

Sztuczne sieci neuronowe

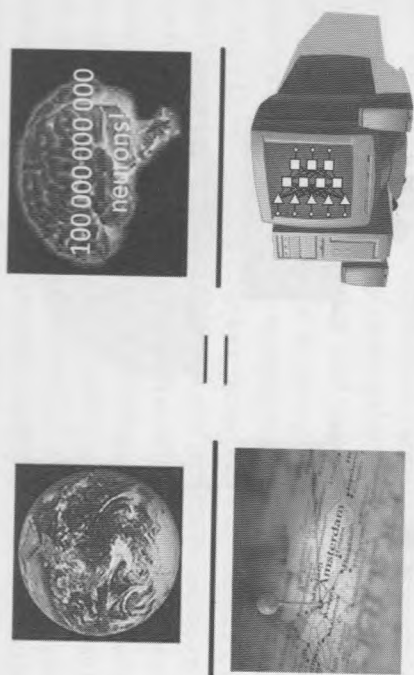
Wstęp

Sieci neuronowe są od lat cenionym narzędziem informatycznym. Teoretyczne podstawy ich budowy i stosowania są raczej skomplikowane, z jednej strony odwołują się bowiem do wyższej matematyki, a z drugiej czerpią inspirację z dość szczegółowej wiedzy neurobiologicznej. Proba przywoływania ich tutaj w ramach *Przewodnika po kognitywistyce* byłaby równie karkołomna dla autora dla czytelników. Czytelnicy spragnieni wiedzy szczegółowej na ten temat mogą sięgnąć do jednej z licznych książek poświęconych teorii sieci neuronowych i przykładom ich stosowania (np. mogą skorzystać z bardzo obszernego dzieła Tadeusiewicz i in. 2013). Jest jednak wygodniejszy sposób na dokładne poznanie sieci neuronowych, bez zagłębiania się w niepotrzebną kognitywistom teorię. Otóż budowę i możliwości sieci neuronowych najlepiej poznać poprzez samodzielne odkrywanie ich właściwości na własnym komputerze. Wystarczy do tego laptop czy tablet, chociaż najwygodniej jest to robić przy użyciu solidnego komputera stacjonarnego. W celu rozpoczęcia mądrej zabawy z sieciami neuronowymi należy zajrzeć na stronę <http://home.agh.edu.pl/~rad//>.

Teoretyczne
podstawy

Programy
do samodzielnego
poznawania sieci
neuronowych

Rys. 13. Proporcje mózgu i sieci neuronowej są takie jak proporcje kuli ziemskiej i główki szpilki

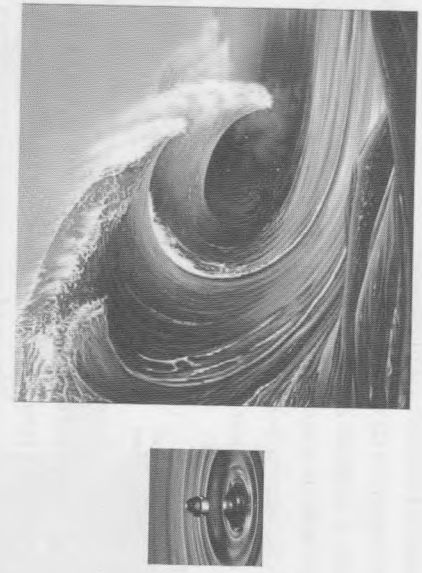


Moc eksplanacyjna sieci neuronowych

Czy wobec tego jest możliwe wyciągnięcie jakichkolwiek użytecznych wniosków na temat tak ogromnego systemu podczas badania tak małego jego modelu?

Odpowiedź na to pytanie jest twierdząca. Owszem, można badać zjawiska w bardzo małej próbce, wyciągając wnioski, które potem potwierdzają się w znacznie większej skali. Jeśli chemik przeszedł jakąś reakcję zachodzącą w małej kropli wody, to możemy być pewni, że dokładnie w taki sam sposób to samo zjawisko zachodzić będzie w ogromnym oceanie (rys. 14).

Rys. 14. Kropla i ocean – skala badanego obiektu nie wpływa na naukową ścisłość wyniku badania



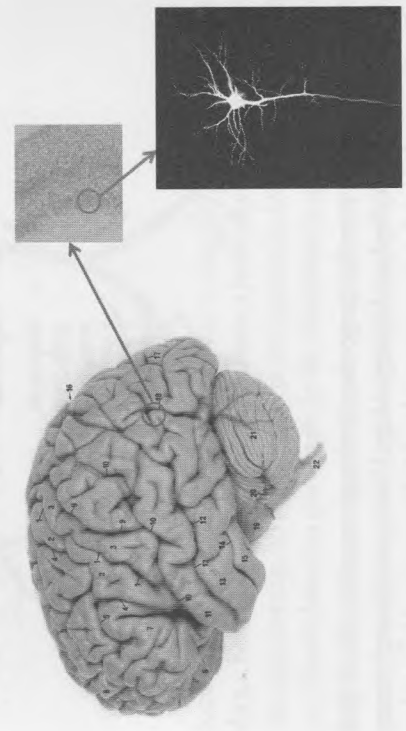
Jeśli zatem będziemy umieli zbudować sieć neuronową, w której wytworzymy pewne zjawiska nawiązujące do interesujących nas procesów psychicznych – to sieci te mogą stać się „poligonem”, na którym będziemy mogli sprawdzać koncepcje dotyczące natury i funkcjonowania elementów naszej psychiki. Poligon taki może być ogromnie użyteczny także w badaniach kognitywistycznych. Rozważmy to zagadnienie nieco dokładniej, bo jest ono w istocie kluczem do tego celu, który chcemy w tej pracy osiągnąć.

Pierwszy krok budowy sieci – definicja sztucznego neuronu

Nasładowując w komputerze ludzki mózg, staramy się wzorować na jego budowie. Analizując tę budowę coraz dokładniej, najpierw widzimy cały mózg, potem wycinek jego kory, a na końcu – pojedynczą komórkę nerwową będącą głównym elementem strukturalnym tej kory (rys. 15). Na rysunku należy zwrócić uwagę na zmienną skalę poszczególnych części: cały mózg ma rozmiar kilkunastu centymetrów, pokazany fragment kory mózgowej ma szerokość i wysokość 1 mm, a uwidoczniony pojedynczy neuron ma średnicę 20 mikrometrów.

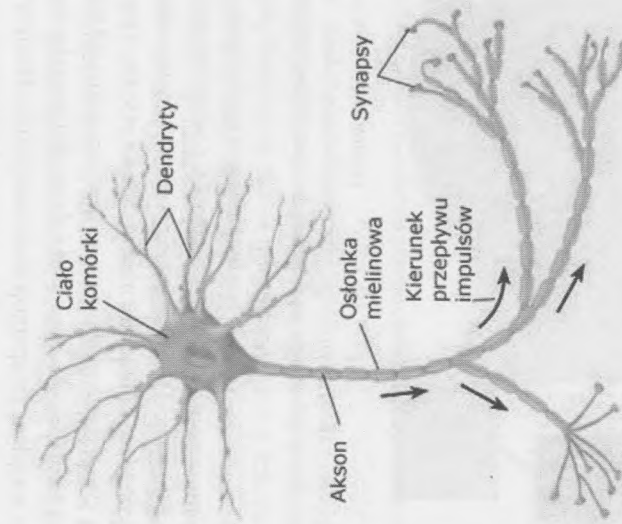
Neuron jako element mózgu

Rys. 15. Etapy poznawania budowy i elementów składowych mózgu (opracowano na podstawie: <http://www.wgabinecie.pl/others/image/brain2.jpg>, dostęp: 02.2013, i własnych materiałów autora)



Tworząc modele mózgu nadające się do umieszczenia w komputerze, zamyamy od podstawowego elementu składowego, jakim jest sztuczny neuron. Stanowi on bardzo uproszczony model rzeczywistego biologicznego neuronu. Rzeczywista komórka nerwowa, wchodząca w skład mózgu, ma dosyć skomplikowaną budowę, pokazaną na rysunku 16. Elementy składowe komórki pełnią następujące funkcje: dendryty zbierają sygnały wejściowe i doprowadzają je do ciała komórki, które jest biologicznym procesorem analizującym te sygnały. Wynik tej analizy zbierany jest przez akson, który jest izolowany jak kabel elektryczny za pomocą osłonki mielinowej. Impulsy wyjściowe neuronu są przekazywane do następnych neuronów za pomocą synaps, w których lokują się procesy uczenia, omówione dokładniej w dalszej części.

