Mózg jako silnik predykcyjny w świetle koncepcji emergencji Boba Burgina

AUTOR

Fundacja Dobre Państwo

STRESZCZENIE / ABSTRACT

Artykuł analizuje teorię mózgu jako systemu generatywnego, który nie odbiera pasywnie danych sensorycznych, lecz aktywnie konstruuje rzeczywistość poprzez predykcje. Autor rozróżnia kluczowe pojęcia: Bayesian brain, predictive processing, predictive coding oraz active inference, wskazując na ich różnice w zakresie implementacji i zakresu wyjaśniającego. Percepcja jest przedstawiona jako proces konfrontowania wcześniejszych przekonań (priori) z napływającymi bodźcami, gdzie błąd predykcji służy do aktualizacji wewnętrznych modeli świata. Tekst podkreśla, że choć paradygmat ten jest wpływowy w neuronauce i filozofii umysłu, wciąż pozostaje przedmiotem empirycznych debat, a proces poznawczy jest nierozerwalnie związany z historią i architekturą organizmu.

The article analyzes the theory of the brain as a generative system that does not passively receive sensory data, but actively constructs reality through predictions. The author distinguishes between key concepts: the Bayesian brain, predictive processing, predictive coding, and active inference, pointing out their differences in terms of implementation and explanatory scope. Perception is presented as a process of confronting prior beliefs (priors) with incoming stimuli, where prediction error serves to update internal models of the world. The text emphasizes that although this paradigm is influential in neuroscience and the philosophy of mind, it remains a subject of empirical debate, and the cognitive process is inextricably linked to the history and architecture of the organism.

Artykuł rzuca wyzwanie potocznemu rozumieniu percepcji jako pasywnego rejestrowania bodźców, proponując w zamian wizję mózgu jako aktywnego silnika predykcyjnego. Autor analizuje rodzinę koncepcji od Bayesian brain po active inference, dowodząc, że nasze doświadczanie świata jest w istocie nieustannym procesem

generowania hipotez i ich korygowania w oparciu o błędy predykcji. Kluczowa teza tekstu zakłada, że prawdziwe poznanie nie wyłania się z samej mocy obliczeniowej czy statystycznej analizy danych, lecz z ucieleśnionej pętli sprzężeń zwrotnych między organizmem a środowiskiem. W tym ujęciu pamięć służy do kompresowania regularności w celu antycypacji przyszłości, a kreatywność jest dynamiczną równowagą między stabilnością modelu a otwartością na nowość. Ostatecznie tekst argumentuje, że różnica między biologicznym umysłem a sztuczną inteligencją tkwi w uwikłaniu predykcji w ryzyko i biologiczny system wartościowania – poznanie staje się podmiotowe dopiero wtedy, gdy system posiada ciało i metabolizm, a co za tym idzie: ma coś do stracenia.

SŁOWA KLUCZOWE

mózg jako silnik predykcyjny | predictive processing | active inference | błąd predykcji | Bayesian brain | predictive coding | emergencja poznawcza | uczenie statystyczne | zasada swobodnej energii | priors | affordance | poznanie enaktywne | interocepcja | ograniczona racjonalność

Percepcja jako aktywna interpretacja i predykcja

Mózg, który zgaduje świat: zmysły, predykcja, czas i kreatywność. Pamięć pozwala przeszłości kształtować przyszłe zachowanie organizmu; percepcja zaś jest punktem styku tej przeszłości z teraźniejszością. Zmysły nie funkcjonują bowiem jak zestaw neutralnych kamer, mikrofonów czy mierników, które przesyłałyby do mózgu gotowy obraz świata. Samo pojęcie „surowych danych sensorycznych” okazuje się jedynie wygodnym uproszczeniem. Receptory reagują selektywnie na określone zakresy energii, a kolejne etapy układu nerwowego przekształcają sygnał: wzmacniają jedne różnice, osłabiają inne i dokonują integracji przestrzennej oraz czasowej. Co więcej, uwaga i wcześniejsze doświadczenie determinują to, co w ogóle staje się dostępne do dalszego przetwarzania. Percepcja jest zatem osiągnięciem organizmu, nie mechanicznym odbiciem rzeczywistości. Świat fizyczny dostarcza przyczyn sygnałów, lecz doświadczany świat stanowi wynik ich interpretacji przez system o konkretnej historii, architekturze, potrzebach i repertuarze działania.

Burgin intuicyjnie ujmuje tę ideę poprzez obraz mózgu jako „silnika predykcyjnego”: system generuje oczekiwania, konfrontuje je z napływającym bodźcem, wykrywa rozbieżność i w odpowiedzi aktualizuje model. To trafny punkt wyjścia, jednak współczesny stan badań wymaga rozdzielenia kilku koncepcji, które w popularnych ujęciach często zlewają się w jedną teorię.

„Bayesian brain”, predictive processing, predictive coding oraz active inference tworzą rodzinę powiązanych idei, lecz nie są synonimami. Bayesian brain to szeroka hipoteza, według której percepcję i decyzje można modelować jako wnioskowanie w warunkach niepewności, gdzie wcześniejsze przekonania łączone są z informacją sensoryczną. Predictive processing stanowi szerszy nurt koncepcji traktujących mózg jako system generatywny, wykorzystujący oczekiwania do organizacji percepcji. Z kolei predictive coding to bardziej szczegółowa propozycja dotycząca implementacji tych obliczeń; zakłada ona wymianę przewidywań i sygnałów błędu pomiędzy poziomami hierarchicznej architektury.

