

Výpočtová neuroveda

Úvod do modelovania
neurofyziologických a behaviorálnych dát

NORBERT KOPČO

Technická univerzita v Košiciach

Autor: doc. Ing. Norbert Kopčo, PhD.

Názov: Výpočtová neuroveda

Podnázov: Úvod do modelovania neurofyziologických a behaviorálnych dát

Vydavateľ: Technická univerzita v Košiciach

Grafická úprava: Norbert Kopčo

Recenzenti: doc. RNDr. Gabriela Andrejková, CSc. a PhDr. Ing. Marek Dobeš, PhD.

Rok vydania: 2011

Miesto vydania: Košice

Vydanie: prvé

Počet strán (číslované): 66

Internetová adresa: <http://pcl.tuke.sk/vakn>

alebo <http://sites.google.com/site/kogneuro/>

ISBN: 978-80-553-0816-6

Kontaktný e-mail: kopco@tuke.sk

PodĎakovanie: Táto kniha vznikla s podporou Kultúrnej a edukačnej grantovej agentúry Slovenskej republiky v rámci projektu č. 3/7300/09 „Kognitívna veda – stredoeurópsky interdisciplinárny magisterský študijný program.“

Obsah

ZOZNAM TABULIEK	V
ZOZNAM OBRÁZKOV	V
ZOZNAM SKRATIEK	VI
KAPITOLA 1 ÚVOD.....	7
KAPITOLA 2 MODELOVANIE A NEUROVEDY	9
2.1 DELENIE MODELOV	9
2.2 ORGANIZÁCIA A MODELOVANIE NA RÔZNYCH ÚROVNIACH	16
2.3 ZDROJE INFORMÁCIÍ PRI MODELOVANÍ A ATRIBÚTY MODELOV	17
2.4 HISTORICKÝ PREHLAD TEÓRIÍ FUNGOVANIA MOZGU A MYSLE	9
2.5 PRÍSTUPY K MODELOVANIU A UMELÁ INTELIGENCIA	13
KAPITOLA 3 LINEÁRNA REGRESIA A NEURÁLNY KÓD.....	19
3.1 LINEÁRNY MODEL PRE POPIS JEDNODIMENZIONÁLNYCH NEUROFYZIOLOGICKÝCH DÁT	20
3.2 ZHRNUTIE A ZÁVEREČNÉ POZNÁMKY	25
KAPITOLA 4 VÝPOČTOVÉ MODELY AKTIVÁCIE NEURÓNU.....	27
4.1 ELEKTRICKÉ CHARAKTERISTIKY NEURÓNU A AKČNÝ POTENCIÁL	27
4.1.1 <i>Elektrické vlastnosti neurónovej membrány</i>	28
4.1.2 <i>Pasívny elektrický obvod ako model pokojových vlastností membrány neurónu</i>	29
4.1.3 <i>Akčný potenciál</i>	33
4.2 MODELY AKTIVÁCIE NEURÓNU	34
4.2.1 <i>Štandardný konekcionistický model neurónu a Grossbergov model neurónu</i> 34	
4.2.2 <i>Integrate-and-fire model neurónu</i>	35
4.2.3 <i>Hodgkin-Huxleyho model neurónu</i>	36
4.3 ZHRNUTIE	39
KAPITOLA 5 KONEKCIONISTICKÉ MODELOVANIE	40
5.1 NEUROVEDECKÁ INŠPIRÁCIA KONEKCIONIZMU.....	40
5.2 HISTÓRIA KONEKCIONISTICKÉHO MODELOVANIA	41
5.3 KONEKCIONISTICKÉ MODELOVANIE A TEÓRIA ADAPTÍVNEJ REZONANCIE (ART).....	44