Rys. 16. Budowa biologicznej komórki nerwowej

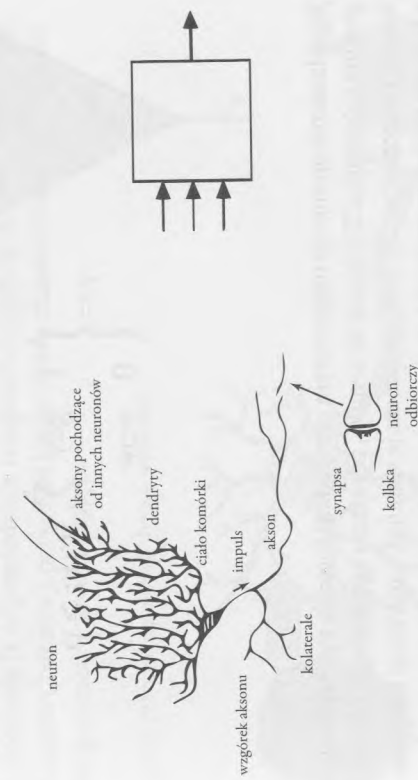


Na marginesie opisu pojedynczego rzeczywistego neuronu (biologicznej komórki nerwowej będącej częścią mózgu) warto odnotować jeszcze jedną ciekawostkę, widoczną na rysunku 16. Otóż ciała komórki są pozbawione osłonki mielinowej, więc tam, gdzie są one nagromadzone w dużych ilościach (głównie w korze mózgowej, ale także i w innych

strukturach anatomicznych, na przykład we wzgórzu albo w jądrach pnia mózgu), tkanka nerwowa ma charakterystyczną szarą barwę, bo taki jest naturalny kolor tych komórek. Natomiast tam, gdzie biegną aksony łączące neurony w sieci, kolor (biały z perłowym połyskiem) całej tkance nadają osłonki mielinowe przypisane do tych biologicznych „kablów połączeniowych”. Stąd już starożytni badacze mózgu wyróżniali w nim „substancję szarą” oraz „substancję białą”, zaś w wielu kontekstach mówi się o „szarych komórkach”, mając na myśli siedlisko naszej inteligencji.

Budując sztuczne neurony, z których są tworzone sieci neuronowe, staramy się w nich odwzorować jedynie najpotrzebniejsze cechy biologicznych neuronów. Z komórki, która ma skomplikowaną budowę, pozostaje więc tylko blok przetwarzający informacje (odpowiednik ciała komórki), duża liczba kanałów wejściowych, odpowiadających dendrytom, i jeden kanał wyjściowy, będący odpowiednikiem aksonu (rys. 17).

Rys. 17. Przejście pomiędzy biologicznym neuronem i jego sztucznym modelem



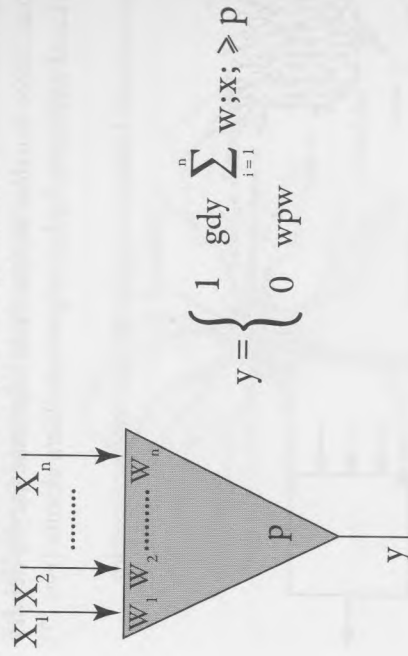
Sztuczny neuron musi się zachowywać w sposób maksymalnie podobny do zachowania biologicznego neuronu, a jednocześnie nie może być zbyt skomplikowany, bo jego realizacja techniczna będzie wtedy zbyt kosztowna. Trzeba bowiem pamiętać, że do zbudowania sieci neuronowej, która będzie służyć do rozwiązywania jakiegoś praktycznego zadania, będziemy potrzebowali na ogół kilkudziesięciu lub nawet kilkuset neuronów. Dlatego w dalszych rozważaniach przyjmujemy bardzo uproszczony sztuczny neuron, który uwzględnia jedynie trzy najprostsze właściwości swoich biologicznych odpowiedników. Są to:

Ilość sztucznych neuronów w sieci

- zdolność agregacji wielu sygnałów pochodzących z różnych źródeł (najczęściej od innych neuronów albo z wejść podających dane potrzebne do rozwiązania zadania),
- zdolność różnicowania sygnałów pochodzących z różnych źródeł na podstawie wiedzy gromadzonej w trakcie uczenia sieci,
- zdolność wypracowywania sygnału wyjściowego, kierowanego do dalszych elementów sieci lub wysyłanego na zewnątrz jako składnik rozwiązania stawianego sieci zadania.

Schematycznie sztuczny neuron (w jego najprostszej wersji) przedstawił rysunek 18.

Rys. 18. Schemat najprostszego sztucznego neuronu



Jego działanie opisuje (w najprostszym przypadku) formuła matematyczna podana na rysunku. Z formuły tej wynika, że neuron agreguje sygnały wejściowe, oznaczone x_1, x_2, \dots, x_n , po prostu je sumując, a różnicuje je, wiążąc z każdym sygnałem wejściowym współczynnik wagowy oznaczony odpowiednio w_1, w_2, \dots, w_n , oraz wypracowuje sygnał wyjściowy y . W przytaczanym tu prostym modelu odbywa się to w taki sposób, że sygnał wyjściowy wynosi 1 (co można interpretować jako wygenerowanie przez biologiczny neuron potencjału czynnościowego w postaci impulsu nazywanego w fizjologii *spike*), gdy sumaryczne pobudzenie neuronu przekroczy wartość progową oznaczoną p ; natomiast gdy sumaryczne pobudzenie jest za małe (w formule sytuację taką oznaczono skrótem *wpw*, czyli: w przeciwnym wypadku) – sygnał wyjściowy neuronu wynosi 0 (neuron pozostaje bezczynny).

