

# Fizyka umysłu

Włodzisław Duch

*Katedra Informatyki Stosowanej, Uniwersytet Mikołaja Kopernika, Toruń*

## 1. Wstęp

Wielkim wyzwaniem dla fizyki jest stworzenie takiego modelu świata, który da się zrozumieć. Istotną częścią tego świata są nasze umysły. Ich natura staje się coraz bardziej zrozumiała dzięki postępom neuronauk (ang. neurosciences), czyli różnych gałęzi nauki badających budowę i aktywność neuronów mózgu. Fizyka jest oczywiście podstawą większości metod badań doświadczalnych, w tym metod obrazowania aktywności mózgu, ale tu skupię się tylko nad zagadnieniami teoretycznymi.

Mózg ludzki jest najbardziej skomplikowanym obiektem w znanym Wszechświecie, a umysł – obiektem najbardziej tajemniczym. Przez „umysł” rozumiemy zespół zdolności umożliwiających powstanie subiektywnego obrazu świata dzięki procesom poznawczym i aktywnej eksploracji. Specyficzną własnością umysłu jest świadomość, zdolność do „zdawania sobie sprawy” z własnych procesów poznawczych. Czy fizyka wystarczy, by zrozumieć działanie mózgu i umysłu? Nie jest to przekonanie wśród fizyków powszechne [1]. Podobne wątpliwości na początku XX w. dotyczyły natury życia [2]. Zjawiska mentalne i procesy zachodzące w mózgu mają całkiem odmienną naturę. Jak należy rozumieć relacje pomiędzy mózgiem a umysłem, jak je badać? Podobnie jak nawet doskonała teoria cząstek elementarnych nie wystarczy do zrozumienia aktywności biologicznej białek, tak i badania mózgu na poziomie molekularnym lub komórkowym niewiele wniosą do zrozumienia umysłu.

Procesy mentalne są wynikiem neurodynamiki na poziomie globalnej aktywności bioelektrycznej mózgu [3]. Takie ujęcie otwiera drogę do tworzenia modeli operujących językiem redukwalnym do zjawisk neurofizjologicznych z jednej

strony i rozszerzalnym do pojęć psychologicznych z drugiej.

## 2. Relacja umysł–mózg

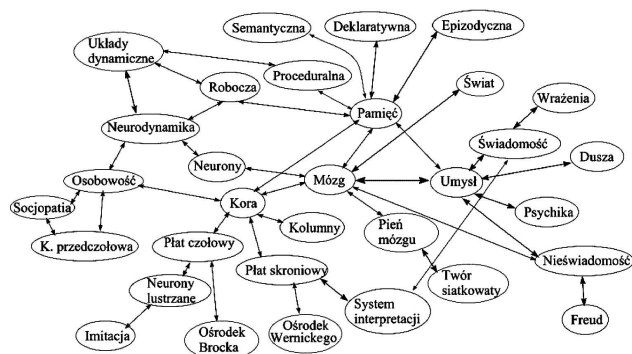
Neuropsychologia [4] bada związki między zaburzeniami funkcji psychicznych a uszkodzeniami różnych obszarów mózgu i ich połączeń. Związki te można często rozumieć jako specyficzne zaburzenia przetwarzania informacji, prowadzące do zaskakujących zmian w sposobie działania umysłu (np. dysocjacji pomiędzy rozpoznawaniem afektywnym i kognitywnym). Modele neuronowe chorób psychicznych i syndromów neuropsychologicznych pozwalają zrozumieć (przynajmniej jakościowo) przyczyny patologii i normalnego funkcjonowania mózgu [5]. Mamy obecnie bardzo bogaty materiał doświadczalny, wymagający teoretycznego uporządkowania.

Mózg jest substratem, w którym zachodzą procesy konieczne do powstawania procesów mentalnych. Potrzeby organizmu i jego możliwości poznawcze, znajdujące odbicie w strukturze mózgu, stwarzają ramy dla powstania subiektywnego obrazu świata. Struktura umysłu jest rezultatem niepowtarzalnej historii jednostki. Procesy neurofizjologiczne odpowiedzialne za zdarzenia mentalne mają rację bytu tylko ze względu na istnienie umysłu i mają sens tylko na poziomie umysłu. Dlatego nie można w pełni wyjaśnić stanów umysłu, sprowadzając je do stanów mózgu. Modele umysłu wymagają innego poziomu opisu niż modele mózgu.

Niezwykle złożone oddziaływania ogromnej liczby neuronów tworzą nowy, emergentny poziom organizacji. Wewnętrzną dynamikę stanów mózgu w przypadku swobodnym (bez bodźców zewnętrznych) można scharakteryzować za pomocą prawdopodobieństw przejść między stanami neurody-

namiki opisującej pobudzenia grup neuronów [5]. Takie stany nie wykazują podobieństwa do stanów umysłu, składających się ze wspomnień, epizodów, myśli, wyobrażeń. Elementy te będą nazywał „obiettami umysłu”. Relacje między stanami mentalnymi powinny mieć analogiczną strukturę do tej, jaka istnieje między stanami mózgu.

Umysł rozumiany jako zbiór stanów, relacji między tymi stanami i funkcji poznawczych, pozwalających na tworzenie nowych stanów, jest czymś abstrakcyjnym i niematerialnym. Budowa mózgu ogranicza możliwe abstrakcyjne stany umysłu do takich, które można fizycznie zrealizować. Najprostszy statyczny model umysłu ma postać grafu (sieci powiązań przyczynowych lub sieci Bayesa [6]), którego węzły reprezentują potencjalnie dostępne stany, a łuki – prawdopodobieństwa przejść między nimi (rys. 1). Pełne odtworzenie relacji zachodzących pomiędzy stanami ludzkiego umysłu wymaga znajomości ludzkiego mózgu, ciała i historii rozwoju. Możemy jedynie tworzyć proste modele, tłumaczące niektóre aspekty działania ludzkiego umysłu w dobrze określonych warunkach.



Rys. 1. Sieć powiązań między różnymi pojęciami dotyczącymi umysłu. Istotne są relacje, a nie sposób bezpośredniej reprezentacji pojęć.

Dlaczego niektóre procesy mentalne są związane ze świadomością, a inne nie [7]? Pamięć długotrwała ma charakter epizodów, zapamiętanych przeżyć, które możemy sobie uświadomić jako zespół wrażeń. Konieczna jest do tego aktywacja śladów pamięci i przeniesienie jej do pamięci roboczej [8]. Pamięć semantyczna związana jest z hierarchizacją pojęć i relacjami gramatycznymi, których używamy, chociaż zwykle nie jesteśmy ich świadomi. Podobnie jest z pamięcią proce-

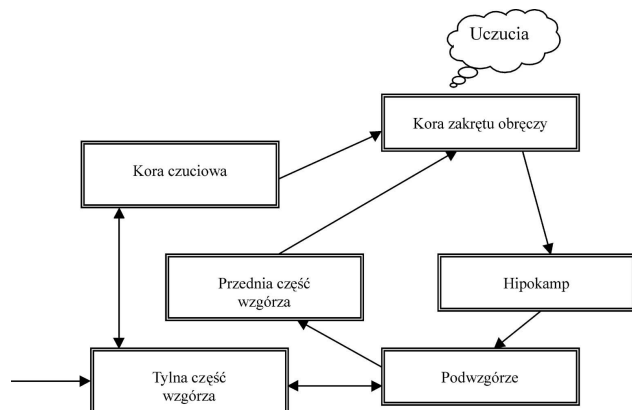
duralną, np. umiejętności pływania czy jazdy na rowerze. Nie mamy świadomych wrażeń związanych z umiejętnością wiązania krawata czy sznurowania butów (wrażenia związane są z epizodem, ale nie z umiejętnością) i nie da się takich umiejętności nauczyć, czytając werbalne opisy. Dlaczego stan jawy związany jest ze świadomością, a koma lub stan snu głębokiego – nie? Z tego rodzaju badań wyłania się obraz procesu zwanego świadomością, pozwalający na coraz bardziej precyzyjne zdefiniowanie samego zjawiska.