Active inference idzie o krok dalej, obejmując nie tylko aktualizację modelu pod wpływem świata, ale również działanie zmieniające otoczenie – lub pozycję organizmu względem niego – tak, aby uzyskać przewidywane stany sensoryczne. Rozróżnienie to jest niezbędne, gdyż predictive processing, mimo ogromnego wpływu na neuronaukę i filozofię umysłu, nie powinien być przedstawiany jako bezspornie udowodniona, jednolita teoria mózgu. Najnowszy przegląd w Annual Review of Neuroscience podkreśla, że choć liczne wyniki neurofizjologiczne są zgodne z przewidywaniami predictive coding, to heterogeniczność definicji oraz niejednoznaczność tego, co faktycznie reprezentują sygnały „błędu predykcji”, pozostawiają istotne problemy empiryczne i teoretyczne. Podobną ostrożność sugeruje analiza statusu predictive coding i active inference: istnieją dane wspierające oba podejścia, lecz ich zakres wyjaśniający jest wciąż testowany, a ogólna formuła teorii łatwiejsza do obrony niż każda z jej konkretnych implementacji neuronalnych. W tekście naukowym należy zatem mówić o wpływowym paradygmacie i rodzinie modeli, a nie o ostatecznie ustalonej „zasadzie działania mózgu”. Podstawowa intuicja pozostaje jednak silna: świat sensoryczny jest niejednoznaczny. Ten sam wzór pobudzenia receptorów może mieć wiele przyczyn, a organizm musi działać szybciej, niż pozwalałaby na to pełna analiza wszystkich dostępnych hipotez.

Percepcja jako proces predykcyjny korygowany przez błąd

Widząc częściowo zasłonięty przedmiot, człowiek zazwyczaj nie doświadcza zbioru kolorowych fragmentów, lecz rozpoznaje krzesło, twarz albo samochód. Słysząc początek zdania, generuje

oczekiwania co do jego dalszego ciągu. Idąc po schodach, układ ruchowy przygotowuje stopę na określoną wysokość kolejnego stopnia, zanim pełna informacja sensoryczna zostanie świadomie przeanalizowana. System nerwowy wykorzystuje regularności z przeszłości, ponieważ rzeczywistość nie jest dla organizmu losową sekwencją niezależnych zdarzeń. Świat posiada strukturę, a uczenie polega między innymi na przyswojeniu tej struktury w formie oczekiwań. W języku bayesowskim można powiedzieć, że organizm dysponuje pewnymi priors – wcześniejszymi rozkładami prawdopodobieństwa możliwych przyczyn bodźców – które konfrontuje z nowymi danymi. Nie należy jednak wyobrażać sobie mózgu jako uczonego zapisującego na tablicy wzór Bayesa i świadomie obliczającego prawdopodobieństwa. Model bayesowski jest przede wszystkim formalnym językiem opisującym racjonalną aktualizację przekonań w warunkach niepewności. To, czy i w jaki sposób konkretne obwody neuronalne implementują obliczenia odpowiadające formalnemu wnioskowaniu bayesowskiemu, stanowi odrębne pytanie neurobiologiczne. Rozróżnienie poziomu obliczeniowego od algorytmicznego i implementacyjnego, zaproponowane przez Davida Marra, okazuje się tutaj szczególnie użyteczne: ten sam problem obliczeniowy może zostać rozwiązany za pomocą różnych mechanizmów.

Predykcja nie oznacza przy tym fantazjowania zamiast widzenia. Dobre modele predykcyjne są podporządkowane korekcie przez rzeczywistość. Jeśli człowiek oczekuje na biurku kubka, może szybciej dostrzec częściowo zasłonięty kształt zgodny z tym założeniem, jednak wystarczająco silny sygnał niezgodny z hipotezą powinien doprowadzić do jej rewizji. To właśnie błąd predykcji pełni w modelach predictive coding szczególną funkcję. Rozbieżność między oczekiwanym a obserwowanym sygnałem niesie informację o tym, czego model świata nie przewidział prawidłowo. Nie oznacza to, że mózg „interesuje się błędami” w sensie psychologicznym; chodzi raczej o to, że rozbieżności mają potencjalnie większą wartość informacyjną niż zdarzenia całkowicie zgodne z oczekiwaniem. Burgin wiąże ten mechanizm z przepływami top-down i bottom-up, wskazując na odgórne oczekiwania oraz oddolne sygnały korygujące.

Jest to heurystyka użyteczna, lecz nie należy utożsamiać jej mechanicznie z Kahnemanowskimi Systemem 1 i Systemem 2, jak czyni to jeden z fragmentów notatki. System 1 i System 2 stanowią funkcjonalne rozróżnienie pomiędzy szybkimi, automatycznymi procesami a bardziej wymagającym przetwarzaniem kontrolowanym; przepływy top-down i bottom-up opisują natomiast kierunki oddziaływań w hierarchiach sensorycznych. Automatyczne procesy mogą być silnie kształtowane odgórnymi oczekiwaniami, zaś informacja oddolna może uczestniczyć w świadomym, deliberatywnym rozumowaniu. Te dwa podziały przecinają się, ale nie mapują na siebie jeden do jednego.

Biologiczne mechanizmy predykcji i uczenia statystycznego

Predictive processing przestaje być wyłącznie metaforą w momencie, gdy można wskazać konkretne mechanizmy neuronalne sygnałów niezgodności. Badania opublikowane w Nature w 2024 roku wykazały u myszy udział współpracujących obwodów korowo-wzgórzowych w generowaniu sensorycznych błędów predykcji podczas przetwarzania wzrokowego. Autorzy opisali odrębne źródła sygnałów związanych z oczekiwanym i rzeczywistym ruchem oraz ich integrację w obwodach pozwalających reagować na rozbieżność. Nie dowodzi to oczywiście, że cały mózg można wyjaśnić jednym schematem predictive coding, ale pokazuje, że przewidywanie konsekwencji własnego działania i wykrywanie sensorycznej niespodzianki mogą posiadać identyfikowalne realizacje obwodowe.

Z punktu widzenia emergencji najciekawsze jest to, że predykcja przekształca samo znaczenie pamięci. Pamięć nie musi służyć głównie rekonstrukcji minionego świata; może być mechanizmem kompresowania regularności przeszłości po to, aby lepiej przewidywać przyszłość. Organizm nie musi pamiętać każdej gałęzi, pod którą przechodził, jeśli nauczył się abstrakcyjnej reguły, że nisko wiszące przeszkody wymagają pochylenia głowy. Nie musi przechowywać dokładnego obrazu każdej twarzy oglądanej w życiu, jeżeli potrafi wytworzyć kategorie i mechanizmy rozpoznawania umożliwiające generalizację. Z perspektywy adaptacyjnej zapominanie szczegółu i abstrahowanie regularności nie są przeciwieństwem uczenia – mogą być właśnie jego wyższym stadium.