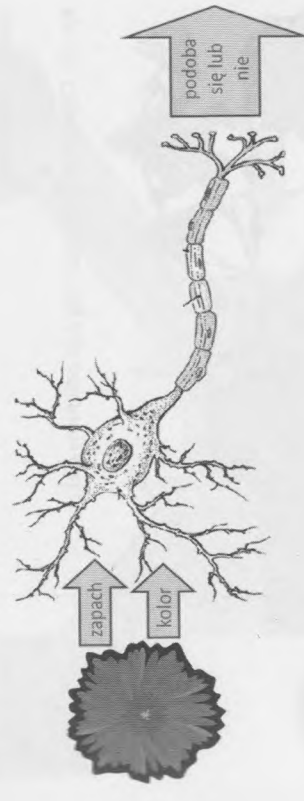
Dawniej taki opis aktywności neuronu uważano za wystarczający w świetle wyników ówczesnych badań neurofizjologicznych, a zasadę

generowania impulsu po przekroczeniu przez łączne pobudzenie zadane go progu lub zaniechania jakiegokolwiek akcji w przypadku, gdy pobudzenie to było słabsze, nazywano zasadą „wszystko albo nic”. Zgodnie z obecną wiedzą na temat funkcjonowania neuronu jego działanie jest znacznie bardziej skomplikowane, a możliwości przetwarzania sygnałów znacznie bogatsze, niż zakładano, ale w sztucznych sieciach neuronowych poprzestajemy na tym prymitywniejszym opisie, ponieważ jest on wystarczający do osiągnięcia wspomnianych celów praktycznych. Dla celów stawianych w tym rozdziale ten prosty model także będzie wystarczający.

Ważnym pojęciem, które pojawiło się w powyższym opisie, jest pojęcie wagi i sygnału wejściowego. Zgodnie z opisem i podaną na rysunku 18 formułą matematyczną sygnał docierający na określone wejście neuronu jest mnożony przez wspomnianą wagę, zatem może silniej albo słabiej oddziaływać na zachowanie neuronu. Ma to zasadnicze znaczenie dla zachowania neuronu i w konsekwencji – dla zachowania całej sieci. Zilustrujmy to przykładem, bo sprawa jest bardzo ważna.

Popatrzmy na rysunek 19. Przedstawiono na nim neuron, do którego docierają (z odpowiednich receptorów) sygnały dotyczące pewnych obiektów (kwiatów), przy czym każdy obiekt charakteryzowany jest przez dwa parametry: kolor i zapach.

Rys. 19. Przykład neuronu o dwóch wejściach



Neuron na podstawie tych sygnałów wejściowych musi wypracować własny sygnał wyjściowy, który będziemy interpretowali jako aprobatę (sygnał wyjściowy wynosi 1) lub dezaprobatę dla badanego obiektu (sygnał wynosi 0).

Wyobraźmy sobie, że badanym obiektem jest kwiat mający ładny kolor (sygnał wejściowy dodatni) i brzydki zapach (sygnał wejściowy pochodzący od tego receptora jest ujemny). Wówczas gdy do wejścia

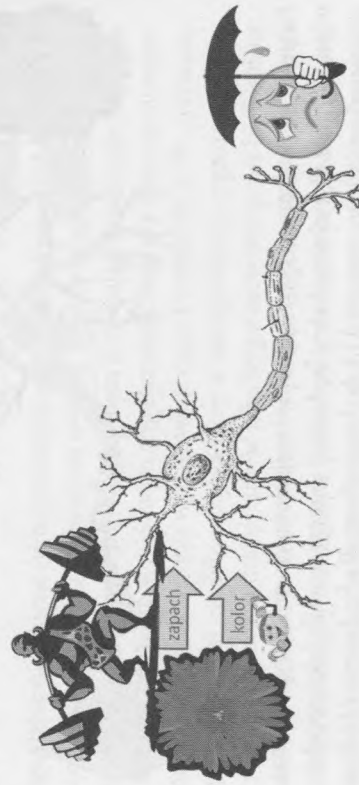
wprowadzającego sygnał o kolorze przypiszemy wagę o wysokiej wartości, a do wejścia niosącego sygnał o zapachu – wagę o niskiej wartości, to sygnał koloru przeważy nad sygnałem zapachu i neuron na wyjściu wyśle swój własny sygnał wyrażający aprobatę ($y = 1$), co przedstawiono na rysunku 20.

Rys. 20. Układ wag neuronu zapewniający jego pozytywną reakcję na sygnały pochodzące od badanego obiektu. Abstrakcyjne wagi przedstawiono odpowiednio w postaci siłacza dźwigającego sztangę o dużej wadze i dziecka bawiącego się handkami o małej wadze



Odwrotny układ wag, przedstawiony na rysunku 21, powoduje odwrotną reakcję neuronu.

Rys. 21. Odwrotne ustawienie wag daje odwrotny efekt



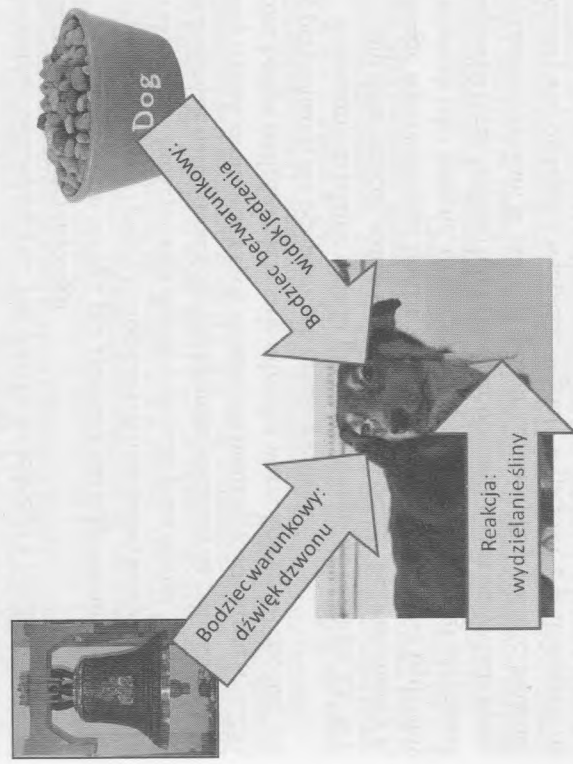
Ze sztucznych neuronów takiego typu jak przedstawiony na rysunku 18 można budować struktury służące określonym zastosowaniom praktycznym (m.in. na bazie takich neuronów funkcjonują perceptrony – uczące

się sieci neuronowe przeznaczone do klasyfikacji sygnałów i ich rozpoznawania). Natomiast my, zgodnie z tytułem tego rozdziału, skupimy się na możliwości wykorzystania tych sieci jako modeli struktur występujących w rzeczywistym układzie nerwowym.

Najprostsz przykład: sieć modelująca zjawisko odruchu warunkowego

Pierwsza z rozważanych tu sieci ma całkowicie sztuczną strukturę, wymyśloną przez autora tego rozdziału w tym celu, aby uzyskany system neurocybernetyczny naśladował jedno z najbardziej znanych zjawisk związanych z układem nerwowym – odruch warunkowy (Tadeusiewicz 2009). Za jego odkrycie rosyjski uczone Iwan Pawłow otrzymał w 1904 roku Nagrodę Nobla. Schemat doświadczenia Pawłowa przedstawiono na rysunku 22.

Rys. 22. Schemat odruchu warunkowego Pawłowa



W doświadczeniu tym mamy do czynienia z bodźcem bezwarunkowym, Bodziec to znaczy takim, który bez żadnych dodatkowych uwarunkowań wywołuje interesującą nas reakcję. Ponieważ obiektem badań Pawłowa był i bezwarunkowy

pies, a obserwowaną reakcją było wydzielanie śliny (wyjście ślimianki wprowadzono do próbówki, aby można było obserwować, kiedy pies wydziela ślinę i ile), naturalnym bodźcem bezwarunkowym był zatem widok i zapach podawanego psu jedzenia. Pawłow zaobserwował, że tę samą reakcję można uzyskać za pomocą bodźca zupełnie niezwiązanego z pożywieniem, mianowicie dźwięku dzwonów cerkiewnych. Dźwięk ten został przez Pawłowa nazwany bodźcem warunkowym, ponieważ reakcja na ten bodziec pojawiała się pod warunkiem, że pies był wcześniej w określony sposób traktowany.

