

MODELOWANIE RZECZYWISTOŚCI

Daniel Wójcik
Instytut Biologii Doświadczalnej PAN

d.wojcik@nencki.gov.pl

tel. 5892 424

<http://www.neuroinf.pl/Members/danek/swps/>

Podręcznik

Iwo Białynicki-Birula
Iwona Białynicka-Birula

ISBN: 83-7255-103-0
Data wydania: 6 maja 2002

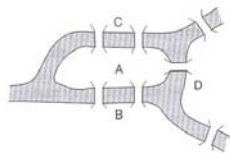


Grafy i sieci

- Mosty Królewca: własności topologiczne
- Problem komiwojażera: własności metryczne
- Języki naturalne: opis obiektów nieskończonych
- Małe światy i inne sieci

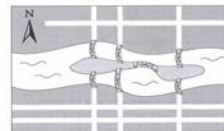
Mosty Królewca

- Czy można przejść Królewiec przechodząc każdy most dokładnie raz?



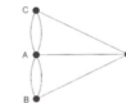
Mosty Królewca

- Czy można usunąć jeden most tak by dało się przejść Królewiec przechodząc każdy most dokładnie raz?
- Czy z każdego miejsca można wtedy wykonać pełen spacer?



Graf reprezentujący mosty Królewca

- Potrzebujemy informacji o połączeniach poszczególnych obszarów łądu
- Własności obiektu, które nie zmieniają się przy jego odkształceniach (bez rozrywania) to **własności topologiczne**
- Mosty Królewca można reprezentować przy pomocy **grafu**
- **Cykl Eulera**

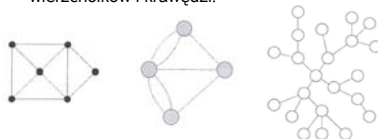


Grafy eulerowskie i półeulerowskie



Graf

- Punkty oznaczające obszary lądu to **wierzchołki grafu**
- Połączenia między obszarami (mosty) to **krawędzie grafu**
- Minimalny model grafu składa się z wierzchołków i krawędzi.

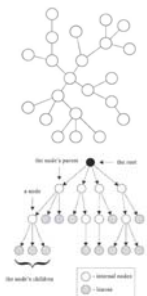


Inne typy grafów

- Minimalny model grafu można rozbudować:
 - Przypisując każdemu z wierzchołków lub krawędzi etykietę (napis) lub wagę (liczbę rzeczywistą)
 - Ustalając kierunek krawędzi (np. ulice jednokierunkowe) – graf skierowany
 - Drzewa to spójne (= niepodzielne bez przecięcia krawędzi) grafy nie zawierające cykli (= zamkniętych dróg)

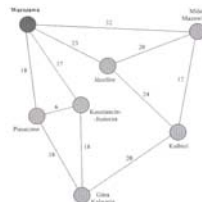
Przykład drzewa

Często wyróżniamy jeden z węzłów drzewa nazywany **korzeniem**. Wtedy każdy węzeł w drzewie ma dokładnie jednego sąsiada, który jest bliżej korzenia od niego. Ten sąsiad nazywa się jego **rodzicem**, pozostali sąsiedzi to **dzieci**. Węzeł bezdzietny nazywa się **liściem**.



Problem komiwojażera

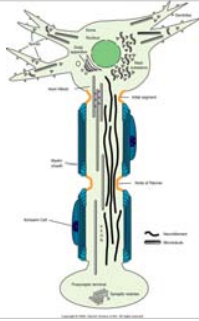
- Która droga jest najlepsza (najkrótsza)?
- minimalny **cykl Hamiltona**



Problem komiwojażera

- Nie istnieje efektywny algorytm rozwiązujący bezbłędnie problem komiwojażera dla dużych grafów w krótkim czasie
- **Problem NP-pełny**
- Istnieją za to efektywne algorytmy heurystyczne, które szybko przybliżają najkrótszą drogę

Przekrój neuronu

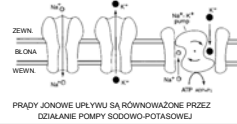


Jak działa neuron?