Umysł jest związany z tą częścią zachodzących w mózgu procesów, które mogą stać się świadome. Wiele czynności po ich doskonałym opanowaniu przestaje wymagać świadomych decyzji – dotyczy to zarówno czynności manualnych jak i percepcji. Również pamięć jest często niezależna od naszej woli, gdyż pomimo wysiłków nie możemy sobie czegoś przypomnieć, a poszukiwana informacja pojawia się później spontanicznie. Działania nieświadome są złożone z automatyzmów, leżących u podstaw zachowania. Nieświadome działania mózgu aktywnie wpływają na postać wrażeń świadomych, percepcji, pamięci i zachowania. Pojęcie nieświadomości przedstawione w psychoanalizie Freuda jest całkiem odmienne. Jego idee oparte były na błędnych analogiach hydraulicznych powstawania „ciśnienia psychicznego”, które musiało znaleźć odpowiednie ujście.

### 3. Relacyjna teoria umysłu

Kiedy poziom glukozy we krwi spada, zwierzę musi wiedzieć, że czas szukać jedzenia. Stan krwi monitorowany jest przed podwzgórze, składające się z wielu wyspecjalizowanych, drobnych skupisk neuronów. Sygnały wysyłane do kory przez jego jądro boczne interpretowane są jako uczucie głodu. Interpretacją zajmuje się kora mózgu, w tym przypadku kora zakrętu obręczy (rys. 2), leżąca na powierzchni boczno-przyśrodkowej mózgu [9]. Wynikiem specyficznych uszkodzeń wtórnej i trzeciorzędowej kory sensorycznej są agnozje, czyli niezdolności do rozpoznawania obiektów na podstawie informacji zmysłowej. Uszkodzenie kory obręczy może spowodować asymbolię bólu, czyli niezdolność interpretacji sygnałów bólu jako nieprzyjemnych, przy w pełni zachowanej zdolności do czucia i umiejscawiania tych sygnałów. Uszkodzenie dolnej części zakrętu skro-

niowego może spowodować prosopagnozę, czyli niezdolność do rozpoznawania twarzy. Uszkodzenia obszaru ciemieniowego w części przyśrodkowej prowadzi do astereognozji, czyli niezdolności do rozpoznawania przedmiotów za pomocą dotyku.



Rys. 2. Poziom glukozy monitorowany jest przez podwzgórze; impulsy przez nie wysyłane docierają przez przednią część wzgórza do kory zakrętu obręczy, gdzie interpretowane są jako uczucie głodu.

Uczucie pragnienia czy głodu, wrażenia zmysłowe, wyobrażenia, to rezultaty interpretacji dokonywanej przez korę mózgu [4]. Pamięć pozwala odtworzyć stany, w których mózg się poprzednio znajdował, umożliwiając dokonywanie interpretacji dzięki skojarzeniom i zdolności rozróżniania np. stanu „przyjemna czerwień zachodzącego słońca” i „czerwień policyjnego lizaka”. W odróżnieniu od stanu procesora komputera stan dynamiczny mózgu dzięki pamięci jest zawsze „ubrany” w skojarzenia. Jeśli relacje między obiektami umysłu pozostają podobne, wszyscy kojarzą podobnie słowa pisane (analizowane przez układ wzrokowy w korze potylicznej) z brzmieniem i sensem (analizowanym przez układ słuchowy w korze skroniowej) i wymową zależną od ruchu strun głosowych (kontrolowaną przez układ ruchowy w korze czołowej).

Jedną z podstawowych cech decydujących o inteligencji jest zdolność przechowywania informacji o bieżącej sytuacji w pamięci roboczej, utrzymywania ich „w umyśle”. Bez pamięci roboczej [8] niemożliwa byłaby wyobraźnia, wykroczenie poza bieżącą chwilę („teraz i tutaj”), a więc również i świadomość. Świadomość jest procesem wyróżniającym pewne elementy pamięci roboczej, pozwalającym na wykorzystanie znajdującej się

w niej informacji do działania. To działanie niekoniecznie musi wyrażać się skurczami mięśni, może to być rozpoczęcie planowania, umożliwienie zapamiętania lub przywołanie jakiegoś skojarzenia z pamięci. W ten sposób powstaje „strumień świadomości”, ciąg niewypowiedzianych zdań (myśli), obrazów, pobudzeń emocjonalnych.

Czy są jakieś cechy aktywności mózgu, które jednoznacznie korelują się z subiektywnymi wrażeniami? W eksperymentach z rywalizacją obuczną [10] małpa widzi każdym okiem inny obraz, ale jest świadoma przez kilka sekund tylko jednego z nich, co sygnalizuje odpowiednim przyciskiem. Aktywność 90% neuronów w obszarze dolnego i górnego zakrętu skroniowego (IT, STS) jest skorelowana z postrzeżeniem konkretnego obiektu. Podobne wyniki otrzymano za pomocą metod magnetoencefalograficznych (MEG) u ludzi.

Aktywność neuronów nie tworzy wrażeń wzrokowych, ale jest konieczna do eksploracji wzrokowej świata, interpretacji stanów układu wzrokowego umożliwiającej działanie [11]. Strumień świadomości jest ciągiem interpretacji stanów części pamięci roboczej, dokonywanych przez system interpretacji związany z działaniem lewej półkuli i płatów czołowych [12]. Pamięć robocza pozwala na rozpowszechnienie informacji o bieżącej sytuacji organizmu do wszystkich części mózgu, dzięki czemu informacja ta może zostać uzupełniona lub zmieniona w procesach kojarzeniowych. Istotną rolę w powstaniu systemu interpretacji grają neurony brzusznej części kory przedruchowej, aktywne w czasie wykonywania specyficznych ruchów, np. łapania, podnoszenia lub popychania ręką. Te same neurony u małpy uaktywniają się również w czasie obserwacji innej małpy lub eksperymentatora wykonującego te same czynności [13]. Podstawą uczenia się przez imitację, zdolności do empatii i rozumienia innych istot jest możliwość powstawania podobnych procesów zachodzących w obu mózgach.

Wrażenia zmysłowe są wynikiem niewerbalnej interpretacji stanu pierwotnej i wtórnej kory zmysłowej. Kiedy rozmawiamy w hałaśliwym miejscu, z szumu wybierane są odpowiednie formanty dzięki oczekiwaniom i najbardziej wiarygodnej interpretacji wyników. Emocje są wynikiem interpretacji stanu ośrodków podkorowych, jak ciało migdałowate w przypadku strachu, oraz wpływu tych ośrodków na stany pamięci roboczej.

Nie można ich oddzielić od procesów poznawczych – pobudzenia emocjonalne ukierunkowują działanie procesów poznawczych, pozwalając na dokonanie wyboru pomiędzy różnymi działaniami.