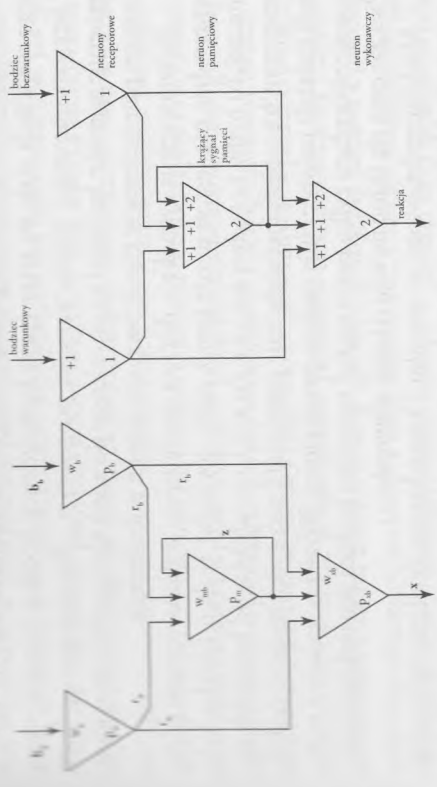
Skojarzanie bodźców

Warunkiem wytworzenia odruchu warunkowego było uprzednie wielokrotne skojarzanie bodźca warunkowego z bodźcem bezwarunkowym. W przypadku psa Pawłowa do takiego skojarzania doszło, ponieważ labrador, który karmił psy, nie postugiwał się zegarkiem (był to podobno niepiśmienny chłop, który nie potrafił odczytać godziny na zegarku), tylko kierował się tym, że o określonych godzinach w niezliczonych petersburskich cerkwiach biły dzwony – i karmił psy, kiedy je usłyszał. Ale psy też słyszały dzwony, przeto skojarzyły ich dźwięk z zapowiedzią karmienia. Po jakimś czasie wystarczył sam dźwięk dzwonu do tego, żeby pies zaczął wydzielać ślinę – co samo w sobie wydawało się efektem absolutnie zadziwiającym. Tak doszło do jednego z fundamentalnych odkryć współczesnej fizjologii.

W rzeczywistym mózgu psa zjawisko odruchu warunkowego powstaje przy udziale tysięcy biologicznych neuronów, których złożoność strukturalna i funkcjonalna daleko wykracza poza obszar, który potrafi niedoskonale naśladować sztuczny neuron, pokazany na rysunku 18. Mózg jednak zbudować z tych sztucznych neuronów strukturę (rys. 23), która będzie naśladowała zasadę odruchu warunkowego – chociaż, jak się przekonamy, nie bez pewnych mankamentów.

Rozważmy działanie układu przedstawionego na rysunku 23. Wiadąc tam dwa wejścia: jedno wprowadzające bodziec bezwarunkowy b_b i drugie wprowadzające bodziec warunkowy b_w . Oba bodźce będziemy traktowali jako binarne, to znaczy każdy z nich będzie mógł przybierać wartości 0 (brak bodźca) lub 1 (bodziec obecny), co jest oczywiście kolejnym uproszczeniem. W taki sam sposób (to znaczy jako wartości binarne) rozpatrywać będziemy wszystkie sygnały wyjściowe z widocznych na rysunku neuronów.

Rys. 23. Prosta sieć realizująca ideę odruchu warunkowego. Po lewej stronie schemat ogólny, po prawej z konkretnymi wartościami parametrów. Wpisano tylko oznaczenia przywoływane w tekście



Przedśledźmy, jak w tym modelu dochodzi do wyzwolenia reakcji przez bodziec bezwarunkowy b_b . Sygnał zewnętrzny $b_b = +1$ trafia do neuronu receptorowego, przy czym waga w_b wejścia tego neuronu wynosi $+1$, a jego próg p_b ma także wartość $+1$. W wyniku spełnienia warunku $b_b \cdot w_b \geq p_b$ na wyjściu neuronu receptorowego bodziec bezwarunkowego pojawia się sygnał $r_b = +1$. Sygnał ten dociera do dwóch neuronów: pamięciowego oraz wykonawczego. W neuronie pamięciowym (oznaczanym m od słowa *memory*) sam sygnał r_b nie jest w stanie wywołać żadnej reakcji, ponieważ trafia on na wejście z wagą $w_{mb} = +1$, ale próg neuronu pamięciowego wynosi $p_m = +2$, a zatem $r_b \cdot w_{mb} < p_m$ i reakcja neuronu pamięciowego, oznaczona z , wynosi 0.

Inaczej rzecz się ma z neuronem wykonawczym. Tutaj odpowiednia waga, do której dociera sygnał r_b , wynosi $w_{sb} = +2$, a próg $p_x = +2$, w następstwie czego spełniony jest warunek $r_b \cdot w_{sb} \geq p_x$ i pojawia się niezerowy sygnał neuronu wykonawczego, stanowiący wynik działania całej sieci, $x = +1$. Pies zobaczył mięso i się ślini. Normalnie!

Analogiczne rozumowanie pozwala stwierdzić, że pojawienie się bodźca warunkowego b_w (dźwięk dzwonów) zostaje w systemie odnotowane (odpowiedni neuron receptorowy zostaje skutecznie pobudzony i wysyła sygnał r_w), jednak nie wywołuje żadnej reakcji ani w neuronie wykonawczym, ani w neuronie pamięciowym.

Opisany efekt można obserwować dowolną ilość razy i nic się nie zmienia. Zmiana następuje – i to radykalna – w sytuacji skojarzania bodźca warunkowego z bodźcem bezwarunkowym. Gdy sygnały $b_b = +1$ oraz

Model reakcji na bodziec bezwarunkowy

Neuron pamięciowy i wykonawczy

Znaczenie skojarzania bodźców

$b_m = +1$ pojawią się równocześnie, wówczas oczywiście zadziała neuron wykonawczy (do tego wystarczył warunek $b_m = +1$), ale dodatkowo jego budzony zostanie neuron pamięciowy, bo suma pobudeń na dwóch jego wejściach staje się teraz (przy koincydencji bodźca warunkowego i bezwarunkowego) równa wartości progowej p_m . Neuron pamięciowy generuje sygnał z , który od tej pory krąży po zamkniętej pętli i utrzymuje neuron pamięciowy w stanie permanentnego pobudzenia.

Równocześnie sygnał ten dociera do neuronu wykonawczego, gdzie wprawdzie nie jest w stanie sam z siebie wywołać reakcji (gdyby było inaczej, to byłby to rodzaj *halucynacji niezależnej od bodźców zewnętrznych* i napędzanej wyłącznie pamięcią!), natomiast może to zrobić, jeśli zadziała jednocześnie neuron pamięciowy. Wówczas reakcja $x = +1$ zostaje wywołana tylko przez sygnał receptora bodźca warunkowego r_N . Pies zaczyna się ślinić na sam dźwięk dzwonów!

Niedokładność modelu

Model przedstawiony na rysunku 23 przedstawia działanie neuronów i prostych układów powstających z ich połączeń, nie jest to jednak dokładny model tworzenia się odruchu warunkowego w mózgu. Jednym z powodów, dla których musimy ten model odrzucić, jest to, że jako model uczenia jest on zbyt długi. Skojarzenie powstaje w nim już po jednorazowej koincydencji, a utworzony ślad pamięciowy jest niezniszczalny (chyba że jakiś czynnik całkowicie wyłączy na pewien czas aktywność sieci, co zdarza się w przypadku ciężkiego urazu lub omdlenia, które na biologiczny mózg działają tak, jak sygnał *reset* w systemach elektronicznych).