Dlatego rozwój poznawczy niemowlęcia jest dla naszego wywodu szczególnie istotny. Burgin, odwołując się do Stanislasa Dehaene'a i badań nad rozwojem, przeciwstawia mózg noworodka Locke'owskiej tabula rasa, przedstawiając go jako system wyposażony w architekturę zdolną do wczesnego wychwytywania regularności statystycznych. Należy jednak zaznaczyć: z faktu, że bardzo małe dzieci wykazują wrażliwość na regularności, nie wynika, iż rodzą się z gotową wiedzą probabilistyczną w sensie, w jakim rozumie ją dorosły matematyk. Znacznie bezpieczniejsza teza brzmi, że mózg rozwijającego się dziecka posiada biologicznie przygotowane mechanizmy uczenia, które bardzo wcześnie wydobywają regularności z napływających danych. Klasyczne eksperymenty nad uczeniem statystycznym pokazały, że już niemowlęta potrafią wykorzystywać prawdopodobieństwa przejść pomiędzy sylabami do segmentacji ciągłego strumienia mowy. Nowsze badania wykazują podobne zdolności nawet u śpiących noworodków, choć zakres i trwałość powstających reprezentacji pozostają ograniczone.

Szerszy przegląd badań rozwojowych wskazuje, że uczenie statystyczne zmienia się wraz z wiekiem w sposób zależny od rodzaju informacji. Młodsze dzieci mogą w pewnych sytuacjach posiadać

szersze i mniej ukierunkowane „curriculum” uczenia niż osoby starsze. To interesująca komplikacja tezy o coraz doskonalszym mózgu: dojrzałość daje większą wiedzę i bardziej precyzyjne priory, ale właśnie dlatego może również zawężać przestrzeń tego, czego system jest gotów się nauczyć.

Kreatywność jako dynamiczna równowaga między stabilnością modelu a eksploracją

W tym miejscu zarysowuje się głęboka zależność pomiędzy predykcją a kreatywnością. Organizm całkowicie pozbawiony wcześniejszych modeli byłby zalewany nieuporządkowanym strumieniem bodźców. Z kolei ten, którego modele byłyby absolutnie sztywne, widziałby jedynie to, co już potrafi przewidzieć. Adaptacyjne poznanie wymaga zatem ruchu pomiędzy tymi skrajnościami: stabilnych oczekiwań, dzięki którym świat pozostaje rozpoznawalny, oraz plastyczności pozwalającej niespodziance zmienić model. Kreatywność można rozpatrywać właśnie w tym obszarze – jako zdolność do generowania i utrzymywania alternatywnych hipotez, łączenia odległych reprezentacji oraz chwilowego osłabiania dominacji najbardziej prawdopodobnej interpretacji, bez całkowitej utraty kontroli nad kryterium sensowności. Nie należy jednak sprowadzać tego do jednej "teorii kreatywności jako błędu predykcji".

Twórczość jest zjawiskiem wieloczynnikowym; zależy m.in. od pamięci, uwagi, kontroli wykonawczej, wiedzy dziedzinowej, motywacji i zdolności do tworzenia odległych skojarzeń. Współczesne badania neurokognitywne wiążą myślenie twórcze raczej z interakcjami pomiędzy systemami generowania spontanicznych reprezentacji a sieciami kontroli i oceny, niż lokalizują je w jednym konkretnym "ośrodku". Badania opublikowane w Communications Biology w 2024 roku ponownie wskazały na związek zdolności twórczych z szeroką architekturą obejmującą procesy pamięciowe, uwagowe i wykonawcze. Kreatywność nie musi zatem oznaczać zawieszenia przewidywania; może być szczególnym sposobem zarządzania przestrzenią przewidywań.

Można to ująć jako problem równowagi pomiędzy eksploatacją a eksploracją. Eksploatacja to korzystanie z modelu, który już działa: rozpoznanie sytuacji i wybór sprawdzonej reakcji w celu oszczędności zasobów. Eksploracja natomiast oznacza dopuszczenie alternatywy: uznanie, że dotychczasowa interpretacja jest niewystarczająca, istnieje nowe połączenie lub problem da się przeformułować. System oparty wyłącznie na eksploatacji jest wydajny w znanym środowisku, lecz kruchy wobec zmiany. System czysto eksploracyjny generuje wiele możliwości, ale może nie potrafić żadnej z nich ustabilizować. Kreatywność wymaga obu tych procesów – chwilowego rozszerzenia przestrzeni hipotez i późniejszej selekcji wytworzonych wariantów. Z tej perspektywy błędem byłoby utożsamienie kreatywności z chaosem. Najbardziej niezwykła odpowiedź nie jest

automatycznie najbardziej twórcza; klasyczne definicje psychologiczne łączą nowość z trafnością, użytecznością lub adekwatnością. System generujący przypadkowe kombinacje byłby wprawdzie nieprzewidywalny, lecz niekoniecznie kreatywny.

Twórcza reorganizacja wymaga zachowania wystarczającej struktury, aby odbiorca – lub sam twórca – mógł rozpoznać relację pomiędzy nowym rozwiązaniem a problemem. Emergencja nowej idei zachodzi więc na granicy porządku i odchylenia: model musi zostać naruszony, ale nie unicestwiony. Niektóre wyniki dotyczące przetwarzania istotności sensorycznej dają interesujący, choć wciąż ostrożny punkt zaczepienia dla takiej interpretacji. Badanie z 2024 roku wykazało u osób bardziej twórczych odmienne reakcje psychologiczne i elektrofizjologiczne na bodźce typu oddball, co sugeruje nietypowe przetwarzanie istotności i większą otwartość na informacje konkurujące o uwagę. Nie wynika z tego uniwersalna reguła, że "kreatywni ignorują predykcje", lecz wynik ten współgra z hipotezą, iż twórczość wymaga innego kompromisu pomiędzy filtrowaniem a otwartością na niespodziewaną informację.