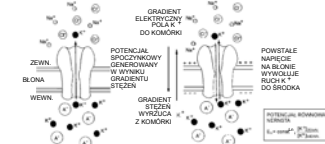
Wnętrze niepobudzonego neuronu ma potencjał ok. -70 mV względem środowiska międzykomórkowego. Jest to wynikiem dwóch mechanizmów:

- Aktywnego transportu biochemicznego dodatnich jonów sodu na zewnątrz komórki
- Dyfuzji tych jonów z powrotem na skutek gradientu elektrycznego

Oba procesy równoważą się właśnie przy napięciu spoczynkowym około -70mV.



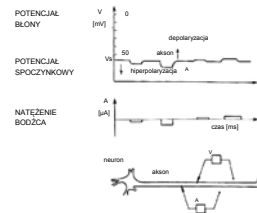
BŁONA PRZEPUSZCZALNA TYLKO DLA K⁺



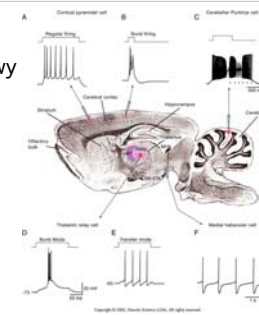
Potencjał czynnościowy (iglicowy)

Wstrzykując do komórki jony *ujemne* zwiększamy napięcie na błonie (**hiperpolaryzacja**). Wstrzykując jony *dodatnie* zmniejszamy napięcie na błonie (**depolaryzacja**). Z chwilą gdy napięcie na błonie przekroczy pewną wartość progową depolaryzacji, jego dalsza zmiana przestaje być proporcjonalna do bodźca – otwierają się kanałki jonowe dla sodu, który błyskawicznie wpływa do wnętrza komórki powodując krótkotrwałą (ok. 1 ms) zmianę polaryzacji błony, nazywaną **potencjałem czynnościowym** lub **iglicowym**. Kiedy na skutek wpływu jonów sodu polaryzacja błony zmieni znak, otwierają się kanałki jonowe potasu, który wypływa z komórki neutralizując zmiany napięcia i przywracając jego początkowy poziom.

Generacja potencjału czynnościowego



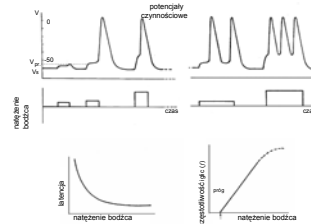
Potencjał czynnościowy



Kodowanie bodźca

Pobudzenie komórki coraz silniejszymi prądami depolaryzującymi powoduje skracanie czasu potrzebnego do powstania kolejnej iglicy, nie zmieniając jej amplitudy. Zatem *siła bodźca jest kodowana częstotliwością iglic generowanych na błonie*.

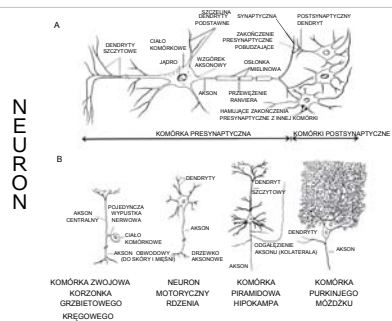
Kodowanie bodźca częstotliwością



Transport informacji

Potencjał czynnościowy w warunkach fizjologicznych powstaje przy wzdórku aksonowym, gdzie jest najniższy próg pobudzenia. Stąd, iglice propagują się wzdłuż aksonu z zachowaniem amplitudy i częstości do jego zakończeń.

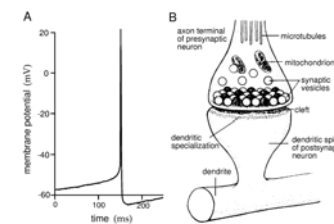
Akson ma często odgałęzienia boczne, a na każdym końcu tzw. drzewko aksonowe, dzięki którym może przekazywać pobudzenie kilku komórkom postsynaptycznym.