5.3.1	<i>Interakcie medzi krátkodobou a dlhodobou pamäťou v jednoduchom neurálnom modeli klasického podmieňovania</i>	44
5.3.2	<i>Zhrnutie</i>	52
KAPITOLA 6	TEÓRIA OPTIMÁLNEHO ROZHODOVANIA A MODELOVANIE PSYCHOFYZIKÁLNYCH DÁT	53
6.1	PSYCHOFYZIKA	53
6.2	MODEL ROZHODOVANIA PRE PSYCHOFYZIKÁLNE DÁTA	54
6.2.1	<i>Definícia modelu</i>	54
6.2.2	<i>Všeobecné vlastnosti modelu</i>	56
6.2.3	<i>ROC krivka pre II-2AFC experimenty</i>	57
6.2.4	<i>Index citlivosti a strannosť pre špeciálny prípad Gaussovských distribúcií a rovnakých variancií</i>	58
6.2.5	<i>Transformácia na jednotky z-scores</i>	60
6.2.6	<i>Zhrnutie a ďalšie modely</i>	60
KAPITOLA 7	ZÁVER	62
LITERATÚRA		63

Zoznam tabuliek

Tab. 1	Koncentrácie iónov podieľajúcich sa na signálnej funkcii neurónov.....	31
Tab. 2	Nernstovo napätie spôsobené na neurónovej membráne rozdielmi v koncentrácii iónov.....	32
Tab. 3	Fyzikálne veličiny a im zodpovedajúce psychologické veličiny, tak ako sa študujú v sluchovej psychofyzike.	53

Zoznam obrázkov

Obr. 2-1	Ilustrácia a definícia frenológie z Webstrovho slovníka zo začiatku 20. storočia.....	10
Obr. 2-2	Nákres Purkyného neurónu a granulórneho neurónu vytvorený Ramonom y Cajalom. Originál: Instituto Santiago Ramón y Cajal, Madrid, Španielsko.....	11
Obr. 2-3	Nákres Brocovej a Wernickeho oblasti mozgu, ktoré sú zodpovedné za produkciu a porozumenie reči u človeka.	12
Obr. 3-1	Príklad časového priebehu stimulu (červená hladká čiara) a odoziev neurónu (modrá séria impulzov) v experimente Wessela a kol.	20
Obr. 3-2	Hĺbka modulácie, ktorou bol neurón stimulovaný (hore) a pravdepodobnosť vygenerovania impulzu neurónom v odozve na stimuláciu (dole).	22
Obr. 3-3	Závislosť medzi hĺbkou modulácie (stimulom) a pravdepodobnosť vygenerovania impulzu neurónom (odozvou).	23
Obr. 4-1	Príklad elektrického správania sa neurónu: pokojový, synaptický a akčný potenciál, meraný vnútro bunkovo na membráne somy neurónu.	29
Obr. 4-2	Preparát pre štúdium elektrických vlastností neurónu.	29
Obr. 4-3	Odozva neurónu (zmena membránového napätia, obr. hore) na skokovú zmenu privedeného prúdu pri malých množstvách privedeného prúdu (obr. dole). V každom obrázku hrúbka čiar udáva zodpovedajúcu silu prúdu a napätia.....	30
Obr. 4-4	Ekvivalentný elektrický obvod popisujúci pasívne elektrické vlastnosti neurónovej membrány.	31
Obr. 4-5	Príkladový experiment pre vysvetlenie vzájomného pôsobenia difúzných a elektrických síl, ktoré spôsobujú vytvorenie náboja na membráne neurónu. ...	32
Obr. 4-6	Schéma Integrate-and-fire modelu aktivácie neurónu.....	34
Obr. 4-7	Príklad aktivácie Integrate-and-fire neurónu pri prezentovaní podprahového stimulu (modrá) a stimulu, ktorý vyvolá prekročenie prahu (červená). Dole: privedený vstupný prúd. Hore: Odozva neurónu.	35
Obr. 4-8	Frekvencia pálenia Integrate-and-fire neurónu pri postupnom zvyšovaní sily vstupného stimulu (prúdu).	36
Obr. 4-9	Schéma Hodgkin-Huxleyho modelu aktivácie neurónu.....	36