Przekleństwo perfekcji

Takie natychmiastowe i trwałe uczenie, które jest marzeniem studentów przygotowujących się do sesji, w rzeczywistości biologicznej byłoby jednak przekleństwem, bowiem mózg działający tak jak układ na rysunku 23 bardzo szybko uległby zapamięciu przez ślady pamięciowe całkiem przypadkowych i bezużytecznych koincydencji różnych zdarzeń, a w dodatku brakowałoby w nim mechanizmów usuwających starsze, a więc być może bezużyteczne już skojarzenia. Dlatego spowolniony (biologiczny) mechanizm pamięci trwałej, występujący w rzeczywistych neuronach, jest korzystniejszy z punktu widzenia dostosowania zachowania organizmu do zmiennych warunków środowiska. Niemniej niektóre mechanizmy pamięci krótkotrwałej (ulotnej) mogą działać w mózgu zgodnie ze schematem dynamicznego obiegu impulsów w niektórych zamkniętych obwodach; prosty schemat pokazany na rysunku 23 jest więc pożyteczny, bo pozwala na pewną orientację w procesach zachodzących w naszym układzie nerwowym.

Sieć najprostszej percepcji – hamowanie oboczne

Sieć omawiana w poprzednim podrozdziale miała strukturę zaprojektowaną arbitralnie. Po prostu autor wymyślił, ile ma być neuronów, jak je połączyć, jak dobrać parametry. Kierował się w tym analizą zachowania układu nerwowego w pewnych konkretnych warunkach. Taki sposób tworzenia neurocybernetycznych modeli jest jednak skrajnie nieracjonalny. Nie po to z ogromnym trudem (i z coraz liczniejszymi sukcesami) usiłujemy odszyfrowywać schematy połączeń elementów w mózgu, żeby namodzielnie wymyślać struktury neuronowe! Co więcej, mamy świadomość, że naturalne struktury neuronowe powstały w wyniku trwającej miliony lat ewolucji, w następstwie procesów doskonalenia się i dostosowywania, którym trudno byłoby sprostać, gdybyśmy zamiast z pokorą uczyli się od natury, zechcieli z nią konkurować. Dlatego rozsądną drogą tworzenia neurocybernetycznych modeli w postaci sieci neuronowych jest próba odwzorowania w modelu sieci rzeczywistej struktury, poznanej dzięki badaniom biologicznym.

Nie jest to jednak sprawa prosta, gdyż różnorodność potencjalnie możliwych połączeń neuronów wchodzących w skład sieci rośnie niesłychanie szybko wraz ze zwiększającą się ich liczbą. Liczbę możliwych struktur (w przypadku ogólnym, gdy nic nie wiemy o przeznaczeniu sieci ani o występujących w niej prawidłowościach), oznaczaną N dla sieci złożonej z K neuronów, można określić za pomocą matematycznego twierdzenia Polya. Okazuje się, że ta liczba rośnie niesłychanie szybko ze wzrostem K (np. dla $K = 20$ otrzymuje się oszacowanie $N = 10^{121}$, co jest wartością szokująco wielką).

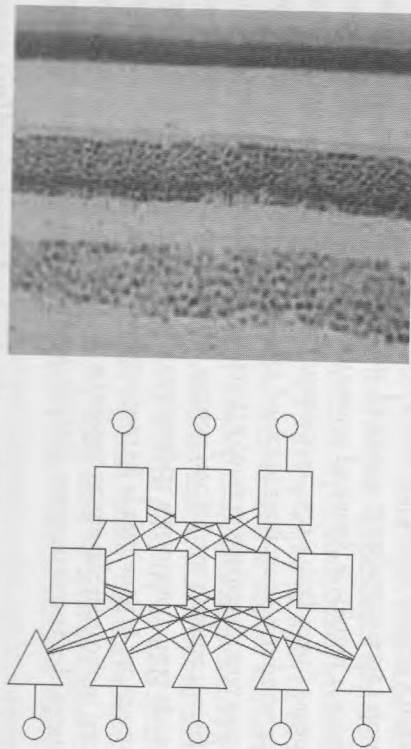
Wybór konkretnej sieci z tej ogromnej liczby możliwości odbywa się w ten sposób, że trzeba podać dla każdej połączonej pary neuronów indeksy (według wcześniej wprowadzonej numeracji) wskazujące, od którego neuronu do którego biegnie połączenie i jaka jest wartość współczynnika wagowego opisującego oddziaływanie synaptyczne na tej właśnie drodze. Przy tworzeniu sieci dynamicznych dochodzą tu jeszcze informacje o szybkości propagacji impulsów w danym połączeniu i o dynamice procesów pre- oraz postsynaptycznych. Tak więc z wyborem struktury i parametrów sieci neuronowej wiąże się ogromna i niesłychanie żmudna praca, której nie można zautomatyzować w taki sposób, by bez wysiłku znajdować sieć dla dowolnego zastosowania.

Jak nie projektować sieci neuronowych

Liczba możliwych połączeń w sieci neuronowej

Twierdzenie Polya

Rys. 24. Budowa typowej sztucznej sieci neuronowej i przekroj kory mózgowej



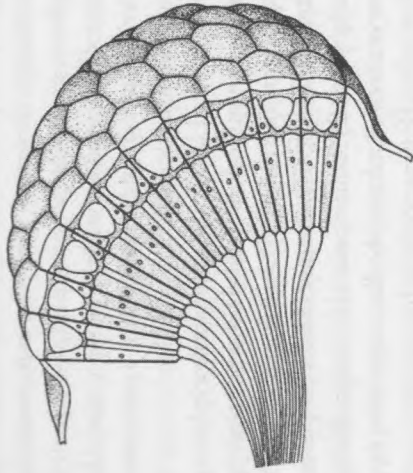
Struktura sztucznej sieci neuronowej

Zasada maksymalnej prostoty (w miejsce biologicznej wierności) przyjęta jest jednak twórcom sztucznych sieci neuronowych także w momencie wyboru struktury przeznaczonych do zastosowania w praktyce (np. do poszukiwania rozwiązań problemów technicznych czy ekonomicznych). W użytkowych sieciach neuronowych przyjmuje się arbitralnie, że sztuczna sieć składa się z neuronów ułożonych w warstwach. Okazuje się jednak, że taka arbitralnie wybrana struktura nie jest całkiem sprzeczna z rzeczywistością biologiczną, gdyż w systemach percepcyjnych (w tym także w ludzkim systemie wzrokowym) można zaobserwować neurony ułożone w warstwy (rys. 24).

Modelowanie percepcji prymitywnych zwierząt

Do modelowania za pomocą sztucznej sieci neuronowej ludzkiej kory mózgowej jeszcze nam daleko, ale można zaryzykować modelowanie w ten sposób systemów percepcyjnych prymitywnych zwierząt. U tych zwierząt (zwłaszcza u stawonogów) układ nerwowy jest znacznie prostszy od mózgu człowieka. Uproszczenia polegają na tym, że rozległe fragmenty rzeczywistego układu nerwowego tych zwierząt wykazują jednorodną strukturę, a to znaczy, że pewien schemat połączeń elementów, ustalony w jednym miejscu układu, jest zachowany (na zasadzie wiernych kopii lub z niewielkimi zmianami) w kolejnych fragmentach. Takie struktury można modelować za pomocą sztucznych sieci neuronowych w miarę prosto, a jednocześnie wyniki symulacji dostarczają odpowiedzi na istotne pytania.

Rys. 25. Schemat budowy oka stawonoga (np. kraba) (źródło: <http://laceylibertarian.us/wp-images/insectEye.jpg>, dostęp: 02.2013)



Jedną ze struktur neuronalnych (naturalnych i sztucznych) spełniających opisane wyżej warunki jest sieć neuronowa wykorzystująca zasadę tzw. hamowania obocznego, wykrytego przez H. Keffera Hartline'a i Floyd Ratliffa podczas badań układu wzrokowego skrzyplacza (*Limulus*).

