przyjąć pluralistyczne stanowisko, które uznaje odrębne, prawomocne koncepcje. Zamiast wywyższać jedną koncepcję ponad inne, pluraliści z radością stosują każdą koncepcję, która wydaje się użyteczna w danym kontekście wyjaśniającym.

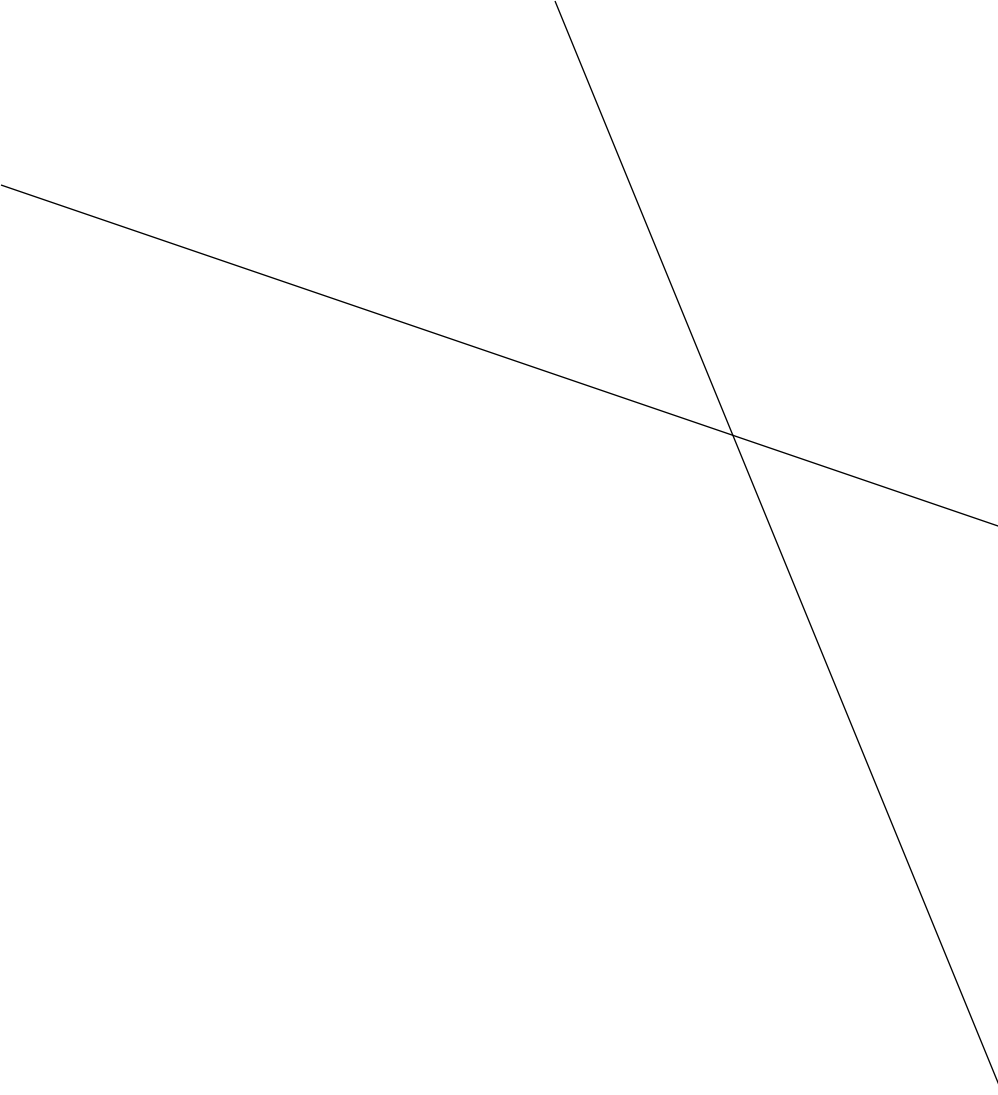
ARGUMENTY PRZECIWKO KOMPUTACJONIZMOWI





Obliczeniowy model mózgu przyciąga liczne zastrzeżenia. W wielu przypadkach zastrzeżenia te dotyczą tylko konkretnych wersji (takich jak klasyczny komputacjonizm czy komputacjonizm koneksjonistyczny). Poniżej przedstawiamy kilka prominentnych zastrzeżeń.