Obr. 4-10 Príklad dynamiky aktivácie Hodgkin-Huxleyho neurónu pri privedení skokového prúdu do neurónu v čase 5 ms. Hore: odpoveď neurónu. Stred: dynamika otvárania a zatvárania sa váhových premenných. Dole: priebeh zmien vodivosti iónových kanálov.	38
Obr. 5-1 Príklad sigmoidálnej funkcie definovanej ako $y = 1/(1+e^{-x})$	42
Obr. 5-2 Časová následnosť pri vytváraní klasického podmieneného reflexu.	44
Obr. 5-3 Jednoduchý neurálny model klasického podmieňovania	45
Obr. 5-4 Priebeh reprezentácie vstupného signálu v modeli CC.	46
Obr. 5-5 Schématické zakreslenie typickej skupiny neurónov, pre ktorú sa využíva učenie Instar a Outstar.	48
Obr. 5-6 Rozšírenie modelu CC o inhibičné spojenia medzi STM neurónmi.	49
Obr. 5-7 Prah-lineárna signálna funkcia	50
Obr. 6-1 Model rozhodovania pre experimentálnu paradigmu 1I-2AFC.	56
Obr. 6-2 Krivka izosensivity ROC pre experimentálnu paradigmu 1I-2AFC.	58
Obr. 6-3 Gaussovská distribúcia pravdepodobnosti.	59

Zoznam skratiek

1I-2AFC	“One-interval, 2-Alternatives Forced Choice”, “1 interval, 2 alternatívy, nutnosť voľby“
ART	Adaptive Resonance Theory, Teória adaptívnej rezonancie
CC	Classical conditioning, Klasické podmieňovanie
NN	Nearest Neighbor, Najbližší sused
ROC	Receiver Operating Characteristic, krivka izosezitivity
LTM	Long-term memory, Dlhodobá pamäť
STM	Short-term memory, Krátkodobá pamäť

Kapitola 1 Úvod

Naše porozumenie tomu, ako funguje ľudský mozog a myseľ, dnes veľmi rýchlo napreduje. Sčasti je to dané vylepšenými metódami pre behaviorálne štúdium správania sa človeka, sčasti pokrokom v biologických technikách, a sčasti dramatickým technologickým pokrokom, ktorý ponúka nové metódy pre sledovanie mozgovej aktivity ako aj pre spracovanie informácií zaznamenaných pri sledovaní mozgu. Ale možno najdôležitejším faktorom spôsobujúcim tento rýchly pokrok je, že k štúdiu mozgu sa pristupuje stále viac interdisciplinárne.

Neuroveda a kognitívna veda sú interdisciplinárne vedné odbory, ktoré majú svoje korene v biologických a psychologických vedách, a v rámci ktorých sa dnes väčšina nového výskumu o mozgu deje. Pri tomto výskume je čoraz dôležitejšie použitie infromatických a matematických metód, a to jednak pre získavanie a spracovanie dát, ale hlavne pre vytváranie a overovanie nových teórií o aktivácii neurálnych obvodov a o kognitívnych funkciách. Výpočtová neuroveda sa zameriava práve na teoretické štúdium neurálnych a kognitívnych funkcií, na vývoj teórií a simuláciu modelov na všetkých úrovniach fungovania mozgu. Účelom tohto textu je poskytnúť základný prehľad metód používaných vo výpočtovej neurovede.

V princípe každý matematický formalizmus a každý infromatický nástroj môže byť užitočný pre spracovanie neurálnych dát a pre popis kognitívnych funkcií. Tu sa v štyroch kapitolách zameriame na detailný popis vybraných typov modelov, ktoré sa používajú na charakterizovanie neurálnych a behaviorálnych dát na rôznych úrovniach abstrakcie. Najprv ilustrujeme použitie lineárnej regresie pre analýzu a popis neurálnych dát, pri ktorom hľadáme vzťah medzi vlastnosťami podnetu a aktiváciou neurónu. Potom podávame prehľad neurálnych modelov aktivácie jednotlivého neurónu. V ďalšej kapitole sa zameriavame na konekcionistické modely, ktoré používajú siete zjednodušených abstraktných výpočtových jednotiek podobných neurónom na popis jednoduchších kognitívnych funkcií a ich neurálnej bázy. Posledná kapitola popisuje štatistický model, ktorý ignoruje neurálnu bázu mozgovej funkcie, a snaží sa popísať len rozhodovanie a správanie sa človeka pri vykonávaní určitej perceptuálnej alebo kognitívnej úlohy.