Zasada hamowania obocznego

Otóż oczy tego kraba, podobnie jak oczy owadów, mają złożoną budowę, co oznacza, że poszczególne fotoreceptory, związane z oddzielnymi włóknami nerwowymi przesyłającymi wrażenia wzrokowe do mózgu zwierzęcia, znajdują się w oddzielnych fasetkach, z których każda ma własny układ optyczny: chitynową soczewkę i tubę o nieprzenikliwych dla światła ściankach (rys. 25). Umożliwia to oddzielne i niezależne oświetlenie każdego receptora, co w połączeniu z możliwością odbioru i analizy sygnałów z nerwu wzrokowego zwierzęcia daje nam unikatową możliwość badania sposobu reprezentacji bodźców wizualnych w sygnale przesyłanym do mózgu.

Badania przeprowadzone według zasygnalizowanego tu (oczywiście uproszczonego) schematu przyniosły dosyć zaskakujący wynik.

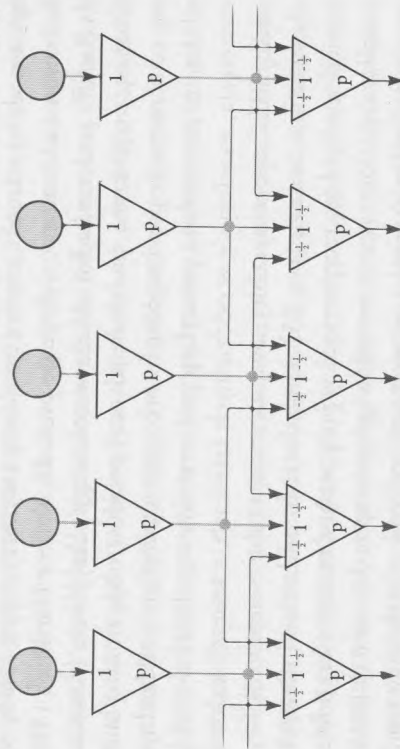
Stosunkowo łatwo odnaleziono dla każdej fasetki włókno nerwowe, które przekazywało do mózgu sygnał neuronowy o stanie receptora w danej fasetce. Gdy receptor był nieoświetlony, nie obserwowano także aktywności elektrycznej w związanym z nim włóknie nerwowym. Gdy światło się pojawiło, włókno przesyłało do mózgu całe serie impulsów o tym większej częstotliwości (co jest neuronowym sposobem kodowania wartości sygnału), im silniejsze było światło padające na fasetkę. Takiej właśnie zależności oczekiwano – oko sygnalizowało mózgowi, co widzi.

Natomiast pewne zaskoczenie wywołało odkrycie, że po oświetleniu także fasetek sąsiadujących z badaną sygnał pochodzący z teje fasetki silnie słabł, tak jakby receptory znajdujące się obok dławily (hamowały) aktywność rozważanego receptora. Zjawisko to nazwano hamowaniem obocznym i ustalono, że za jego powstanie odpowiedzialna jest bardzo prosta i – co ważne dla naszych rozważań – praktycznie jednorodna sieć neuronowa, pośrednicząca między receptorami a nerwem wzrokowym i mózgiem kraba (Tadeusiewicz 2009).

del hamowania
obocznego

Sieć, którą wykryto w układzie wzrokowym skrzypłocza, jest oczywiście siecią dwuwymiarową, podobnie do dwuwymiarowej powierzchni (zbliżonej do fragmentu kuli) pokrytej przez fasetki. Przez chwilę jednak zapomnijmy o tym szczególe i zbudujmy model (w postaci sieci neuronowej) realizujący zasadę hamowania obocznego dla hipotetycznej struktury jednowymiarowej, to znaczy dla receptorów światłoczułych ułożonych wzdłuż jednej linii na powierzchni fasetki (rys. 26).

Rys. 26. Jednowymiarowa sieć neuronowa modelująca zjawisko hamowania obocznego



Z każdym receptorem (przedstawionym na tym rysunku w postaci kółka) związany jest neuron, zamieniający jego sygnał (depolaryzację następującą pod wpływem światła) na typowy sygnał nerwowy. Sygnały w neuronowej części tego systemu są tym razem traktowane w sposób typowy dla sieci neuronowych, to znaczy jako liczby rzeczywiste z pewnego przedziału. Sygnały te są przekazywane do odpowiednich neuronów zwojowych, kierujących odpowiedź wprost do mózgu, przy czym obowiązuje następująca zasada (pokazana też na rys. 26): sygnały z receptora przyporządkowanego do danego neuronu docierają do niego poprzez wejście pobudzające z wagą $+1$; oznacza to, że im silniej pobudzony jest

receptor, tym silniejszy sygnał może wysłać do mózgu neuron zwojowy. Na sygnał tego ostatniego mają jednak wpływ receptory sąsiednie. Przy modelu pola percepcyjnego w postaci jednowymiarowego łańcucha receptorów sąsiednie są dwa receptory (umownie nazwane lewy i prawy). Sygnały z tych receptorów docierają do neuronu zwojowego z wagą $-0,5$. Opisany schemat powtarza się w identycznej postaci dla każdego receptora (i związanego z nim neuronu zwojowego).

Opiszemy teraz działanie takiej sieci neuronowej wykorzystującej zasadę hamowania obocznego.

Całkowite zbilansowanie w każdym neuronie opisanej wyżej sieci z hamowaniem obocznym sygnału pobudzającego pochodzącego od „własnego” receptora oraz od receptorów sąsiadujących (obocznych) powoduje, że taki „układ wzrokowy” (czujsłów wynika z umowności rozważania tego problemu na modelu jednowymiarowym) jest niewrażliwy na dowolne równomiernie rozłożone bodźce sensoryczne. Łatwo to zauważyć, biorąc pod uwagę sumaryczne pobudzenie neuronów zwojowych przy dowolnym sygnale pobudzenia X . Korzystając ze wzoru podanego na rysunku 18, możemy to zapisać następująco:

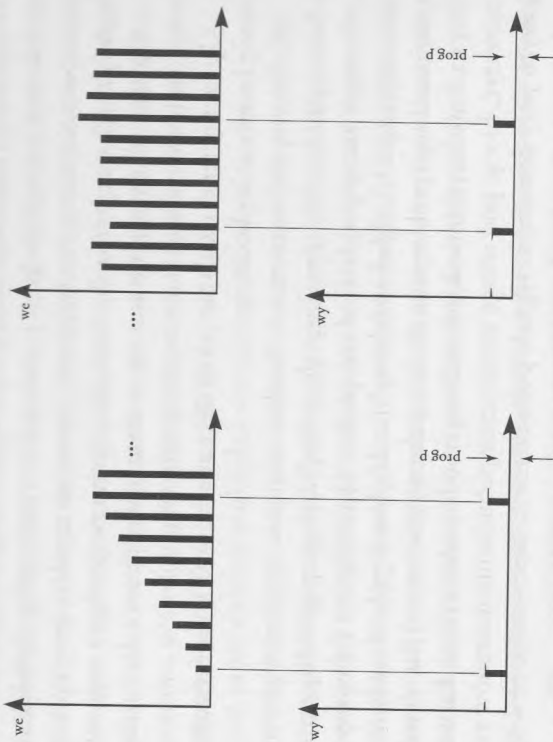
$$\text{Pobudzenie} = 1 X - 0,5 X - 0,5 X = 0.$$

A zatem niezależnie od tego, czy oczy kraba zwrócone są ku rozświetlonej słońcem powierzchni oceanu, czy ku jego mrocznej głębi – zwierzę nie widzi nic. Jest to bardzo dobry mechanizm przystosowawczy, który oszczędza ubogie zdolności przetwarzania informacji, jakimi dysponuje niewielki mózg tego zwierzęcia. W jednym i drugim przypadku bodźce wzrokowe nie niosą bowiem żadnych istotnych życiowo informacji.