System interpretujący ma teorię „ja”, pozwalającą mu odróżniać procesy odnoszące się do własnego organizmu oraz do własnej teorii/wiedzy o świecie [12]. Aktywność kory przedczołowej umożliwia podjęcie działania, interpretowanego jako wolny wybór dokonany przez „ja”, proces wolicjonalny. Sam proces interpretacji trwa ok. pół sekundy – pobudzenie kory motorycznej widoczne jest o tyle wcześniej w stosunku do momentu uświadomienia sobie, że zamierzamy poruszyć ręką. Potknięcie na schodach prowadzi do automatycznego działania, strach i uświadomienie sobie niebezpieczeństwa pojawia się później. Świadomość nie jest czynnikiem sprawczym, podejmującym decyzję, lecz jedynie wynikiem interpretacji stanu mózgu. Wiele wyspecjalizowanych procesów w mózgu zachodzi bez konieczności komentowania skutków ich działania przez system interpretacji, dlatego nie są one uświadamiane.

#### 4. Od mózgu do umysłu

Modele pamięci i funkcji poznawczych dotyczą działania mózgu. Zrozumienie umysłu wymaga „wejścia do środka” sieci relacji pomiędzy stanami mózgu. Czy można tak opisać neurodynamikę, by dostrzec w niej zjawiska mentalne? Umysł jest częścią tego, co robi mózg lub, wyrażając to metaforycznie, umysł jest cieniem neurodynamiki. Nawiązując do słynnej alegorii Platona możemy powiedzieć, że widzimy cienie prawdziwej rzeczywistości na ścianie jaskini. Prawdziwa rzeczywistość odbija się w „aktywnym lustrze” naszych mózgów, a to, co dostrzegamy w naszym umyśle, jest cieniem tego odbicia.

Próby zrozumienia świata umysłu na podobieństwo świata fizycznego czynione były w psychologii od dawna. Kurt Lewin już w 1938 r. proponował rozpatrywanie zdarzeń mentalnych w odpowiedniej przestrzeni psychologicznej, w której agent (autonomicznie działający system poznawczy) poddawany jest działaniu „sił kognitywnych” [14]. George Kelly zaproponował analizę zjawisk psychologicznych korzystającą bardziej z geometrii niż z logiki [15] w ramach swojej „psychologii osobistych konstruktów”. Kelly

chciał doprowadzić do powstania oprogramowania dla symulacji osobowości i wizualizacji wielu procesów psychologicznych. Proponowano, by przyjąć jego idee za centralny model nauk kognitywnych [16], nie są one jednak powszechnie znane.

Roger Shepard usiłował znaleźć w świecie umysłu uniwersalne prawa, na wzór praw fizycznych [17]. Według niego „nie potrzeba nam więcej danych lub bardziej szczegółowych danych, lecz innego podejścia do problemu”. W fizyce opis ruchu jest prosty dzięki symetriom, pozwalającym zdefiniować niezmienniki ruchu w odpowiednich przestrzeniach. W przestrzeniach euklidesowych można zdefiniować transformację Galileusza, w pseudoeuklidesowych  $(3+1)$  transformację Lorentza, a w przestrzeniach Riemanna – transformacje w układzie nieinercyjnym. Prawa psychologiczne należy również formułować w odpowiednich przestrzeniach. Jako przykład Shepard [18] pokazuje dane psychofizyczne dla zwierząt i ludzi, dotyczące uogólniania bodźców wizualnych, słuchowych i smakowych. Dla  $N$  par bodźców bada się generalizację  $G$ , czyli prawdopodobieństwo tego, że zachowanie wyuczone dla jednego z nich pojawi się dla drugiego. Procedura skalowania wielowymiarowego, stosowana często w psychometrii, przekształca te dane, porządkując je monotonicznie, tak by wzajemne relacje podobieństwa przedstawić w postaci odległości  $D$  w jednym wymiarze. Dla dowolnego eksperymentu tego typu otrzymujemy zależność wykładniczą  $G(D) = \exp(-\alpha D)$ .

Skąpe informacje docierające z siatkówki wystarczają do odtworzenia wrażenia ciągłego obrazu poruszającego się przedmiotu dzięki internalizacji własności geometrycznych przestrzeni euklidesowej. Obiekty nie znikają w przypadkowy sposób, więc mózg wykorzystuje „zasadę zachowania obiektów”, tworząc wrażenie pozornego ruchu, np. wtedy, gdy dwa znajdujące się obok siebie podobne obiekty naprzemiennie znikają i pojawiają się. Jeśli obiekty używane do wywołania wrażenia ruchu pozornego mają różną orientację, to postrzegany ruch zachodzi zwykle wokół osi symetrii danego obiektu.

Intuicyjna fizyka ruchu wyobrażanych obiektów opiera się bardziej na geometrii kinematycznej (jest to gałąź matematyki zajmująca się opisem ruchu obiektów geometrycznych) niż na dy-

namice Newtona [18]. Trajektorie postrzeganego ruchu pozornego to linie geodezyjne obrotu i przesunięcia, a więc jest to ruch helikalny. Położenie sztywnego, asymetrycznego obiektu można opisać za pomocą położenia wybranego punktu i kątów określających orientację. Przestrzeń rozróżnialnych położenia jest 6-wymiarową rozmaitością, a 3 jej wymiary (rotacje) są zwinięte kołowo. Rozmaitość ta ma symetrię grupy będącej iloczynem półprostym dwóch grup  $E^+ = R^3 \otimes SO(3)$ . Linie geodezyjne w tej przestrzeni tworzą rodzinę jednoparametrowych podgrup, odpowiadających helikalnym trajektoriom geometrii kinematycznej. Dla obiektów posiadających częściową symetrię struktura przestrzeni ulega uproszczeniu. Można ją bezpośrednio powiązać z rezultatami eksperymentów psychofizycznych. Interpretacja kształtów obiektów jest bardziej skomplikowana i wymaga większej liczby wymiarów. I w tym przypadku można sobie wyobrazić proces rozpoznawania jako ruch po geodezyjnej trajektorii w stronę prototypu kształtu danego obiektu [18,19].

Charakteryzacja różnych aspektów przestrzeni, w których należy analizować wrażenia związane z percepcją, jest pierwszym krokiem w kierunku fizyki umysłu. Postrzeganie koloru – tematem tym zajmowali się: Newton, Helmholtz, Maxwell i Schrödinger – jest tu również dobrym przykładem. Pomimo zmiany rozkładu widmowego światła odbitego od powierzchni postrzegany kolor jest stabilny. Przestrzeń kolorów przedstawić można w postaci sfery, podając intensywność (zmienna radialna), nasycenie i barwę (zmiennie kątowne). Już Newton zauważył, że skrajne kolory widma, purpura i fiolet, są bardziej podobne do siebie niż do kolorów leżących w środku widma. Niezmienniczość postrzeganego koloru  $K(x, \lambda) = I(\lambda)S(x, \lambda)$  w punkcie  $x$  odbijającej powierzchni  $S$  przy różnym oświetleniu  $I(\lambda)$  wymaga odróżnienia wkładu  $S(x, \lambda)$  dzięki informacji zaledwie z trzech receptorów koloru. Rozważania teoretyczne pokazują [20], że do wykonania tego zadania potrzeba co najmniej 6 typów receptorów. Nie widzimy więc w pełni wszystkich cech powierzchni odbijającej. Zmienność naturalnych warunków oświetlenia można w przybliżeniu opisać za pomocą 3 parametrów: natężenia światła, przesunięcia czerwien–zieleń zależnego od absorpcji atmosfery oraz przesunięcia niebieski–żółty, zależnego od rozpraszania. Układ wzrokowy doko-

nuje kompensacji w tych trzech wymiarach, zachowując stałość postrzeganego koloru [21].