Účelom tohto textu nie je (a nemôže byť) podať vyčerpávajúci prehľad o výpočtovej neurovede. Text predpokladá u čitateľa základné matematické a infromatické znalosti na úrovni úvodných všeobecných predmetov pre študentov technických a prírodovedných smerov. Taktiež užitočné sú základy neurobiológie a kognitívnej psychológie. Po preštudovaní textu by čitateľ mal získať prehľad o šírke metód používaných vo výpočtovej neurovede, a mal by vedieť implementovať a použiť modely tu popísané.

Ku knihe existuje web stránka, prístupná cez adresy <http://pcl.tuke.sk/vakn> alebo <http://sites.google.com/site/kogneuro/>. Na stránke sú doplnkové materiály, a to jednak multimédiá objasňujúce špecifické príklady niektorých konceptov, a jednak dáta alebo simulačné skripty, ktoré čitateľom umožnia priamo sa oboznámiť s niektorými modelmi. Táto kniha vznikla aj ako doplnkový text, založený na časti prednášok z predmetov Úvod do neurovied a Výpočtová a kognitívna neuroveda, ponúkaných na TU v Košiciach. Veľmi dôležitými doplnkovými materiálmi sú aj slidy a ostatné materiály z týchto dvoch predmetov, na ktoré sa webstránka knihy

odkazuje, a ktoré ponúkajú okrem výpočtovej neurovedy aj všeobecnejší úvod do kognitívnej neurovedy z biologického aj psychologického hľadiska. A na záver sú na stránke k dispozícii aj skriptá pre cvičenia z predmetov Úvod do neurovied a Výpočtová a kognitívna neuroveda, v ktorých sa riešia príklady priamo súvisiace s výkladom v tomto texte.

Dúfam, že táto kniha poslúži ako užitočný úvod, ktorý vyvolá u študentov neurovedy, kognitívnych vied, informatiky, kybernetiky, či umelej inteligencie záujem a zapálenie pre štúdium výpočtovej neurovedy, a priblíži problematiku teoretickej neurovedy aj širokému spektru ľudí, ktorí sa zaoberajú štúdiom mozgu a mysle z najrôznejších hľadísk.

Kapitola 2 Modelovanie a neuroveda

Výpočtová neuroveda sa dá charakterizovať ako aplikácia matematických a počítačových nástrojov pre štúdium nervovej sústavy. Výpočtová neuroveda (tiež označovaná ako teoretická neuroveda alebo kognitívne modelovanie) predstavuje jednu z troch základných metód neurovedeckého výskumu. Ďalšími dvomi základnými metódami sú **behaviorálne** psychologické experimenty, ktoré sa snažia pochopiť jednotlivé kognitívne funkcie pozorovaním správania sa človeka a iných živočíchov, a **neurobiologické** experimenty, ktoré skúmajú anatómiu (štruktúru) alebo fyziológiu (funkciu) nervového systému. Tieto tri základné prístupy k štúdiu nervovej sústavy sú vzájomne prepojené, a to tak, že nové poznatky, ktoré získame jedným prístupom, informujú a ovplyvňujú výskum založený na zostávajúcich dvoch prístupoch. Napríklad, nový fyziologický poznatok o tom, ako funguje nejaké mozgové centrum, môže ovplyvniť ako behaviorálny výskum zodpovedajúcej kognitívnej funkcie, tak aj neurálne modely príslušného mozgového obvodu. Podobne, návrh novej výpočtovej teórie nejakej kognitívnej funkcie je často podnetom pre nové fyziologické a behaviorálne experimenty, ktoré testujú predikcie založené na tejto teórii.