Synapsa

Zakończenie aksonu (**kolbka synaptyczna**) zawiera zapas **neurotransmitera**, który wyzwala się w momencie depolaryzacji kolbki aksonu przez potencjał czynnościowy. Neurotransmitter wychodzi do szczeliny międzysynaptycznej, dyfunduje w kierunku błony neuronu postsynaptycznego, łączy się z nią, przy czym zmienia na pewien czas jej przepuszczalność i polaryzację.

Synapsa

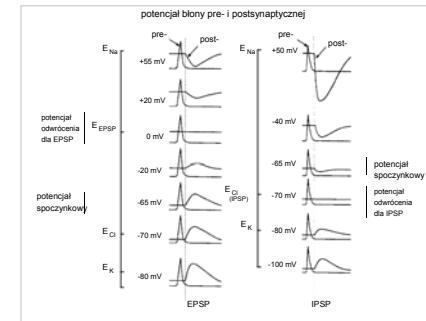


Dendryty

W zależności od rodzaju neurotransmitera napięcie na błonie dendrytu maleje albo rośnie. Ponieważ potencjał progowy pobudzenia błony dendrytu jest znacznie wyższy niż w aksonie, depolaryzacja błony nie wywołuje iglicy ale propaguje się w stronę ciała komórkowego ze stratami (amplituda maleje). Potencjał ten zanika w czasie od kilku do kilkunastu sekund. Jeżeli w tym czasie nadejdzie kolejny impuls od tej samej lub innej komórki, sumują się one.

Wolny potencjał

Analogową zmianę potencjału postsynaptycznego błony ciała i dendrytów neuronu w wyniku docierających kolejnych porcji neurotransmitera nazywamy **wolnym potencjałem**. Jest on *proporcjonalny do częstotliwości iglic w zakończeniu aksonu*. Podobnie sumują się hamulcowe potencjały postsynaptyczne, tylko *większa aktywność* komórki hamulcowej presynaptycznej *zmniejsza wolny potencjał* dendrytu postsynaptycznego, zatem komórki hamujące w pewnym sensie zmieniają znak informacji. Gdy suma dyfundujących do wzdórka aksonowego prądów jonowych jest dodatnia, błona ulega depolaryzacji, a gdy potencjał przekroczy wartość progową, może zapoczątkować aktywację potencjałów czynnościowych.



Waga synaptyczna

Każde połączenie synaptyczne między komórkami wytwarza potencjały postsynaptyczne o charakterystycznej wielkości. Ich amplituda zależy od aktualnej liczby kontaktów synaptycznych drzewka aksonowego z komórką postsynaptyczną oraz od procesów plastycznych, które mogły zmienić efektywność tych kontaktów w inny sposób.

Udział danego potencjału postsynaptycznego w wytworzeniu potencjału czynnościowego neuronu nazywa się wagą synaptyczną. Dla charakterystyki sieci w danej chwili zakłada się, że jest ona stała, w rzeczywistości zmienia się ona z czasem. Zmiana wag synaptycznych prowadzi do trwałej reorganizacji struktury sieci neuronowej i leży u podłoża mechanizmów uczenia się i pamięci.

Podsumowanie neuronu

Przeszliśmy tak przez cały transfer informacji od potencjałów czynnościowych generowanych w jednym neuronie, przez synapsę, do następnego neuronu. Widzimy, że w pierwszym przybliżeniu można przyjąć, że sieć nerwowa działa według prostej zasady sumowania wpływów pobudzających i hamujących, a wynikiem tej operacji jest wolny potencjał neuronu. Ten podstawowy model transportu informacji jest używany w teorii sztucznych sieci neuronowych, o których będzie mowa dalej.

Sztuczne sieci neuronowe

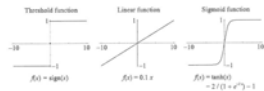
Sieci neuronowe to modele obliczeniowe, których funkcjonalność i struktura wzorowana jest na mózgu.