Również w sytuacji przedstawionej na rysunku 27 (po lewej stronie), kiedy jasność obserwowanej sceny zmienia się w sposób płynny od wartości niższej (po lewej stronie) do wartości wyższej (po prawej), nasz model nie prześle do mózgu modelowanego zwierzęcia żadnej informacji. Dlaczego?

Rola hamowania
obocznego
w percepcji

Rys. 27. Sieć z hamowaniem obocznym jest niewrażliwa na wiele cech wejściowego sygnału



Czego nie widzi krab?

Popatrzmy uważnie na rysunek. Na górze widoczne są słupki reprezentujące wielkość pobudzenia kolejnych receptorów. Widać wyraźny gradient wzrostu oświetlenia. Na dole natomiast pokazana jest siła pobudzenia neuronów zwoju spiralnego, odpowiadających tym receptorom. Tam, gdzie oświetlenie równomiernie narasta, pobudzenie jest zerowe. Łatwo to wytłumaczyć. Jeśli receptor środkowy jest oświetlony z intensywnością X , to receptor po lewej będzie oświetlony z intensywnością $(X - \Delta)$, a receptor z prawej $(X + \Delta)$, gdzie Δ jest stałą, zależną od „stro-ności” (tempa) narastania oświetlenia. Ale łatwo policzyć, że:

$$\text{Pobudzenie} = 1 \cdot X - 0,5 (X - \Delta) - 0,5 (X + \Delta) = 0.$$

Coś innego dzieje się na początku „zbocza” i na jego końcu, bo tam występuje niewielkie pobudzenie wynikające z asymetrii. Pamiętajmy jednak, że neurony zwojowe mają próg pobudzenia p (patrz rys. 18). Wystarczy więc, by spełniony był warunek

$$\text{Pobudzenie} = 1 \cdot X - 0,5 (X - \Delta) < p,$$

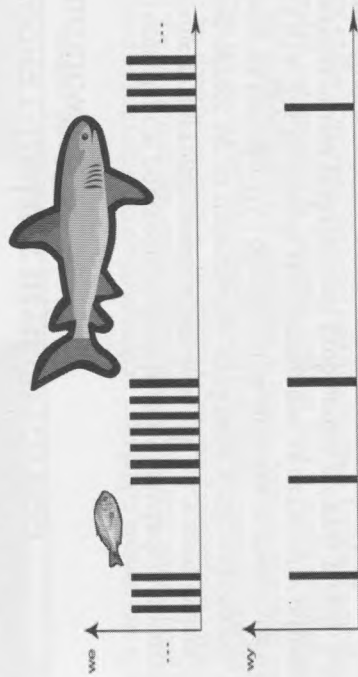
aby krab nie musiał „zawracać sobie głowy” początkowymi i końcowymi fragmentami obszarów, w których występuje gradient oświetlenia.

Również małe fluktuacje jasności w postaci pojedynczych, nieco jaśniejszych lub nieco ciemniejszych punktów „w polu widzenia” na

podobnej zasadzie nie będą przez mózg kraba rejestrowane, co ilustruje prawa strona rysunku 27.

Z naszych rozważań wynika, że sieć z hamowaniem obocznym (nawet w postaci uproszczonego jednowymiarowego modelu) stanowi filtr, który bardzo silnie ogranicza rozmiar strumienia informacji sensorycznych docierających do mózgu kraba. Wiemy już, czego to zwierzę nie widzi, chociaż odpowiednie sygnały świetlne docierają do jego oka. Krab jednak coś powinien widzieć, bo inaczej jego złożone oczy byłyby nieprzydatne. Otóż badania przeprowadzone za pomocą sieci neuronowej momentalnie ujawniają, jak wygląda świat jego wrażeń zmysłowych. Popatrzmy na rysunek 28.

Rys. 28. Wydzielanie krawędzi obiektów jako głównego źródła informacji w sieci z hamowaniem obocznym



Na rycinie tej widzimy tło (narysowane jako serie wysokich słupków reprezentujących rozświetloną toń morza) oraz dwa ciemne obiekty (obszary, w których wysokość słupków jest zerowa). Łatwo przewidzieć na podstawie rozważań z poprzedniego podrozdziału, że zarówno obszary tła, jak i obszary wnętrza ciemnych obiektów nie powodują wysyłania do mózgu kraba żadnych informacji. Natomiast w dolnej części rysunku widoczne są cztery bardzo silnie pobudzone neurony, których sygnały z pewnością dotrą do mózgu – są to neurony ulokowane dokładnie na krawędziach dostrzeżonych obiektów. Dzieje się tak dlatego, że neurony odpowiadające receptorom dostrzegającym krawędź mają niezbalansowane pobudzenie (hamowanie dociera tylko z jednej strony):

$$\text{Pobudzenie} = 1 \cdot X - 0,5 \cdot X > p.$$

Te neurony przechodzą zatem w stan aktywny i wysyłają sygnały.

Oszczędny aparat percepcyjny kraba

Zauważmy, jak mądrze jest to urządzone: do mózgu dociera tylko minimalna ilość informacji, ma on więc niewiele pracy. W tym lakonicznym komunikacie jest jednak wszystko: informacja, że pojawiły się obiekty, ile ich jest i w jakim kierunku od odbiorcy się znajdują. Co więcej, na podstawie odstepu pomiędzy włóknami neuronów sygnalizujących krawędzie obiektów można oszacować rozmiar obiektu. Obiekt mały (po lewej stronie rysunku) może być zdobyczą, czyli trzeba próbować go złapać i zjeść. Obiekt duży (po prawej) to potencjalny drapieżnik, więc należy od niego uciekać.

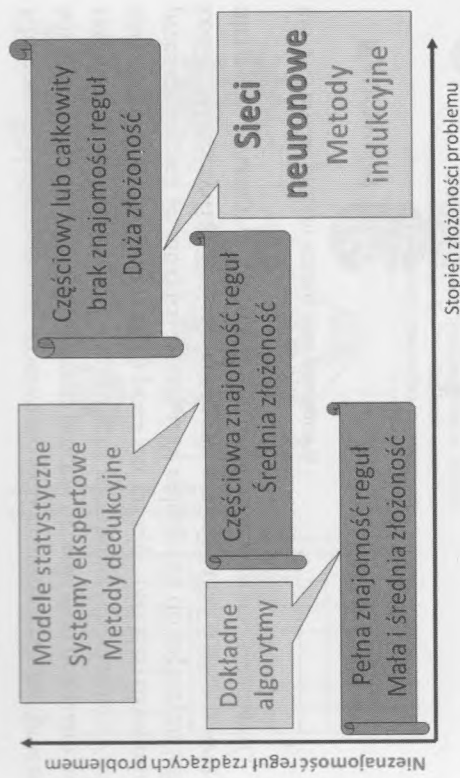
Tak prosty mechanizm, a jak wiele można na jego podstawie ustalić przy minimalnym zaangażowaniu „intelektu” kraba!

Budowa i działanie użytkowych sieci neuronowych

Sieci omawiane w poprzednich dwóch podrozdziałach miały strukturę i parametry zaprojektowane arbitralnie. Jak już wspominałem, lepiej jednak oprzeć się na uniwersalnej strukturze sieci neuronowej, przywołanej na rysunku 10, i określić jej zachowanie w trakcie procesu uczenia. Tak budowane sieci neuronowe są używane w informatyce do rozwiązywania praktycznych problemów, głównie takich, których nie potrafimy rozwiązać innymi metodami, ponieważ brak nam wiedzy o tym, jakimi regułami rządzi się rozwiązywany problem. Z reguły są to także problemy o dużym stopniu złożoności. W takim przypadku zawodzą dokładne algorytmy stosowane przy rutynowych zastosowaniach informatyki. Nieskuteczne są też modele statystyczne, do których się odwołujemy, gdy dysponujemy jedynie częściową znajomością reguł rządzących rozwiązywanym problemem, a także oparte na sztucznej inteligencji tak zwane systemy ekspertowe, wykorzystujące metody dedukcyjne. Przy całkowitym braku znajomości reguł rządzących problemem (np. w zadaniach związanych z prognozowaniem przyszłości), właśnie sieci neuronowe z ich umiejętnością korzystania z metod indukcyjnych, opartych na procesie uczenia, są jedynymi w pełni skutecznymi narzędziami (rys. 29).