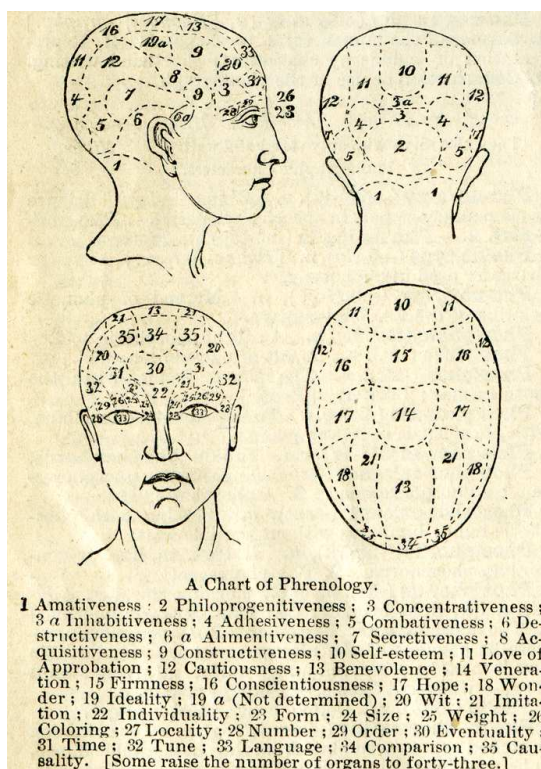
Základnou **funkciou** neurálneho modelu je zvyčajne **popísať experimentálne dáta**, či už anatomické, fyziologické, alebo behaviorálne. Druhou rovnako dôležitou úlohou modelov je **generovať predikcie**, ktoré je možné následne experimentálne testovať. Väčšinou existuje viacero modelov aj pre jednu a tú istú kognitívnu funkciu či neurálnu štruktúru. Tieto modely sa ale zvyčajne líšia, a to jednak tým, aké veľké množstvo existujúcich neurálnych dát vedia popísať (či sú zamerané len na jeden konkrétny fenomén alebo na celú všeobecnú funkciu), jednak svojou prediktívnou silou (niektoré modely sú len popisné a negenerujú žiadne nové predikcie, iné sú schopné vyprodukovať testovateľné hypotézy, overovaním ktorých strávia experimentátori roky). Preto je na začiatok dôležité uviesť si niekoľko základných rozdelení modelov, ktoré neurovedcom umožňujú rýchlejšie sa zorientovať v tom, čo daný model môže popísať a na aký typ dát a fenoménov je vhodný. Predtým sa ale pozrime trocha do histórie.

2.1 Historický prehľad teórií fungovania mozgu a mysle

Ludia sa otázkami ako *kde sídli duša* či *ako funguje myseľ* zaoberali od nepamäti. Zaujímavé historické zhrnutie na túto tému podáva Daugman (Daugman 2001), ktorého hlavným záverom je, že keďže myseľ je jedným z najkomplexnejších systémov aké poznáme, väčšina teórií o fungovaní mozgu navrhnutých v priebehu storočí bola vždy založená na metafore vychádzajúcej z vtedajších najnovších výsledkov vedeckého poznania a technického pokroku. Napr., keď antickí Gréci pozdvihli svoje technologické zručnosti vybudovaním fontán, kanálov a púmp, práve toto poznanie bolo základom pre grécky pneumatický koncept duše (grécky „pneuma“, alebo dych života) či pre Galenovu teóriu štyroch temperamentov. Podobne v dobe osvietenstva, zložité a presné mechanizmy hodínok motivovali úvahy o mysli ako zložitom mechanickom systéme (La Mettrie 1748) a vo Viktoriánskom období konštrukcia parných strojov motivovala napr. Freuda k úvahám o podvedomom fungovaní mysle na základe hydraulických prúdení v nervových vláknach (Freud 1913). Neskôr, konštrukcia telegrafných sietí s ich

zložitými prepojavacími reléovými členmi boli základom pre Helmholtzovu neurálnu metaforu (Helmholtz and Ellis 1875) a pre Hebbovu teóriu pamäti (Hebb 1949). Z tohto pohľadu nie je prekvapivé, že dnešný prudký pokrok v počítačových technológiách vytvára priestor pre hľadanie mnohých nových metafor pripodobňujúcich fungovanie mozgu k počítačovým systémom.