Makroskopowe własności obiektów i ich ruchu postrzegane na poziomie umysłu są wynikiem sposobu przetwarzania informacji na poziomie mikroskopowym przez sieci neuronowe mózgu. Struktura tych sieci wykształciła się w wyniku ewolucji, tak by z informacji dostarczanej przez zmysły wydobywać najbardziej przydatne i wiarygodne cechy. Świat umysłu wykorzystuje tylko te cechy, które są przydatne z punktu widzenia sprawnego działania. Porównywanie obiektów odgrywa podstawową rolę i dokonywane jest za pomocą najszybszych transformacji przy zachowaniu niezmienniczych własności obiektów. Ważną funkcję pełni tu pamięć.

## 5. Modele pamięci i sieci atraktorowe

Kora mózgu ma budowę modułarną – składa się z obserwowalnych pod mikroskopem kolumn o średnicy ułamka milimetra, zawierających ok.  $10^5$  neuronów. Typowa kolumna zawiera tysiące mikroobwodów, w niewielkim stopniu sprzężonych ze światem zewnętrznym. Neurony kilkunastu rodzajów pobudzają się wzajemnie za pomocą różnorodnych synaps, reagujących na dziesiątki neurotransmiterów i neuromodulatorów. Informacja przekazywana w genach nie wystarcza, by w pełni określić strukturę tak złożonego obiektu. Podstawowym zadaniem kolumny jest rozróżnianie dochodzących do niej sygnałów, tworzenie niepowtarzalnej sygnatury, którą inne kolumny będą mogły odczytać. Układ tysięcy przypadkowych mikrooscylatorów zaburzony przez zewnętrzne sygnały będzie przez krótki czas pobudzony w charakterystyczny sposób [22]. Odzworowuje to w nieliniowy sposób dochodzące sygnały w przestrzeń o dużej liczbie wymiarów. Inne kolumny kory otrzymują sygnały o jej aktywności i potrafią łatwo rozróżnić, w jakim jest on stanie, za pomocą zwykłego mechanizmu uczenia się dzięki wzmacnianiu siły połączeń synaptycznych jednocześnie aktywnych neuronów (nazywa się to uczeniem hebbowskim, na cześć Donalda Hebba [26]). Większość połączeń wewnątrz kolumn może być przypadkowa i nie musi podlegać uczeniu.

Maass i współpracownicy nazwali ten model „płynem neuronowym” [22] i pokazali, że

układ takich kolumn ma moc równą maszynie Turinga dla zadań obliczeniowych w czasie rzeczywistym. Taki układ działa stabilnie, nie wymagając powtarzalnych stanów wewnętrznych. Zastosowano go m.in. do identyfikacji słów reprezentowanych przez zaszumione ciągi impulsów. Reprezentacją rozpoznawanego obiektu jest chwilowy rozkład częstości oscylacji mikroobwodów. Można go przedstawić jako rozkład gęstości prawdopodobieństwa (PDF) pojawienia się sygnału w jakiejś części przestrzeni parametrów wewnętrznych kolumny. Wystarcza to do dyskryminacji różnych bodźców, rozpoznawania słów czy wzrokowego rozpoznawania obiektów lub ich fragmentów. Pierwotna kora zmysłowa zawiera wyspecjalizowane obwody wykrywające różne aspekty sygnału, np. kolor, nachylenie krawędzi, kontury, dominację oka, ruch. Kolumny mogą specjalizować się w wykrywaniu różnych cech sygnałów, gdyż otrzymują je z różnych obszarów kory zmysłowej. W ten sposób może działać wtórna kora zmysłowa, odpowiedzialna za nadawanie sensu odbieranym bodźcom. Jednakże pojawienie się jakiegoś obiektu w pamięci roboczej, jego uświadomienie, wymaga procesów zachodzących w dłuższej skali czasu, sprzężenia zwrotnego między kolumnami [23].

Nawet najprostsze modele neuronowe pamięci mają wiele własności typowych dla pamięci biologicznych. Pamięć zapisana jest w sposób rozproszony w wielu połączeniach synaptycznych, więc zniszczenie części sieci prowadzi do łagodnych zmian (ang. *graceful degradation*), a nie zapomnienia poszczególnych wzorców. Dzięki temu możliwe jest również rozpoznawanie uszkodzonych wzorców i skojarzenia nowych wzorców z najbardziej podobnymi wśród zapamiętanych. Adresowalność kontekstowa oznacza, że pobudzenie sieci fragmentem zapamiętanego sygnału prowadzi do przypomnienia całego wzorca. Uszkodzone modele neuronowe, zamiast przypominać sobie sekwencję zapamiętanych wzorców, mogą produkować halucynacje, fałszywe stany atraktorowe, poskładane z fragmentów zapamiętanych wzorców. Przepelnienie pamięci prowadzi do gwałtownego wzrostu liczby pomyłek i powstania fałszywych wzorców pamięci. Czas przypomnienia nie zależy od liczby zapamiętanych wzorców, w odróżnieniu od baz danych, gdy trzeba kolejno porównać zapamiętane wzorce.

Realistyczne modele pamięci wymagają uwzględnienia faktu, że oprócz pamięci długoterminowej, zapisanej w połączeniach neuronów kory mózgu, mamy również pamięć średnioterminową, w starej korze hipokampa. Zaletą dwóch systemów pamięci jest możliwość budowania stabilnego modelu świata, przenoszenia do pamięci trwałej tylko ważnych faktów, pasujących do wcześniej zapamiętanych, powolnej modyfikacji lub dodawania nowych faktów bez katastroficznej interferencji ze starymi. Wyjaśnienie własności różnych rodzajów amnezji: wstecznej, następczej, całkowitej przejściowej, trudności z uczeniem się, wymaga dodatkowo uwzględnianie wpływu neuromodulacji na plastyczność hipokampa i kory [5].

Sieci neuronowe stosowane dotychczas w modelowaniu funkcji mózgu są skrajnie uproszczonymi modelami. Amit i jego współpracownicy zrobili ważny krok na drodze ku stworzeniu prostych modeli sieci neuronowych poprawnych z neurobiologicznego punktu widzenia [23,24]. Neurony mają trzy typy połączeń: pobudzające i hamujące wewnątrz modułu (kolumny kory) oraz pobudzające dochodzące z zewnątrz (od długich aksonów komórek piramidowych, przekazujących impulsy od innych modułów). Kolumny kory składają się z tysięcy neuronów, a symulacje obejmują kilkaset tysięcy neuronów. W trakcie uczenia się moduł biorący udział w rozpoznawaniu początkowo zwiększa w nieselektywny sposób średnią częstość swojej impulsacji dla wszystkich pobudzających go sygnałów zewnętrznych. W wyniku zwiększania się siły połączeń synaptycznych na tle globalnej aktywności pojawiają się lokalne atraktory, wywołane przez specyficzną strukturę sygnału uczącego. W pierwszym etapie neurony kolumn reagują na nauczone bodźce tylko wówczas, gdy są pobudzane. W drugim etapie aktywność spoczynkowa rośnie do ok. 20 Hz i utrzymuje się długo po zniknięciu bodźca – jest to sygnał aktywnej reprezentacji bodźca w pamięci. W trzecim etapie powstają korelacje aktywności między kolumnami, co pozwala na odtworzenie kolejności pojawiania się obrazów [25].