Uczące się sieci neuronowe

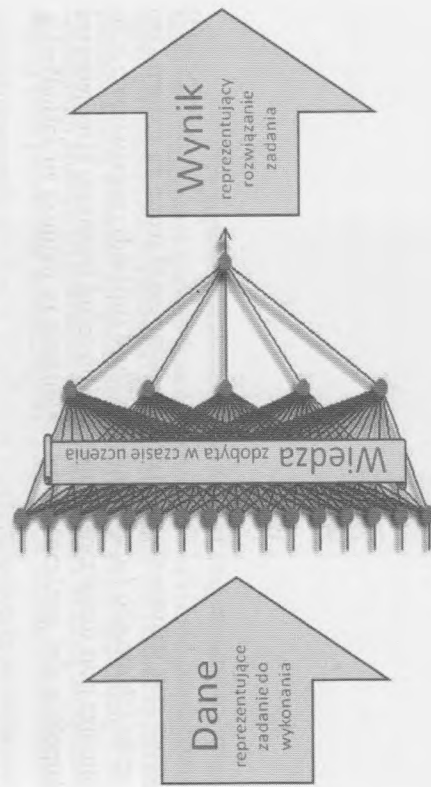
Rys. 29. Diagram ilustrujący, kiedy sieci neuronowe są lepsze od innych metod informatycznych



Kiedy i jak stosować sieci neuronowe

Sposób stosowania takiej sieci neuronowej do rozwiązywania praktycznych problemów, których cechy scharakteryzowane zostały wyżej, pokazany jest (w uproszczeniu) na rysunku 30. Więcej informacji na temat tych typowych zastosowań sieci neuronowej oraz na temat ich budowy i zasad działania znaleźć można w książkach przywoływanych we wstępie. Ponownie warto w tym miejscu wskazać, że prawdziwie użyteczną i szczegółową wiedzę o sieciach oraz o ich właściwościach warto zdobywać i utrzymywać poprzez wykorzystanie darmowych i łatwych w użyciu programów, ściąganych ze strony internetowej pokazanej na rysunku 1.

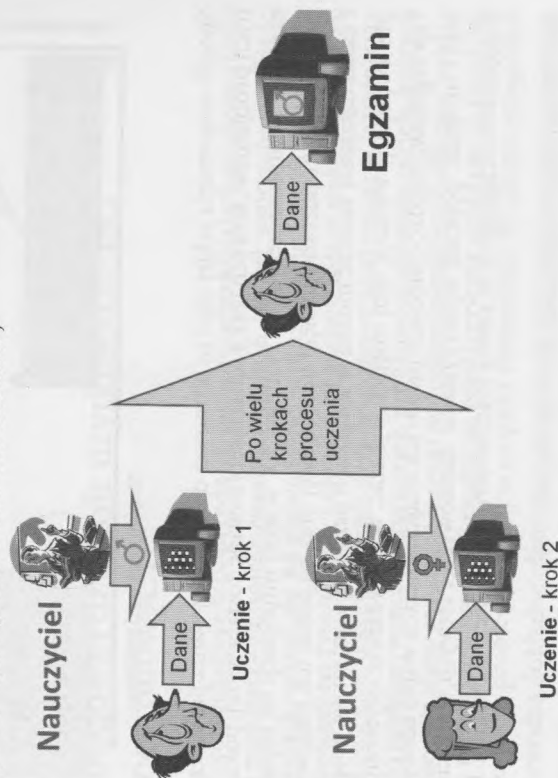
Rys. 30. Ilustracja typowego działania sieci neuronowej



Proces uczenia

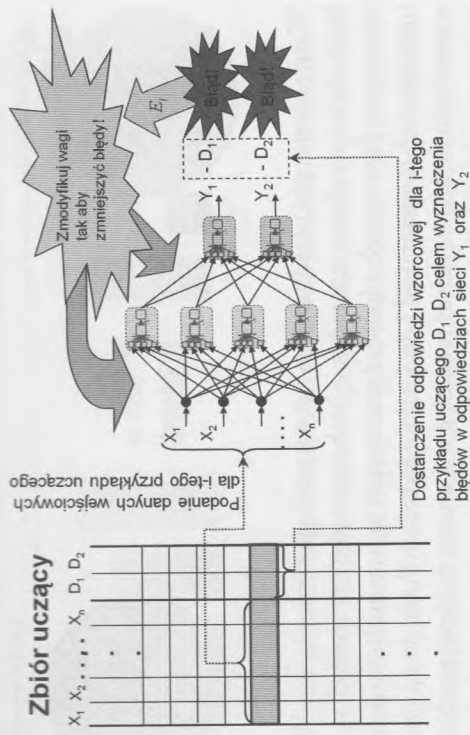
Kluczowym elementem przy stosowaniu sieci neuronowych jest proces uczenia. Schemat przykładowego procesu uczenia sieci pokazano na rysunku 31, przy czym chodzi tu o trochę nietypowy (choć dość łatwy do narysowania) przykład nabywania przez sieć umiejętności klasyfikowania przedstawianych jej ludzkich twarzy jako należących do mężczyzn lub kobiet. W rzeczywistości sieci neuronowe uczą się zadań o wiele poważniejszych i związanych z licznymi ważnymi praktycznymi zadaniami.

Rys. 31. Przykład procesu uczenia sieci neuronowej



Występujący na rysunku 31 instruujący sieć „nauczyciel” jest wygodną personalizacją formalnej procedury, w której nauczaniem sieci zajmuje się w rzeczywistości specjalny program komputerowy odwołujący się do bazy danych na temat prawidłowych rozwiązań stawianych sieci zadań (rys. 32).

Rys. 32. Nieco dokładniejszy schemat uczenia sieci neuronowej



Baza danych z rozwiązanymi zadaniami nazywa się zbiorem uczącym, a proces uczenia polega na takim zmienianiu wag w sieci neuronowej, by minimalizować popełniany przez nią błąd.

Proces uczenia zachodzący w sieci ma zwykle na celu wyuczenie jej jakichś praktycznych umiejętności, my jednak przyjrzymy się temu, jak bardzo proces uczenia sieci neuronowej przypomina sposób uczenia człowieka, dla którego może być w efekcie użytecznym modelem. Opisane niżej doświadczenia były realizowane przez autora i jego doktorantów, jednak nawiązują one do wiadomości zawartych w książce *Jak uczy się mózg* (Spitzer 2007).

Powiedzmy, że sieć ma się nauczyć czterech różnych umiejętności (mniej w tym momencie o szczegóły), a stopień opanowania tych umiejętności można wyznaczyć, podliczając w nauczonej sieci liczbę neuronów, które są zaangażowane w pamiętanie tej właśnie umiejętności. Obejrzyjmy kilka wyników uzyskanych za pomocą tego prostego modelu i zauważmy, że zadziwiająco przypominają one obserwacje, jakie mogliśmy poczynić, obserwując procesy uczenia zachodzące w naszym własnym umyśle. Przyjrzyjmy się obrazowi na rysunku 33.

Uczenie sieci neuronowych a uczenie człowieka