Užitočné môže byť pozrieť sa na niekoľko teórií mysle od konca 18. storočia, t.j. od doby kedy už existovala viac-menej zhoda, že existencia ľudskej mysle nejakým spôsobom súvisí s fungovaním mozgu. Tu spomenieme štyri staršie teórie, frenológiu, teóriu celkového poľa, neurónovú doktrínu a celulárny konekcionizmus, a definujeme dnešný pohľad na fungovanie mozgu a mysle. Zmienené teórie sú všeobecnými teóriami fungovania mozgu, nie detailnými (kvantitatívnymi) teóriami, ktorými sa budeme zaoberať neskôr v tejto knihe.



Obr. 2-1 Ilustrácia a definícia frenológie z Webstrovho slovníka zo začiatku 20. storočia.

Frenológia

Frenológia bola vedecká teória (dnes už sa za vedu nepovažuje), ktorú koncom 18. storočia definoval nemecký lekár Franz Joseph Gall, a ktorá bola v 19. storočí veľmi populárna. Na základe pozorovaní svojich pacientov prišiel Gall s vtedy radikálnym tvrdením, že:

1. mozog je základom správania sa,
2. kortex nefunguje ako jeden orgán, ale má špecializované oblasti,

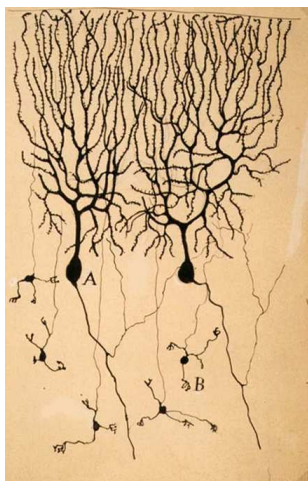
3. každá oblasť zodpovedá špecifickej mentálnej funkcii (Gall ich našiel 35),
4. pri častom používaní určitej funkcie sa zodpovedajúca oblasť zväčšuje, preto sú na lebke vyduté oblasti spojené švami.

Gall potom zostrojil nákres s rozložením mozgových „orgánov“ (viď Obr. 2-1), pri vytváraní ktorého postupoval tak, že funkcie mozgových orgánov určoval podľa povahy skúmaných ľudí a podľa toho, ktoré oblasti ich lebky sa mu javili ako vyduté, čo malo naznačovať, že funkcie v nich uložené sú u daného človeka rozvinutejšie. Samozrejme, dnes už málokto verí, že Gallova mapa je správna, a frenológia sa už vedecky neštuduje. Dôležité je ale pripomenúť, že v z uvedených štyroch bodov navrhnutých Gallom, prvé tri platia aj dnes.

Teória celkového poľa

Teória celkového poľa (angl. Aggregate-Field Theory) prišla ako protiteória k frenológii. Francúzsky psychológ Pierre Flourens odstránil resp. predelil niektoré mozgové oblasti experimentálnych zvierat, napr. králikov, čím chcel izolovať Gallove mozgové „orgány“. Keďže sa mu nepodarilo pozorovať žiadne zmeny v ich správaní, prišiel k záveru, že mozog nefunguje ako súbor špecializovaných orgánov, ale že celý kortex sa podieľa na každej mentálnej aktivite. Teóriu celkového poľa potom Flourens formuloval ako:

“All perceptions, all volitions occupy the same seat in these (cerebral) organs; the faculty of perceiving, of conceiving, of willing merely constitutes therefore a faculty which is essentially one.” voľne preložiteľné ako *“Všetky naše vnemy, všetky rozhodnutia sú výsledkom aktivity týchto (cerebrálnych) orgánov, ktoré zdieľajú rovnaký priestor v mozgu; naše schopnosti vnímať, myslieť, či chcieť sú teda v skutočnosti len jednou schopnosťou.”*



Obr. 2-2 Nákres Purkyného neurónu a granulárneho neurónu vytvorený Ramonom y Cajalom. Originál: Instituto Santiago Ramón y Cajal, Madrid, Španielsko.