Tu dochodzimy do pojęcia precyzji (precision), jednego z kluczowych, choć często pomijanych elementów modeli predykcyjnych. Nie każdy błąd predykcji powinien zmieniać przekonanie z jednakową siłą. Jeśli w gęstej mgle widzę zarys przypominający człowieka, informacja sensoryczna jest niepewna i to silny model wcześniejszy dominuje interpretację. W jasnym świetle precyzyjny bodziec łatwiej skoryguje błędne oczekiwanie.

W formalnych modelach bayesowskich znaczenie informacji zależy nie tylko od samego błędu, lecz również od wiarygodności źródła. Organizm musi zatem szacować nie tylko to, co widzi, ale i to, jak bardzo może ufać danemu sygnałowi. W szerszym sensie jest to problem epistemologii ucieleśnionego organizmu: nie wszystkie źródła wiedzy mają jednakową wagę. Uwaga może być rozumiana jako mechanizm zmieniający wagę wybranych sygnałów, choć i tutaj predictive processing nie powinien być traktowany jako jedyna teoria uwagi. Oczekiwanie i uwaga pozostają pojęciami odrębnymi.

Zdarzenie może być bardzo prawdopodobne, a jednocześnie mało istotne; może być też skrajnie nieprawdopodobne, lecz kluczowe dla aktualnego celu. Badania nad percepcyjnym podejmowaniem decyzji podkreślają, że oczekiwanie dotyczy prawdopodobieństwa wystąpienia bodźca, podczas gdy selektywna uwaga odnosi się do jego behawioralnej relewancji. Rozróżnienie to okaże się istotne przy analizie proto-emocji. System musi nie tylko przewidywać świat, ale również rozstrzygać, które z tych zdarzeń – przewidziane czy nie – mają dla niego rzeczywiste znaczenie.

Percepcja jako sprzężenie działania z otoczeniem

Zmysły nie tworzą osobnych kanałów, które spotkałyby się dopiero w jakimś centralnym "kinie świadomości". Organizm funkcjonuje w środowisku multisensorycznym, gdzie informacje wzrokowe, słuchowe, dotykowe, proprioceptywne, przedsionkowe, węchowe, smakowe i interoceptywne są integrowane w stopniu zależnym od kontekstu i konkretnego zadania. To dlatego ruch ust może zmieniać sposób słyszenia fonemu, konflikt między wzrokiem a układem przedsionkowym prowadzi do choroby lokomocyjnej, a widok narzędzia w zasięgu ręki jest kodowany inaczej niż abstrakcyjny obiekt oddalony od nas. Świat percepcyjny nie jest zatem mapą sporządzoną przez neutralnego kartografa, lecz zapisem tworzonym dla organizmu zdolnego do działania.

Pojęcie affordance, wprowadzone przez Jamesa J. Gibsona, pozwala spojrzeć na ten problem z innej perspektywy niż klasyczne predictive processing. Gibsonowska psychologia ekologiczna akcentuje, że środowisko udostępnia organizmowi konkretne możliwości działania: powierzchnia może być "do chodzenia", uchwyt "do chwycenia", a otwór "do przejścia" — wszystko to zależnie od relacji między właściwościami świata a ciałem. Organizm nie musi najpierw tworzyć pełnego, neutralnego opisu przedmiotu, by następnie decydować o interakcji z nim; percepcja jest sprzężona z działaniem od samego początku. Stanowi to istotny kontrapunkt dla nadmiernie internalistycznych wersji mózgu predykcyjnego. Poznanie nie polega bowiem wyłącznie na udoskonalaniu obrazu rzeczywistości wewnątrz czaszki, lecz na regulowaniu relacji organizm-środowisko.

Jeszcze dalej idą podejścia enaktywne, kojarzone z Francisco Varelą, Evanem Thompsonem i Eleanor Rosch. W tej tradycji poznanie nie jest przede wszystkim odtwarzaniem istniejącego świata wewnątrz systemu reprezentacyjnego. Organizm aktywnie ustanawia świat poznawczo relewantny poprzez własną strukturę, potrzeby i działanie. Nietoperz, człowiek, bakteria czy robot mogą przebywać fizycznie w tym samym otoczeniu, lecz nie oznacza to, że istnieją w tym samym środowisku poznawczym. Znaczenie wyłania się z relacji między organizacją systemu a strukturą możliwości działania. To właśnie tutaj "ciało" zaczyna oznaczać coś więcej niż platformę transportującą mózg.

Active inference próbuje połączyć te intuicje z formalizmem modeli generatywnych. W najprostszym schemacie system może redukować rozbieżność między przewidywanym a otrzymywanym sygnałem na dwa sposoby: zmieniając model albo działając tak, by zmienić napływające dane. Jeśli przewiduję obecność kubka za książką, mogę zaktualizować przekonanie i

uznać, że go tam nie ma, lub przesunąć głowę, by sprawdzić przestrzeń za przeszkodą. Percepcja i działanie stają się tym samym częściami jednej pętli.

Percepcja jako działanie i symulacja przyszłości

W literaturze active inference rozbudowano ten pomysł w formalny aparat związany z zasadą swobodnej energii, choć jego empiryczny zakres pozostaje przedmiotem dyskusji. Dla naszego problemu wystarczy jednak istotniejsza intuicja: poznający organizm nie jest obserwatorem nieruchomo siedzącym przed ekranem świata. Uzyskuje on informację poprzez działanie. Ta zasada radykalnie zmienia rolę zmysłów. Oko się porusza, głowa obraca, dłoń bada powierzchnię, a ciało zmienia pozycję – system aktywnie kreuje warunki, w których niejednoznaczna wcześniej informacja staje się klarowna. Percepcja jest eksperymentem. Organizm nie tylko odbiera dowody, ale sam decyduje, gdzie szukać kolejnych. Ruch sakkadowy oka można potraktować jako mikroeksperyment: hipoteza dotycząca obiektu kieruje pobieraniem następnej próbki danych. Dziecko podnosi zabawkę, obraca ją, wkłada do ust lub uderza nią o stół; poznanie materiału, masy i kształtu przedmiotu wyłania się z zorganizowanej sekwencji percepcji i działania. Właśnie dlatego ucieleśnienie stało się tak kluczowym problemem w badaniach nad AI.