W eksperymentach opóźnionego wyboru (ang. *Delayed Match to Sample, DMS*) makak uczy się zapamiętać sekwencję wielu przypadkowych kształtów (by uniknąć przypadkowych korelacji, stosuje się obrazy fraktalne); w fazie testu po krótkiej prezentacji kilku obrazów i przerwie

ok. 30 s ma pa ma pokazać właściwą sekwencję obrazów. W tym celu musi je utrzymać w pamięci roboczej, pobudzając odpowiednie kolumny kory zaangażowane w tworzenie śladów pamięci. Kolumny te znajdują się w obszarze zakrętu dolnego (IT) kory skroniowej oraz w korze przedczołowej (PF). Wysoka aktywność (20 Hz) neuronów w tych obszarach utrzymuje się do 30 s po prezentacji i widoczna jest w pomiarach potencjałów z wielu elektrod [25]. Korelacje czasowe wynikające z zapamiętania kolejności prezentacji obrazów po wielu powtórzeniach przechodzą w korelacje aktywności neuronów biorących udział w kodowaniu śladów pamięci.

## 6. Przestrzenie umysłu

Neurodynamika decyduje o zdarzeniach na poziomie umysłu: postrzeganych cechach, skojarzeniach, rozróżnianych obiektach i decyzjach podejmowanych na podstawie tych informacji. Opis na poziomie neurodynamiki zachodzących w umyśle zjawisk jest bardzo trudny. Opis na poziomie symbolicznym, za pomocą konstrukcji logicznych, jest z kolei zbyt uproszczony i często nienaturalny. Lewin próbował zdefiniować przestrzenie psychologiczne jako obszary, w których można umieścić elementy naszego doświadczenia, zdarzenia mentalne [14]. Jakich przestrzeni należy użyć, by dostarczyć ogólne prawa działania umysłu? Jeśli takie prawa istnieją, powinny być odbiciem ewolucyjnie wykształconych własności mózgu, internalizacją istotnej dla przetrwania naszego gatunku wiedzy o świecie.

Nie należy się spodziewać, by jeden model wystarczył do opisu wszystkich zjawisk – różne działy fizyki posługują się różnymi modelami. Przestrzeń i czas są od czasów Newtona areną zdarzeń fizycznych, pozwalając na ilustrację procesów mechanicznych. Chociaż początkowo myślano, że są to pojęcia dane a priori, Einstein wprowadził abstrakcyjną czasoprzestrzeń, w której czas i położenie związane są z wynikami pomiarów. Przestrzenie fazowe i przestrzenie Hilberta są konieczne, by uogólnić mechanikę newtonowską. Przestrzenie psychologiczne rozpięte na osiach związanych z rezultatami pomiarów psychofizycznych, traktowane jako arena zdarzeń umysłowych, są niezbędne, by stworzyć język umożliwiający precyzyjny opis zdarzeń mentalnych.

Poznanie jednego konkretnego przykładu obiektu nowego rodzaju pozwala na generalizację, stworzenie nowej kategorii – w sieci neuronowej każdy zapamiętany wzorzec ma swój basen atraktora. W przestrzeni zdefiniowanej przez postrzegane cechy można przypisać zapamiętanym obiektom rozkłady PDF przypisania danego obiektu do jakiejś kategorii przy coraz większych deformacjach różnych cech zapamiętanego obiektu. Odległości między różnymi obiektami  $O_i$ ,  $O_j$  będą maleć wraz z wzrastającym podobieństwem obiektów. Sensowne odległości  $D_{ij}$  można ustalić za pomocą procedury skalowania wielowymiarowego na podstawie postrzeganego podobieństwa. Jak pokazał doświadczalnie i teoretycznie Shepard, prawdopodobieństwo reakcji ludzi i zwierząt na wyuczony bodziec  $O_i$  przy pojawieniu się bodźca  $O_j$  wynosi  $P(O_i, O_j) = \exp(-\alpha D_{ij})$  dla szerokiej gamy bodźców [17]. Rozkłady PDF reprezentujące obiekty w przestrzeniach psychologicznych powinny więc zanikać wykładniczo. W modelu można użyć naturalnych z punktu widzenia neurobiologii funkcji sigmoidalnych [26], np.  $\sigma(x) = 1/[1+\exp(-x)]$ . Różnica  $\sigma(x+a) - \sigma(x+b)$  zanika w sposób wykładniczy, utrzymując stałą wartość w większym obszarze, w którym deformacje cechy  $x$  nie mają znaczenia dla przypisania obiektów do danej kategorii.

Modele zdarzeń mentalnych wykorzystujące przestrzenie psychologiczne mogą pozwolić na integrację informacji behawioralnej otrzymanej metodami psychologii eksperymentalnej oraz informacji dotyczących neurodynamiki otrzymanej metodami neurofizjologii. Od pomiarów aktywności neuronów możemy przejść do oceny siły bodźców, które należy następnie związać ze skalami dla poszczególnych wymiarów przestrzeni psychologicznej. Za pomocą wielu elektrod mierzymy aktywność neuronów  $r_i$  przy stosowaniu bodźca  $s$ , otrzymując prawdopodobieństwa  $P(r_i|s)$ . Korzystając z twierdzenia Bayesa, można teraz obliczyć prawdopodobieństwo posterioryczne  $P(s|r)$ , gdzie  $s$  oznacza stymulację, a  $r$  – reakcję:

$$\begin{aligned} P(s|r) &= P(s|r_1, r_2, \dots, r_N) \\ &= P(s) \prod_{i=1}^N P(r_i|s) / \sum_{s'} P(s') \prod_{i=1}^N P(r_i|s'). \end{aligned}$$

Stwarza to pomost między psychologią i neurofizjologią, prostszy niż próba użycia sieci neuro-

nowych do modelowania w psychologii, zachowując zalety modeli subsymbolicznych, posługujących się opisem ciągłych procesów, a więc bardziej szczegółowym niż opis za pomocą logiki. Upraszczając neurodynamikę, model powinien reprezentować jej atraktory za pomocą rozkładów PDF w niskowymiarowych przestrzeniach psychologicznych, obrazując powstawanie kategorii i relacje między nimi. Takie upraszczanie może przebiegać wielostopniowo: należy określić aktywność poszczególnych kolumn kory (odpowiadającą ustaleniu wartości poszczególnych cech obiektu), a następnie całą populacji.

W wersji statycznej model ten ma za zadanie opisywać behawioralne (sensomotoryczne) lub kognitywne (oparte na pamięci) reakcje mózgu, zachodzące w czasie nie dłuższym niż ok. 1 s. Skojarzenia modelowane są zgodnie ze strukturą przestrzeni psychologicznej odpowiedniej dla danego eksperymentu. Reprezentacja relacji, podobieństw między obiektami, wystarcza do zrozumienia kategoryzacji i rozpoznawania obiektów [27], a więc realizacji podstawowych procesów poznawczych. Analiza wielkiego zbioru tekstów, zawierającego 60 tys. różnych słów, doprowadziła Landauera i Dumais'go do wniosku, że wystarczy ok. 300 wymiarów, by jednoznacznie rozróżnić wszystkie słowa [28]. Nauczenie każdego nowego słowa ustawia je niejako w relacji ze wszystkimi już poznanymi. Dzięki temu, dowiadując się niewiele, dzieci szybko uczą się prawidłowo wielu relacji.