Charakteryzuje je:

- Rozproszone i równoległe przetwarzanie informacji
- Odporność na uszkodzenia części elementów
- Możliwość uczenia

Sztuczny neuron

Podstawowym elementem sztucznych sieci neuronowych jest najprostszy model neuronu, którego stan odpowiada częstotliwości potencjałów czynnościowych komórki. Sieci neuronowe składają się z połączonych sztucznych neuronów. Połączenia odpowiadają synapsom, z każdym jest skojarzona pewna waga, odpowiednik wagi synaptycznej. Kiedy sygnał przechodzi przez połączenie, mnożony jest przez wagę. Sztuczny neuron na wyjściu podaje ważoną sumę przychodzących sygnałów, którą czasem dodatkowo się moduluje (**funkcją aktywacji**).



Sieci neuronowe

Sieci neuronowe jednokierunkowe (**feedforward neural networks**) są budowane warstwowo. Składają się z warstwy wejściowej, która odpowiada neuronom sensorycznym, i wyjściowej, która odpowiada neuronom motorycznym. Często także z kilku tzw. warstw ukrytych, które odpowiadają interneuronom układu nerwowego.

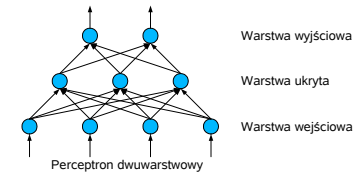
Kiedy chcemy rozwiązać dany problem przy pomocy sieci neuronowej napotykamy dwa główne problemy:

- Jak zaprojektować sieć (znaleźć graf skierowany reprezentujący sieć)
- Jak nauczyć sieć (dobrać wagi by sieć wykonywała żądane zadanie)

Nie wiemy jak natura rozwiązuje te problemy (ewolucja), nie mamy też ogólnej metody konstrukcji sztucznych sieci neuronowych.

Perceptrony

- Najprostszą i najpopularniejszą siecią neuronową jest **perceptron wielowarstwowy**. Jest siecią jednokierunkową (**feedforward**), to znaczy nie zawierającą cykli. Informacja podana jej na wejściu przechodzi taką sieć raz i stabilizuje się na wyjściu.



Jak działa perceptron?

- Ustaw stan S_j^k każdego węzła N_j^k warstwy wejściowej na I_j – i -tą składową sygnału wejściowego
- W każdej warstwie k , od drugiej do wyjściowej (warstwa nr M):
 - W każdym węźle N_j^k w warstwie k policz ważoną sumę wchodzących sygnałów, zastosuj funkcję aktywacji f , wynik przypisz stanowi S_j^k tego węzła. Inaczej mówiąc

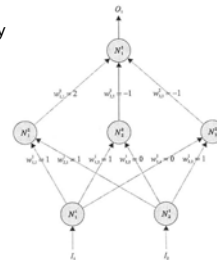
$$S_j^k = f\left(\sum_{i=1}^{L_{k-1}} w_{ji}^{k-1} S_i^{k-1}\right)$$

gdzie L_{k-1} to liczba węzłów w warstwie $k-1$, w_{ji}^{k-1} jest wagą połączenia od węzła N_i^{k-1} do węzła N_j^k

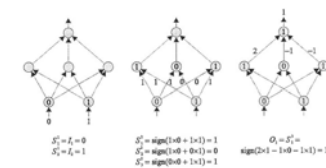
- Skopiuj stan S_j^k każdego węzła N_j^k na O_j – i -tą składową sygnału wyjściowego.

Przykład: reguła XOR

- W roku 1969 M. Minsky i S. Papert udowodnili, że żaden perceptron jednowarstwowy nie może wyliczyć funkcji XOR. Okazało się, że dwuwarstwowy perceptron potrafi to zrobić.