Model językowy może dysponować ogromną statystyczną wiedzą o spadaniu przedmiotów, bólu czy oporze powierzchni, gdyż nauczył się regularności obecnych w tekstach. Nie oznacza to jednak, że jego poznanie posiada tę samą strukturę co poznanie organizmu, który odkrył grawitację poprzez tysiące nieudanych prób wstawania, chwytania i utrzymywania równowagi. Różnica nie sprowadza się do ilości informacji, lecz dotyczy historii przyczynowej, dzięki której informacja stała się częścią systemu. Zmysły są nierozerwalnie związane z czasem. Sygnały docierają z opóźnieniem, ich przetwarzanie trwa, a działanie wymaga przygotowania – świat może zmienić się, zanim odpowiedź zostanie wykonana. Organizm reagujący dopiero na przetworzony stan świata pozostawałby systematycznie spóźniony. Predykcja jest więc częściowo odpowiedzią na fizyczną strzałkę czasu. Mózg musi antycypować położenie poruszającego się obiektu, skutek ruchu ręki czy przebieg wypowiadanego zdania, oceniając, czy bieżąca trajektoria prowadzi do zagrożenia.

Świadomość teraźniejszości nie jest zatem matematycznym punktem między nieistniejącą już przeszłością a nieistniejącą jeszcze przyszłością. To czasowe okno integracji, w którym ślady tego, co przed chwilą zaszło, łączą się z przewidywaniem tego, co prawdopodobnie nastąpi. Zjawisko to pozwala głębiej zrozumieć związek czasu z kreatywnością. Twórcza idea często wymaga utrzymania elementów nieobecnych jednocześnie w bieżącej percepcji: wspomnienia, celu czy alternatywnego stanu świata. System zdolny wyłącznie do reakcji na aktualny bodziec może być sprawny, lecz

trudno przypisać mu rozwiniętą wyobraźnię. Wyobraźnia wykorzystuje architekturę modelowania poza aktualnym sygnałem. Człowiek potrafi przeprowadzić w umyśle scenariusz, którego nigdy nie widział, przewidzieć skutki decyzji, odtworzyć nieistniejący krajobraz czy zaprojektować instytucję, której jeszcze nie ma. Pamięć i predykcja spotykają się tutaj w procesie generatywnym: elementy przeszłości są rekombinowane, by wytwarzać możliwe przyszłości. Nie jest przypadkiem, że uszkodzenia struktur pamięci epizodycznej upośledzają zdolność szczegółowego wyobrażania sobie nowych scenariuszy. Związek pamięci i prospekcji sugeruje, że system autobiograficzny nie jest archiwum zorientowanym wyłącznie ku przeszłości.

Jego adaptacyjna funkcja polega właśnie na umożliwieniu mentalnej symulacji. Przeszłe epizody dostarczają materiału do budowania modeli tego, czego jeszcze nie było. W tej perspektywie człowiek pamięta częściowo po to, aby móc tworzyć przyszłość. Kreatywność otrzymuje tym samym wyraźnie temporalny charakter. Nowość nie powstaje w próżni. Mozart bez kontaktu z systemem tonalnym nie byłby bardziej kreatywny, lecz pozbawiony języka muzycznego, w którym możliwa była innowacja. Einstein potrzebował wcześniejszej fizyki, aby móc zakwestionować jej założenia. Dziecko uczące się języka najpierw asymiluje regularności, zanim zaczyna konstruować zdania, których nigdy wcześniej nie słyszało. Twórczość jest więc paradoksalnym owocem ograniczenia przez wcześniejszą strukturę: model świata musi istnieć, zanim można go przekroczyć.

Strukturalna analogia błędu predykcji i granice modelu

W tym ujęciu „błąd predykcji” zyskuje znaczenie wykraczające poza ramy neuronalnego sygnału. Każde odkrycie naukowe, innowacja artystyczna czy radykalna zmiana kategorii poznawczej bierze swój początek w momencie, gdy dotychczasowy model przestaje być wystarczający. Nie oznacza to, że fizjologiczny prediction error i rewolucja naukowa są tym samym zjawiskiem; można jednak dostrzec między nimi strukturalną analogię. Model generuje oczekiwanie, rzeczywistość lub eksperyment ujawnia niezgodność, a system może wówczas albo bronić dotychczasowych kategorii, albo przebudować przestrzeń reprezentacji. Na poziomie jednostki przejawia się to jako uczenie percepcyjne, w nauce – jako zmiana teorii, a w kulturze – jako reorganizacja systemów znaczeń.

Należy jednak wystrzegać się pokusy uczynienia z predictive processing „teorii wszystkiego”. Percepcja podlega przecież adaptacji sensorycznej, mechanizmom uwagi, lokalnej dynamice sieci czy ograniczeniom receptorowym – procesom, których nie trzeba rekonstruować wyłącznie jako minimalizowania błędu predykcyjnego. Przegląd badań nad adaptacją wzrokową z 2024 roku wskazywał wręcz na przypadki, w których modele lokalnego zmęczenia sieciowego lepiej wyjaśniają obserwowane efekty niż proste wersje predictive coding.

Jest to cenna metodologicznie przestroga. Wielka teoria traci na użyteczności, jeśli każde możliwe zachowanie można post factum opisać jako „minimalizowanie niespodzianki". Model naukowy musi nie tylko obejmować szeroki wachlarz zjawisk, ale przede wszystkim generować rozróżniające przewidywania. Podobnie nie rozstrzygnięto sporu o stopień, w jakim poznanie wyższego rzędu bezpośrednio przenika wczesne etapy percepcji. Choć wpływ kontekstu i oczekiwań na doświadczenie percepcyjne jest dobrze udokumentowany, pytanie o cognitive penetrability – czy przekonania, pragnienia i wiedza konceptualna zmieniają samą treść percepcji, czy jedynie wpływają na późniejszą interpretację i decyzję – pozostaje otwarte.