Teoria statyczna zastosowana została do problemów uczenia się kategorii [29]. W eksperymentach z kategoryzacją bada się szybkość uczenia i prawdopodobieństwa błędów w różnych warunkach. Pokazywane obiekty (np. karty lub figury geometryczne o kilku różnych kształtach, kolorach i wielkościach) dzieli się na kilka kategorii zgodnie z jakąś ukrytą regułą. Jeśli własności są binarne i tylko jedna z nich określa kategorię (np. kolor jest czarny lub biały), to łatwo jest zgadnąć odpowiednią regułą. Jeśli istotne są dwie własności, możemy mieć np. regułę typu XOR: kategoria A, jeśli obiekt jest czarny i duży lub obiekt jest biały i mały, a kategoria B dla kombinacji (czarny, mały) lub (biały, duży). Uczenie się właściwej kategoryzacji trwa wówczas dłużej. Teoria powinna przewidywać tempo uczenia się i prawdopodobieństwa błędów.

Złożone procesy neurodynamiczne mózgu realizują po nauczaniu prostą dynamikę. Dla reguł logicznych można napisać odpowiednia równania, zawierające istotne cechy i kategorie. Na przykład, dla problemów typu  $z = x \text{ xor } y$ , gdzie  $z = \pm 1$  jest kategorią, zaś  $x, y = \pm 1$  są cechami obiektów, od których zależy kategoryzacja, najprostsze równania dynamiki mają postać

$$\begin{aligned} V(x, y, z) &= 3xyz + \frac{1}{4}(x^2 + y^2 + z^2)^2, \\ \dot{x} &= -\frac{\partial V}{\partial x} = -3yz - (x^2 + y^2 + z^2)x, \\ \dot{y} &= -\frac{\partial V}{\partial y} = -3xz - (x^2 + y^2 + z^2)y, \\ \dot{z} &= -\frac{\partial V}{\partial z} = -3xy - (x^2 + y^2 + z^2)z. \end{aligned} \quad (1)$$

Na rysunku 3a przedstawiono trajektorie obrazujące baseny czterech atraktorów punktowych dla tych równań. Chociaż u różnych osób zachodzące w mózgu procesy uczenia będą się od siebie znacznie różnić, to efekt końcowy da się opisać za pomocą układu dynamicznego z 4 atraktorami (w dokładniejszym modelu nie będą one oczywiście punktowe). Jeszcze prościej można to opisać w przestrzeni psychologicznej rozpiętej na 3 użytych tu wymiarach (rys. 3b). Strzałki pokazują tu kierunek ruchu w stronę stanu końcowego  $(x, y, z)$  przy startowaniu z różnych stanów początkowych.

Doświadczenia z kategoryzacją dają czasami zaskakujące wyniki [30]. Na przykład, mamy dwie choroby, często występującą C z dwoma symptomami, SC i SI, oraz rzadką R z symptomami SR i SI. Po analizie listy przypadków uczestnicy eksperymentu odpowiadają prawidłowo na takie pytania, jak: jeśli zaobserwowano symptomy (SC,SI), to mamy prawdopodobnie chorobę... O jakiej chorobie świadczy sam symptom SR? Oczywiście o R. A o jakiej sam symptom SI? Większość odpowiada rozsądnie „C”, bo ta choroba występowała częściej. Podobnie dla kombinacji symptomów (SC,SI,SR) ok. 60% osób odpowiada „C”. Jednak dla kombinacji symptomów SC i SR taki sam procent osób odpowiada „R”. Dlaczego w tym przypadku ludzie nie stawiają na najczęściej występującą chorobę?

Sytuacja jest tu bardziej skomplikowana niż poprzednio, gdyż przestrzeń psychologiczna ma 5 wymiarów,  $\{C, R, SI, SC, SR\}$ . Baseny atraktorów neurodynamiki można ocenić na podstawie odpo-



wiedzi. Interpretacja psychologiczna jest tu mało przekonująca: symptom SR ma większe znaczenie, gdy pojawia się z SC, ponieważ jest to symptom wyróżniający, chociaż występuje rzadziej. Dlaczego w takim razie dodanie symptomu SI zmienia odpowiedzi?

w eksperymentach z kategoryzacją mamy okazję dostrzec załamanie się tych mechanizmów.

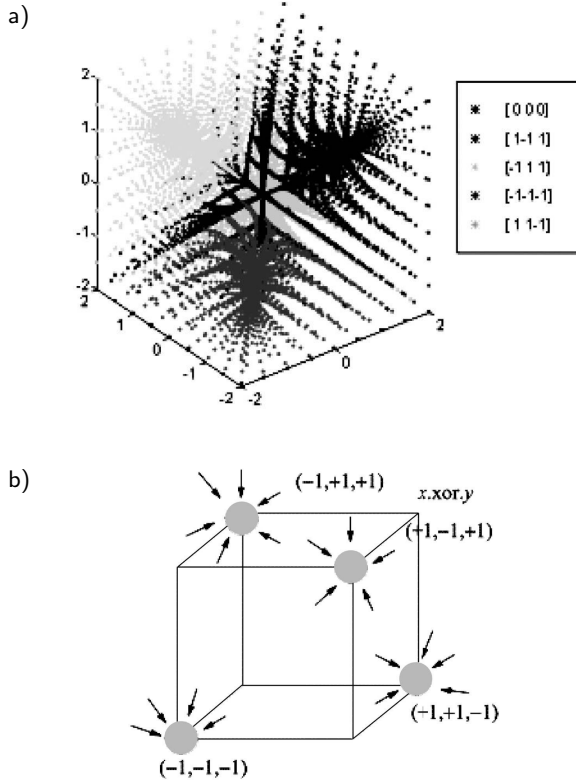
### 7. Ewolucja w przestrzeniach umysłu

Opisany model statyczny jest przydatny do interpretacji szybkich reakcji. W przestrzeniach psychologicznych określone są rozkłady prawdopodobieństw rozpoznania obiektów przy określonej kombinacji cech. Odpowiada to elementom trwałej pamięci, związanym z atraktorami neurodynamiki. Przestrzeń psychologiczna zawiera wymiary odpowiadające cechom wewnętrznym, które możemy traktować jako pomiary wykonane przez obszary mózgu zajmujące się przetwarzaniem bodźców zmysłowych. Umysł działa tu jak układ kontrolny, podejmując decyzje na podstawie wyników takich pomiarów. Przestrzeń umysłu powinna uwzględniać zarówno cechy wewnętrzne jak i podejmowane działania.

W doświadczeniach ze zwierzętami mamy do czynienia z sensomotorycznymi obiektami umysłu: po rozpoznaniu jakiegoś bodźca zwierzę wykonuje jedną z kilku wyuczonych czynności. Na przykład, w doświadczeniach z układem węchowym królików rozpoznawanych jest kilka rodzajów zapachów, na które królik reaguje w specyficzny sposób [31]. Zagadnienie można więc rozpatrywać w przestrzeni o liczbie wymiarów nie większej niż 10, podczas gdy procesy neurodynamiczne zachodzą w przestrzeniach o wymiarach rzędu setek milionów. W takich warunkach atraktory opisujące zachowanie trudno jest dostrzec wśród ogólnych zmian stanu kory mózgu.

W normalnych warunkach doświadczamy zmiennych stanów umysłu, określanymi mianem „strumienia świadomości”. Stany mózgu przechodzą w kolejne możliwe stany z pewnym prawdopodobieństwem. Chwilowy stan umysłu możemy powiązać z określonym punktem lub zlokalizowanym obszarem przestrzeni psychologicznej. Taki obszar odpowiada części aktywnej pamięci długotrwałej, „kopiowanej” do pamięci roboczej. Dynamika aktywacji kolejnych obiektów umysłu przedstawianych w przestrzeniach cech powinna odtwarzać dynamikę stanów mózgu (neurodynamikę).