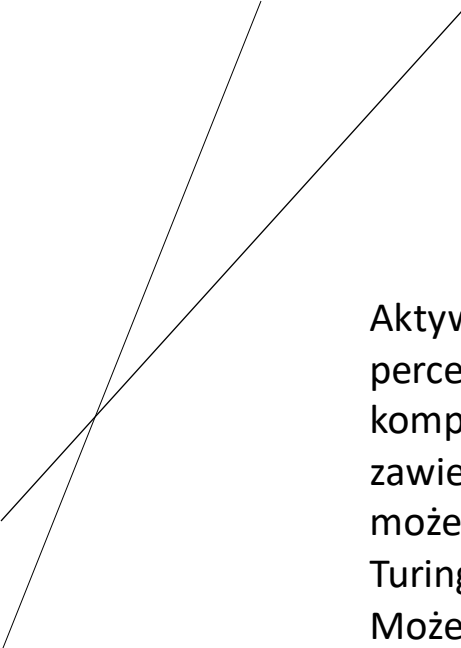
Powtarzającym się problemem jest to, że obliczeniowy model mózgu jest trywialny, ponieważ możemy opisać prawie każdy system fizyczny jako wykonujący obliczenia. Amerykański filozof John Searle twierdzi, że ściana implementuje dowolny program komputerowy, ponieważ możemy dostrzec pewien wzór ruchów molekularnych w ścianie, który jest izomorficzny z formalną strukturą programu.



Według niektórych autorów, twierdzenie niezupełności Gödla pokazuje, że ludzkie zdolności matematyczne przewyższają zdolności jakiegokolwiek maszyny Turinga. Różni filozofowie i logicy odpowiedzieli na tę krytykę, argumentując, że istniejące sformułowania są obarczone błędami, założeniami kwestionującymi, a nawet jawnymi błędami matematycznymi. Może się okazać, że pewne ludzkie zdolności umysłowe przewyższają możliwości obliczeniowe Turinga, ale twierdzenia o niezupełności Gödla nie dają powodu, by przewidywać taki rezultat.

Granice modelowania obliczeniowego

Czy komputer mógłby skomponować symfonię Eroica? Albo odkryć ogólną względność? Albo nawet odtworzyć bezwysiłkową zdolność dziecka do postrzegania otoczenia, wiązania sznurowadeł i rozpoznawania emocji innych? Intuicyjne, twórcze lub zręczne ludzkie działanie może wydawać się odporne na formalizację przez program komputerowy. Bardziej ogólnie, można się martwić, że kluczowe aspekty ludzkiego poznania wymykają się zwłaszcza klasycznemu modelowaniu obliczeniowemu.



Aktywność umysłowa rozwija się w czasie. Co więcej, umysł wykonuje skomplikowane zadania (np. estymację percepcyjną) bardzo szybko. Wielu krytyków martwi się, że komputacjonizm, a zwłaszcza klasyczny komputacjonizm, nie uwzględnia odpowiednio temporalnych aspektów poznania. Model w stylu Turinga nie zawiera wyraźnej wzmianki o skali czasowej, w której zachodzą obliczenia. Obliczeniowcy odpowiadają, że możemy uzupełnić abstrakcyjny model obliczeniowy o względy temporalne. Na przykład model maszyny Turinga zakłada dyskretne "etapy obliczeń", bez opisywania, jak te etapy odnoszą się do czasu fizycznego. Możemy jednak uzupełnić nasz model opisując, jak długo trwa każdy etap, przekształcając w ten sposób nasz nieczasowy model maszyny Turinga w teorię, która daje szczegółowe przewidywania czasowe.



BIBLIOGRAFIA

<https://plato.stanford.edu/entries/computational-mind/#ComFor>



DZIĘKUJEMY ZA UWAGĘ

Maciej Zawisz
Szymon Żylski
Hubert Wątorczyk