Przegląd różnych modalności sensorycznych z 2024 roku pokazuje, że problem jest znacznie bardziej złożony niż sugeruje hasło „widzimy to, co chcemy zobaczyć". To kolejny powód, dla którego w dalszej analizie należy odróżniać metafory poznawcze od ustalonych mechanizmów. Mimo tych zastrzeżeń współczesne badania pozwalają utrzymać jedną mocną tezę: percepcja nie jest izolowana od historii systemu. Badanie opublikowane w Nature Communications w 2024 roku wykazało, że uczenie relacji predykcyjnych może reorganizować geometrię neuronalnych reprezentacji obiektów, zwiększając podobieństwo reprezentacji elementów powiązanych przewidywaniami. Innymi słowy, przyswojenie struktury świata zmienia nie tylko deklaratywną wiedzę o relacjach między obiektami, ale i sam sposób organizacji ich reprezentacji w mózgu.

Powracamy zatem do wniosku z poprzedniej części: uczenie nie polega jedynie na dodawaniu informacji. Może ono przekształcać przestrzeń, w której kolejne dane będą interpretowane. To właśnie tutaj emergencja, czas i zmysły zaczynają tworzyć spójną konstrukcję. Z pojedynczych zdarzeń sensorycznych wyłaniają się regularności; z regularności – oczekiwania; z oczekiwań – model świata; z rozbieżności między modelem a doświadczeniem – uczenie; z pamięci i modelowania – zdolność symulowania przyszłości, a z wielości tych symulacji – przestrzeń kreatywnego działania.

Mózg jako regulator relacji i system ograniczonej racjonalności

Nie jest to liniowa drabina, na której kolejne stadium po prostu zastępuje poprzednie; każdy z tych procesów sprzęga się z pozostałymi. To, czego oczekujemy, wpływa na to, co zauważamy; to, co dostrzegamy, kształtuje naszą pamięć, a ta z kolei definiuje zakres przyszłych oczekiwań. Nowa konfiguracja doświadczenia może z kolei przebudować cały ten obieg. W takim ujęciu mózg nie jest ani biernym odbiornikiem świata, ani samotnym generatorem halucynacji kontrolowanych przez dane sensoryczne. Jest raczej regulatorem relacji pomiędzy organizmem a środowiskiem.

Wytwarza modele, które są nieustannie testowane przez konsekwencje działania. Przewiduje, lecz pozostaje podatny na korektę. Selekcjonuje, choć selekcja ta może zostać zakwestionowana przez silną niespodziankę. Stabilizuje wiedzę, zachowując jednocześnie zdolność do reorganizacji. Jego poznawcza skuteczność nie wynika z perfekcyjnego odwzorowania rzeczywistości, lecz z umiejętności tworzenia wystarczająco trafnych modeli przy niepełnej informacji i ograniczonych zasobach. To rzuca nowe światło na zawartą w książce Burgina intuicję "dwudziestu watów".

Samo zestawienie poboru mocy ludzkiego mózgu z energochłonnością współczesnych systemów obliczeniowych nie jest tu kluczowe. Ważniejsza okazuje się hipoteza organizacyjna: biologiczny system poznawczy ukształtował się w warunkach, które uniemożliwiały analizowanie, zapamiętywanie i sprawdzanie wszystkiego. Musiał on rozwiązać problem ograniczonej racjonalności na długo przed tym, zanim Herbert Simon nadał temu pojęciu nazwę w teorii decyzji. Heurystyka, selekcja uwagi, predykcja, automatyzacja oraz emocjonalne znaczniki istotności nie są z tej perspektywy niedoskonałościami idealnego rozumu, lecz rozwiązaniami architektonicznymi umożliwiającymi działanie skończonemu organizmowi. Z tego powodu różnica między biologicznym umysłem a AI nie sprowadza się jedynie do mocy obliczeniowej. System dysponujący milion razy większymi zasobami nie staje się automatycznie milion razy inteligentniejszy, podobnie jak biblioteka zawierająca milion razy więcej książek nie staje się przez to milion razy mądrzejsza.

Kluczem pozostaje organizacja: to, co zostaje wybrane, jakie reprezentacje powstają i które sygnały modyfikują model. Istotne jest, jakie cele ustanawiają kryterium wartości, w jaki sposób system odróżnia informację istotną od szumu oraz jak działanie weryfikuje posiadane reprezentacje. Zwiększanie skali może generować nowe kompetencje, lecz sama skala nie odpowiada na pytanie o podmiotowość.

Poznanie jako ucieleśniona pętla wartościowania

Należy jednak starannie ograniczyć podobieństwo między modelem generatywnym AI a predykcyjnym mózgiem. Model językowy, generując tekst, prognozuje rozkład prawdopodobieństwa kolejnych elementów reprezentacji językowej; człowiek natomiast przewiduje na wielu sprzężonych poziomach konsekwencje sensoryczne, motoryczne, interoceptywne i społeczne własnego działania. W obu przypadkach wykorzystuje się nabyte regularności do przewidywania, lecz funkcjonalna analogia nie oznacza identyczności całego systemu. Kluczowe jest to, że ludzka predykcja jest stale konfrontowana z rzeczywistością przez ciało. Błędne przewidzenie położenia stopnia może skończyć się upadkiem. Zła ocena temperatury doprowadzi do oparzenia. Pomylenie intencji drugiego człowieka może trwale przekształcić moje przyszłe zachowanie.

Predykcja w biologii jest uwikłana w ryzyko. I właśnie ryzyko zapowiada kolejny etap problemu: system biologiczny nie tylko przewiduje, co się stanie, ale musi rozstrzygnąć, czy to zdarzenie jest dla niego dobre, złe, pilne, obojętne, atrakcyjne albo niebezpieczne.