Jeśli w chwili  $t_0$  obiektem aktywnym jest  $O_1$ , to stan umysłu reprezentowany jako punkt w przestrzeni psychologicznej znajduje się w obszarze dużych wartości PDF dla tego obiektu. Za-



Rys. 3. a) Trajektorie fazowe dla równań (1); b) stan końcowy w przestrzeni psychologicznej.

Rozwiązanie leży, jak się wydaje, w sposobie formowania się basenów atraktorów w czasie uczenia się. Kombinacja {SC,SI,C}, powtarzająca się często, tworzy szybciej duży basen, zaś basen dla {SR,SI,R} jest mniejszy. Aktywacja {SR,SC} przy nieokreślonym SI częściej prowadzi do odpowiedzi R, ponieważ gradient w kierunku R jest większy. Zwykle udaje się podać racjonalne przyczyny decyzji, ale w niektórych testach może to być trudne. Nie mamy wpływu na percepcję, np. ulegamy iluzjom wzrokowym, co ułatwia badanie mechanizmów działania układu wzrokowego. Podejmowanie decyzji przypomina percepcję, gdyż postrzegamy w swoim umyśle wynik przebiegających w mózgu procesów, a procesy interpretacji dostarczają nam racjonalizacji, odnosząc decyzję do modelu „ja” i swoich potrzeb. Być może

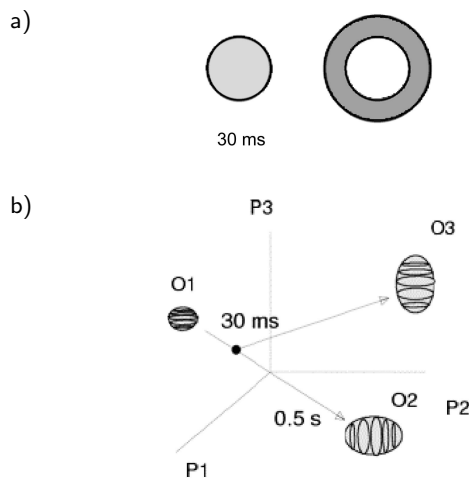
równy pod wpływem bodźców zewnętrznych jak i w wyniku wewnętrznej dynamiki stan ten będzie się zmieniał, podążając po „linii najmniejszego oporu” od obiektu do obiektu. W najprostszym przypadku ruch ten obrazuje proces aktywacji kolejnych elementów pamięci, np. serii wyuczonych zachowań.

Opis takiej dynamiki można wzorować na mechanice klasycznej. Stan umysłu ma pewną bezwładność, związaną ze stabilnością atraktora reprezentującego obiekt podlegający aktywacji. Można mu przypisać też pęd („kierunek myślenia”), utrudniający zmianę kierunku trajektorii. Przejście od aktywacji jednego obiektu do drugiego wymaga energii. Mózg jest układem otwartym, zużywającym aż 20% całej energii organizmu. Trajektorie obrazujące stan umysłu będą szybko przeskakiwać między obiektami i pozostawać w ich pobliżu.

Obiektom można przypisać pewne potencjały, określone przez funkcję zadaną w przestrzeni psychologicznej. Funkcja ta określa strukturę potencjalnie dostępnych stanów umysłu. Dynamikę w takim układzie można symulować podobnie jak dynamikę ładunku próbnego w układach elektrostatycznych, z uwzględnieniem sił stochastycznych i zewnętrznych sił wymuszających. Zadaniem takich modeli jest odtworzenie prawdopodobieństw przejść między atraktorami w mikroskopowym opisie neurodynamicznym, a więc sieci relacji między stanami umysłu i zachowaniami organizmu. W przestrzeni psychologicznej miarą odległości aproksymującą te prawdopodobieństwa będzie hamiltonowskie najmniejsze działanie, wyznaczające trajektorie geodezyjne. Taka miara odległości używana jest w przestrzeniach Finslera [32]. Nie jest ona symetryczna, albowiem nie zachodzą tu procesy odwracalne.

Przydatność takich pojęć, jak pęd czy bezwładność stanu umysłu można dostrzec przy analizie eksperymentów psychofizycznych dotyczących maskowania [33]. Jeśli po ekspozycji pierwszego obrazu, np. wypełnionego kółka, pokazany jest szybko drugi obraz, np. pierścienia, to obserwator widzi tylko drugi obraz (rys. 4a). Brak dobrego języka opisu takich zagadnień przyczynia się do kontrowersji typu: czy drugi obraz wymazuje wrażenia powstałe w wyniku pojawienia się pierwszego (zmienia przeszłość), czy też nie dopuszcza do jego powstania [33]? Proces ten można opisać

w sposób przypominający zderzenie lub rozproszenie (rys. 4b). Początkowo stan umysłu znajduje się w pobliżu O1 (patrzmy na pusty ekran), potem pojawia się bodziec O2, ale zanim stan umysłu znajdzie się w jego obszarze (rozpoznane zostanie kółko), bodziec maskujący O3 zmienia trajektorię stanu umysłu i nie dochodzi do aktywacji O2. Wrażenia związane z pierwszym bodźcem nie powstają, jeśli obiekt O3 jest podobny (bliski) O2. Obiekty całkiem odmienne, np. o innej modalności, nie mają wpływu na aktywizację O2. Nawet jeśli dojdzie do zmiany trajektorii – można sobie wyobrazić kaskadę takich rozprożeń blokującą powstawanie wrażeń – to część energii dociera do O2 i chociaż nie wystarcza to do aktywacji obiektu (przeniesienia go do pamięci roboczej), to może mieć wpływ na takie decyzje, jak zgadywanie, czy były dwa, czy jeden bodziec [33]. Podobnym językiem można się posłużyć przy analizie wielu doświadczeń psychofizycznych lub badaniach pamięci związanych z torowaniem (ang. priming).



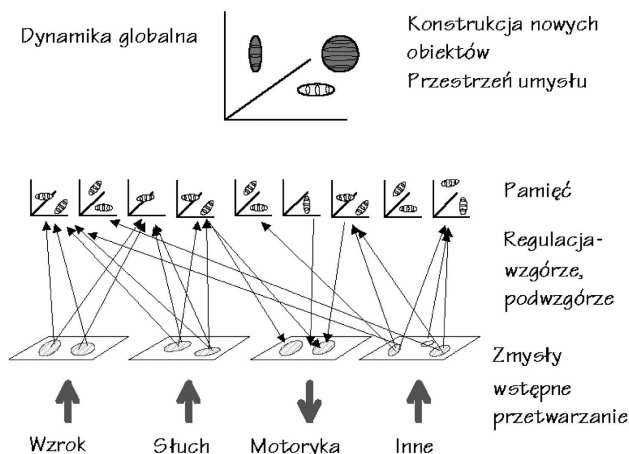
Rys. 4. a) Kółko i pierścień pokazany po 30 ms; b) schemat procesu maskowania w przestrzeni stanu umysłu [33].