Perceptron XOR w akcji



Uczenie sieci

- Uczenie sieci polega na zmianie wag sieci tak, by funkcjonowała w pożądanym sposób. Na ogół mamy na myśli automatyczne dopasowywanie wag zgodnie z pewnym **algorytmem uczenia**.
- Wyróżniamy trzy typy uczenia sieci:
 - Supervised learning (uczenie z nauczycielem)
 - Reinforcement learning (uczenie z forsowaniem): oceniamy sieć na podstawie tego, jak sobie radzi z zadaniem
 - Unsupervised learning (uczenie bez nauczyciela, samouczenie): sieć nie dostaje żadnych wskazówek jak ma działać; na ogół takie sieci uczą się dzielić grupę sygnałów wejściowych na klasy sygnałów o „podobnym” charakterze.

Supervised learning

- Uczenie z nauczycielem polega na prezentacji sieci par wejście – pożądanie wyjście. Wagi są tak zmieniane by minimalizować funkcję błędu. Sposób minimalizacji i funkcja błędu zasadniczo określają algorytm uczenia. Do najpopularniejszych metod należą **perceptron learning rule** (PLR, metoda uczenia perceptronu), **delta rule** (reguła delty) oraz **backpropagation** (metoda wstecznej propagacji błędów). Najpotężniejsza z nich jest metoda ostatnia, ale opiszemy tu najprostszą, regułę uczenia perceptronu.

Reguła uczenia perceptronu

PLR korzysta z par wejście – pożądanie wyjście. Początkowo wszystkie wagi sieci wybieramy losowo, następnie iterujemy cykl operacji: W każdej iteracji podajemy wzór wejściowy $\{I_1, \dots, I_n\}$ ze zbioru uczącego. Sieć produkuje wyjście $\{O_1, \dots, O_n\}$, od którego odejmujemy pożądanie wyjście $\{D_1, \dots, D_n\}$ otrzymując wektor błędów $\{E_1, \dots, E_n\}$. Długość tego wektora jest miarą sukcesu sieci. Jeżeli wektor ten ma długość 0, to znaczy, że sieć wygenerowała żądany wynik dokładnie. Im dłuższy wektor błędów tym gorzej poradziła sobie sieć. Każdą wagę sieci zmieniamy w zależności od wartości składowej błędów, z którą jest związana. Zmodyfikowana waga przyjmuje wartość

$$w_{i,j} + \eta I_i E_j$$

gdzie η jest parametrem szybkości uczenia. Powinien on wynosić ok. 0.1. Jeżeli ta stała jest za duża, wówczas sieć szybko „zapomina” to, czego się nauczyła. Jeżeli jest za mały, to sieć uczy się bardzo wolno.

Metoda propagacji wstecznej

- Uogólnienie reguły uczenia perceptronu
- Zmieniamy wagi począwszy od warstwy wyjściowej cofając się do wejścia
- Zmiany wag zachodzą zgodnie z PLR

Sieci ze sprzężeniem zwrotnym

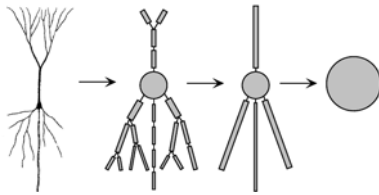
- Uogólnienie sieci jednokierunkowych.
- Sieć Elmana: pamięć poprzednie wejście



Modelowanie neuronu

- Firing-rate model (najprostszy)
- Integrate-and-fire model (proste równanie różniczkowe)
- Równania Hodgkina-Huxleya (układ równań różniczkowych)

Redukcja neuronu



Źródła rysunków i literatura

- „Theoretical neuroscience”, Dayan, Abbott
- „Fundamental neuroscience”, Squire et al.
- „Principles of Neural Science” Kandel et al.
- „Neuron i sieci neuronowe”, A. Wróbel, rozdział III w „Mózg a zachowanie”
- „Modelowanie rzeczywistości” IBB+IBB, rozdział 13
- „Wprowadzenie do teorii grafów” R. J. Wilson
- „Understanding nonlinear dynamics” Kaplan, Glass

Użyte programy

- Euler
- Hopfield

Źródła

- <http://www.wiw.pl/modelowanie/>
-