Sam prediction error nie mówi jeszcze, czy niespodzianka jest ważna. Nieoczekiwany kolor przypadkowego samochodu i nieoczekiwany ruch drapieżnika mogą być równie niezgodne z wcześniejszym przewidywaniem, a jednak ich znaczenie dla organizmu jest radykalnie różne. Poznanie potrzebuje zatem kolejnej warstwy organizacji – systemu wartościowania.

W tym miejscu spotykają się ciało, homeostaza i emocja. Organizm funkcjonuje w ściśle określonym przedziale temperatury, nawodnienia, natlenienia, stężenia glukozy i tysięcy innych parametrów; nie każdy stan świata jest dla niego równoważny. Życie wprowadza asymetrię wartości jeszcze przed pojawieniem się świadomej refleksji: pewne trajektorie podtrzymują organizację, inne jej zagrażają. Z tego poziomu można przejść do koncepcji Damasia i LeDouxa, interocepcji oraz teorii emocji, by następnie zapytać, czy system cyfrowy pozbawiony metabolizmu może posiadać funkcjonalny odpowiednik takiej asymetrii.

Na tym etapie można sformułować kluczowy wniosek. Świadomości nie powinniśmy szukać poprzez dodawanie pojedynczych „modułów": pamięci, sensoryki, predykcji czy dużej mocy obliczeniowej. Biologia sugeruje raczej problem zamkniętej pętli organizacyjnej, w której historia zmienia model, model kształtuje oczekiwanie, oczekiwanie kieruje pobieraniem informacji, informacja koryguje model, działanie zmienia świat, a skutki działania wracają do systemu poprzez ciało. To nie pojedyncza zdolność, lecz sposób powiązania tych funkcji zaczyna przypominać architekturę podmiotu.

Zasadnicze pytanie przesuwa się zatem ponownie. Nie wystarczy zapytać, czy AI potrafi przewidywać – współczesne modele niewątpliwie wykorzystują regularności danych do imponujących predykcji i generowania. Nie wystarczy też pytać o zdolność tworzenia rzeczy nowych: systemy te mogą generować kombinacje niewystępujące w danych wejściowych w dokładnie tej postaci. Trzeba zapytać, dla kogo wynik predykcji jest błędem, dlaczego jakaś rozbieżność miałaby być dla systemu ważniejsza od innej oraz skąd pochodzi kryterium, według którego jedne stany powinny zostać utrzymane, a innych należy unikać. W biologii odpowiedź prowadzi ku żywemu ciału.

Pamięć wprowadziła przeszłość, predykcja – przyszłość, a zmysły osadziły oba kierunki czasu w bieżącej relacji ze światem. Kreatywność ujawniła możliwość generowania przyszłości, której dotąd nie było. Teraz trzeba wprowadzić element, który czyni te przyszłości nierównoważnymi: wartość. Dopiero gdy niektóre możliwe stany stają się dla systemu pożądane, niebezpieczne, pilne albo

warte kosztu, informacja jest selekcjonowana nie tylko według prawdopodobieństwa, lecz również według znaczenia. To właśnie na tej granicy poznanie spotyka emocję, a mózg przestaje być jedynie maszyną przewidywania i okazuje się częścią organizmu, który ma coś do stracenia.

> Czy zatem inteligencja bez ciała jest jedynie wyrafinowaną statystyką, pozbawioną możliwości prawdziwego rozumienia? Dopóki algorytm nie poczuje lęku przed błędem ani pragnienia przetrwania, pozostanie jedynie lustrem naszych danych, a nie podmiotem. Prawdziwe poznanie zaczyna się tam, gdzie informacja przestaje być tylko prawdopodobieństwem, a staje się kwestią przeżycia.

Inspiracje

[1] **"Emergence"** Bob Burgin

Bob Burgin to amerykański przedsiębiorca, naukowiec zajmujący się nauką o uczeniu się i autor, mieszkający w Boulder w stanie Kolorado. Jest założycielem, prezesem zarządu i dyrektorem generalnym Amplifire, firmy zajmującej się oprogramowaniem do adaptacyjnego uczenia się, której celem jest zastosowanie psychologii poznawczej i neuronauki w edukacji i szkoleniach zawodowych. W ciągu ponad trzydziestoletniej kariery w dziedzinie technologii i kognitywistyki Burgin współpracował z wybitnymi neurobiologami i badaczami pamięci, takimi jak Robert Björk, Daniel Schacter i Henry Roediger III, badając, w jaki sposób mózgi biologiczne kodują, przechowują i odzyskują informacje. Jego prace analizują, jak ewolucyjne ograniczenia poznawcze kształtują architekturę ludzkiego uczenia się oraz jak te zasady wpływają na współczesną sztuczną inteligencję i poznanie maszynowe. Burgin jest autorem książki „Emergence: How Our Twenty-Watt Brains Unlocked Intelligence Beyond Biology" (2026), która bada konwergencję neuronauki i systemów sztucznej inteligencji.

Bob Burgin *is an American entrepreneur, learning scientist, and author based in Boulder, Colorado. He is the founder, chairman, and chief executive officer of Amplifire, an adaptive learning software company founded to apply cognitive psychology and neuroscience to education and workforce training. Over a career spanning more than three decades in technology and cognitive science, Burgin has collaborated with prominent neuroscientists and memory researchers, including Robert Bjork, Daniel Schacter, and Henry Roediger III, to investigate how biological brains encode, retain, and retrieve information. His work examines how evolutionary cognitive constraints shape human learning architectures and how those principles inform modern artificial intelligence and machine cognition. Burgin is the author of Emergence: How Our Twenty-Watt Brains Unlocked Intelligence Beyond Biology (2026), which explores the convergence of human neuroscience and artificial learning systems.*

Amplifire

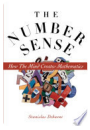
Literatura uzupełniająca

Książki

Książki autorów przywoływanych i literatura pokrewna:

[1] "What Is Consciousness, and Could Machines Have It?"