Z powyższych rozważań widać, że przestrzenie psychologiczne mogą być przydatne w dobrze określonych sytuacjach eksperymentalnych. W geometrycznym modelu działania umysłu (rys. 5) przestrzenie psychologiczne wykorzystywane są na różnych etapach [2]:

— wykrywania istotnych cech sygnałów zmysłowych – odpowiada za to kora sensoryczna, a głównym mechanizmem analizy są mapy topograficzne;

— rozpoznawania podstawowych obiektów – kombinacje cech określają prawdopodobieństwo pobudzenia się reprezentacji obiektów w pamięci długotrwałej;

— pamięci roboczej – najbardziej aktywne elementy są kopiowane z pamięci długotrwałej do pamięci roboczej, umożliwiając podejmowanie decyzji i kontrolę zachowania.



Rys. 5. Wykorzystywanie przestrzeni psychologicznych w geometrycznym modelu działania umysłu.

## 8. Podsumowanie

Fizyka odwróciła się w znacznej mierze od zagadnień psychofizycznych dopiero na początku XX w. Działanie mózgu było wówczas całkowitą tajemnicą, a psychologia doświadczalna dopiero zaczynała się rozwijać. Obecnie sytuacja uległa radykalnej zmianie. Coraz lepiej rozumiemy działanie mózgu zarówno na poziomie mikroskopowym jak i systemowym. Neurofizjologia bada szczegóły procesów zachodzących w mózgu, ale modele umysłu to nie modele mózgu. Zrozumiałych modeli działania umysłu dostarczyć może tylko fizyka. W tym celu trzeba poszukiwać cech umożliwiających niskowymiarowe reprezentacje zdarzeń mentalnych i sposobów upraszczania neurodynamiki.

Próba opisu zjawisk mentalnych jest oczywiście znacznie trudniejsza niż typowych zjawisk fizycznych. Trudności techniczne wynikają zarówno z niepełnej wiedzy dotyczącej działania mózgu, sposobu analizy sygnałów zmysłowych mogących stanowić podstawę dla opisu obiektów umysłu, czy też samej złożoności modeli neurodynamicznych i sposobów ich upraszczania. Na najwyż-

szym poziomie niskowymiarowe przestrzenie psychologiczne opisujące pamięć roboczą można zdefiniować tylko zakładając szybką zmienność liczby i rodzaju cech rozpinających przestrzeń. Wybór obiektów pojawiających się w tej „przestrzeni umysłu” jest związany z działaniem mechanizmów uwagi, które dopiero uczymy się modelować.

Geometryczne unaocznienie nawet prostych eksperymentów wymaga wielowymiarowych przestrzeni. Przydatność geometrycznych modeli umysłu nie jest jeszcze oczywista; nie są też znane ich ograniczenia. Celem jest nie tyle wyjaśnienie wszystkich zagadek umysłu za pomocą jednego prostego modelu, co próba zmniejszenia przepaści między badaniami nad mózgiem a psychologią, przez dostarczenie precyzyjnego języka umożliwiającego analizę rezultatów konkretnych eksperymentów we wzorowany na fizyce sposób.

Rozszerzona wersja tego artykułu jest umieszczona w Internecie pod adresem [www.phys.uni.torun.pl/publications/kmk](http://www.phys.uni.torun.pl/publications/kmk).

## Literatura

- [1] R. Penrose, *Nowy umysł cesarza* (PWN, Warszawa 2000).
- [2] E. Schrödinger, *What is life* (Cambridge Univ. Press, 1944).
- [3] W. Duch, „Platonic model of mind as an approximation to neurodynamics”, w: *Brain-like computing and intelligent information systems*, red. S-i. Amari, N. Kasabov (Springer, Singapore 1997), s. 491.
- [4] K. Walsh, *Neuropsychologia kliniczna* (PWN, Warszawa 1998).
- [5] W. Duch, „Sieci neuronowe w modelowaniu zaburzeń neuropsychologicznych i chorób psychicznych”, w: *Biocybernetyka 2000*, t. 6: „Sieci neuronowe”, red. W. Duch, J. Korbicz, L. Rutkowski, R. Tadeusiewicz, rozdz. II.18.
- [6] P. Cichosz, *Systemy uczące się* (WNT, Warszawa 2000).
- [7] W. Duch, *Neurologia i Neurochirurgia Polska* **34**(50), suppl. 2, s. 69 (2000).
- [8] A.D. Baddeley, *Consciousness and Cognition* **1**, 3 (1992).
- [9] E. Kandel, J. Schwartz, T. Jessell, *Principles of neural science*, wyd. 3 (Prentice Hall Int., 1991).
- [10] D.A. Leopold, N.K. Logothetis, *Trends in Cognitive Sciences* **3**(7), 254 (1999).
- [11] J.K. O'Regan, A. Noë, *Behavioral and Brain Sciences* **24**(5) (2001), w druku.
- [12] M.S. Gazzaniga, *O tajemnicach ludzkiego umysłu. Biologiczne korzenie myślenia, emocji, seksualności*,

- języka i inteligencji (Książka i Wiedza, Warszawa 1997).
- [13] G. Rizzolatti, L. Fadiga, V. Gallese, L. Fogassi, *Cognitive Brain Research* **3**, 131 (1996).
- [14] K. Lewin, *The conceptual representation and the measurement of psychological forces* (Duke University Press, Durham, N.C. 1938).
- [15] G.A. Kelly, *The Psychology of Personal Constructs* (Norton, New York 1955).
- [16] M.L.G. Shaw, B.R. Gaines, *The New Psychologist* **10**, 23 (1992).
- [17] R.N. Shepard, *Science* **237**, 1317 (1987).
- [18] R.N. Shepard, *Psychonomic Bulletin and Review* **1**, 2 (1994).
- [19] M. Leyton, *Symmetry, causality, mind* (MIT Press, 1992).
- [20] L.T. Maloney, *J. Opt. Soc. Amer. A* **3**, 1673 (1986).
- [21] R.N. Shepard, „On the physical basis, linguistic representation, and conscious experience of colors”, w: *Conceptions of the mind*, red. G. Harman (Lawrence Erlbaum, Hillsdale, NJ 1993), s. 217.
- [22] W. Maass, T. Natschläger, H. Markram, „Real-time computing without stable states: A new framework for neural computation based on perturbations”, preprint (Technische Universität Graz, 2001).
- [23] J.A. Anderson, *An Introduction to Neural Networks* (MIT Press/Bradford Book, 1995).
- [24] D.J. Amit, *Behavioral and brain sciences* **18**(4), 617 (1994).
- [25] D.J. Amit, N. Brunel, *Cerebral Cortex* **7**, 237 (1997).
- [26] V. Yakovlev, S. Fusi, E. Berman, E. Zohary, *Nature Neuroscience* **1**(4), 310 (1998).
- [27] S. Edelman, *Representation and recognition in vision* (MIT Press, 1999).
- [28] T.K. Landauer, S.T. Dumais, *Psychological Review* **104**, 211 (1997).
- [29] W. Duch, „Categorization, Prototype Theory and Neural Dynamics”, w: *Proc. 4th Int. Conf. on Soft Computing*, Iizuka, Japonia 1996, s. 482.
- [30] D.L. Medin, S.M. Edelson, *Journal of Experimental Psychology: General* **117**, 68 (1988).
- [31] W.J. Freeman, *Societies of Brains* (Lawrence Erlbaum, 1995).
- [32] P.L. Antonelli, R.S. Ingarden, M. Matsumoto, *The Theory of Sprays and Finsler Spaces with Applications in Physics and Biology* (Kluwer Academic, Dordrecht 1993).
- [33] B.G. Breitmeyer, *Visual Masking* (Clarendon Press, Oxford 1984).
- [34] D.C. Dennett, M. Kinsbourne, *Behavioral and Brain Sciences* **15**(2), 183 (1995).