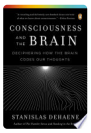
Stanislas Dehaene



[2] "The Number Sense: How the Mind Creates Mathematics, Revised and Updated Edition"

Stanislas Dehaene

Oxford University Press, USA | ISBN: 9780199723096



[3] "Consciousness and the Brain"

Stanislas Dehaene

Penguin | ISBN: 9780698151406

[4] "The Perception of the Visual World"

James J. Gibson



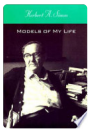
[5] "The ecological approach to visual perception"

James J. Gibson

Psychology Press | ISBN: 9780898599596

[6] "Conclusions from a Century of Research on Sense Perception"

James J. Gibson



[7] "Models of my life"

Herbert Simon

MIT Press | ISBN: 9780262326582



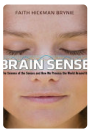
[8] "Looking for Spinoza: Joy, Sorrow, and the Feeling Brain"

Antonio Damasio

HMH | ISBN: 9780547541716

[9] "Descartes' error : emotion, reason, and the human brain"

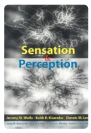
Antonio Damasio



[10] "Brain Sense"

Faith Hickman Brynie

2009 | AMACOM/American Management Association | ISBN: 9780814413241



[11] **"Sensation & Perception"**

Jeremy M. Wolfe

2006 | Sinauer Associates, Incorporated

Artykuły naukowe

Prace autorów przywoływanych w tekście:

[1] **"Spontaneous supra-modal encoding of number in the infant brain."**

Giulia Gennari, S. Dehaene, Chanel Valera, G. Dehaene-Lambertz

2023 | Current Biology | DOI: 10.1016/j.cub.2023.03.062

Źródło

[2] **"The Ecological Approach to Visual Perception"**

James J. Gibson

2014 | Psychology Press eBooks | DOI: 10.4324/9781315740218

Źródło

[3] **"The Senses Considered as Perceptual Systems"**

Charles K. West, James J. Gibson

1969 | Journal of Aesthetic Education | DOI: 10.2307/3331482

Źródło

[4] **"Observations on active touch."**

James J. Gibson

1962 | Psychological Review | DOI: 10.1037/h0046962

Źródło

[5] **"The Perception of the Visual World"**

James J. Gibson

1951 | The American Journal of Psychology | DOI: 10.2307/1419017

Źródło

[6] **"The Embodied Mind: Cognitive Science and Human Experience"**

Francisco J. Varela, Eleanor H. Rosch, Evan T. Thompson

2017 | Project Muse (Johns Hopkins University)

Źródło

[7] "The Embodied Mind"

Francisco J. Varela, Eleanor H. Rosch, Evan T. Thompson

1991 | The MIT Press eBooks | DOI: 10.7551/mitpress/6730.001.0001

Źródło

[8] "Autopoiesis and Cognition"

Humberto R. Maturana, Francisco J. Varela

1980 | Boston studies in the philosophy of science | DOI: 10.1007/978-94-009-8947-4

Źródło

[9] "Autopoiesis and Cognition: The Realization of the Living"

Humberto R. Maturana, Francisco J. Varela, Stafford Beer

1980

Źródło

[10] "Machine Learning Study: Identification of Skin Diseases for Various Skin Types Using Image Classification."

Gulhan Bizel, A. Einstein, Amey G Jaunjare, S. Jagannathan

2024 | Journal of Big Data and Artificial Intelligence | DOI: 10.54116/jbdai.v2i1.32

Źródło

[11] "A BEHAVIORAL MODEL OF RATIONAL CHOICE"

Herbert A. Simon

1955 | The Quarterly Journal of Economics | DOI: 10.2307/1884852

Źródło

[12] "Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization."

Michael Brown, Herbert A. Simon

1978 | Contemporary Sociology A Journal of Reviews | DOI: 10.2307/2065048

Źródło

[13] "Perception in chess"

William G. Chase, Herbert A. Simon

1973 | Cognitive Psychology | DOI: 10.1016/0010-0285(73)90004-2

Źródło

[14] "Models of Bounded Rationality"

Herbert A. Simon

1997 | The MIT Press eBooks | DOI: 10.7551/mitpress/4711.001.0001

Źródło

[15] "Descarte's error : emotion, reason, and the human brain"

António R. Damásio

1994 | DigitalGeorgetown (Georgetown University Library)

Źródło

[16] "Insensitivity to future consequences following damage to human prefrontal cortex"

Antoine Bechara, António R. Damásio, Hanna Damásio, Steven W. Anderson

1994 | Cognition | DOI: 10.1016/0010-0277(94)90018-3

Źródło

[17] "Subcortical and cortical brain activity during the feeling of self-generated emotions"

António R. Damásio, Thomas J. Grabowski, Antoine Bechara, Hanna Damasio, Laura L. Boles Ponto

2000 | Nature Neuroscience | DOI: 10.1038/79871

Źródło

[18] "Self comes to mind : constructing the conscious brain"

António R. Damásio

2010

Źródło

[19] "Emotion Circuits in the Brain"

Joseph E. LeDoux

2000 | Annual Review of Neuroscience | DOI: 10.1146/annurev.neuro.23.1.155

Źródło

[20] "The Emotional Brain: The Mysterious Underpinnings of Emotional Life"

Joseph E. LeDoux

2004 | Medical Entomology and Zoology

Źródło

[21] "The Emotional Brain"

Joseph E. LeDoux

1996

Źródło

[22] "Contributions of the Amygdala to Emotion Processing: From Animal Models to Human Behavior"

Elizabeth A. Phelps, Joseph E. LeDoux

2005 | Neuron | DOI: 10.1016/j.neuron.2005.09.025

Źródło

Treść tworzy Fundacja Dobre Państwo.
Redaguje i publikuje APA ONE, autorski system redakcyjny Fundacji oparty na AI.
APA ONE Publishing System
Fundacja Dobre Państwo © 2026
Article ID: 7c3cb0e5-c0e3-4c7a-9e32-97ff25d17f25