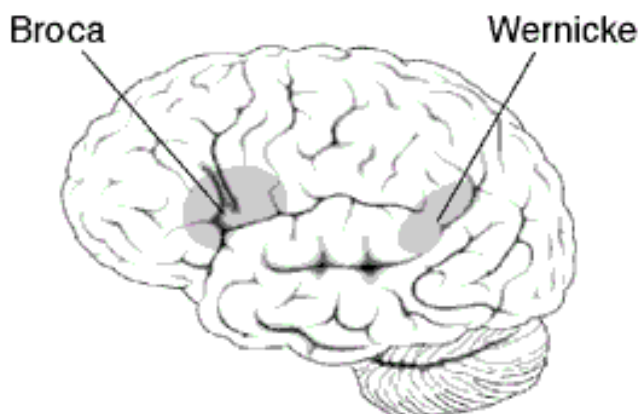
Neurónová doktrína

Až do konca 19. storočia sa verilo, že mozog je veľká žľaza, ktorej výlučky sa šíria v tele prostredníctvom nervov (v súlade s Teóriou celkového poľa; viď aj Freudovu

citáciu vyššie). Santiago Ramon y Cayal použil Golgiho farbivovú techniku, a ukázal, že neuróny sú nezávislé jednotky, ktoré majú zložitú štruktúru, a teda, že neplatí, že by bol mozog jednoliatym celkom (viď Obr. 2-2). Na základe Cayalových výsledkov začala v teóriách mozgu a mysle prevažovať **Neurónová doktrína**, ktorá hovorí, že jednotlivé neuróny sú hlavnými signálnymi jednotkami nervového systému.

Celulárny koneccionizmus

Po Ramónovi y Cajal mnoho vedcov prezentovalo výsledky, ktoré preukázali špecializáciu jednotlivých mozgových oblastí pre rôzne kognitívne funkcie. Veľmi významnými medzi nimi boli francúzsky fyziológ Pierre Paul Broca a nemecký fyziológ Carl Wernicke, ktorí pozorovali pacientov s lokálnym poškodením mozgu. Na základe týchto pozorovaní došli k záveru, že oblasť v ľavom frontálnom laloku je nevyhnutná pre normálnu schopnosť hovoriť, zatiaľ čo oblasť superiórneho temporálneho gyrusu ľavej hemisféry je nevyhnutná pre schopnosť človeka porozumieť jazyku (viď Obr. 2-3).



Obr. 2-3 Nákres Brocovej a Wernickeho oblasti mozgu, ktoré sú zodpovedné za produkciu a porozumenie reči u človeka (zdroj Národný inštitút zdravia, USA).

Slávne, na základe svojich pozorovaní, Broca v r. 1865 vyhlásil „Nous parlons avec l’hémisphere gauche”, čiže „Hovoríme ľavou hemisférou nášho mozgu“. Dnes sa už vie, že toto tvrdenie nie je až také presné, keďže rečové centrá sú prevažne v ľavej hemisfére len u pravákov, aj to nie u všetkých.

Na základe týchto výsledkov Wernicke, Sherrington a Ramon y Cayal navrhli, že jednotlivé neuróny sú signálnymi jednotkami mozgu, a že neuróny sú organizované vo funkčne špecifických skupinách, ktoré sú navzájom špecificky poprepájané. Tento pohľad sa nazýva **celulárny koneccionizmus**, a je vlastne rozšírením neurálnej doktríny y Cayala.

Súčasný pohľad

Aj keď mnohé otázky o fungovaní mysle sú nateraz nezodpovedané, súčasný rozvoj nových invazívnych a neinvazívnych techník pre štúdium mozgu nám

poskytuje stále nové a nové informácie. Na ich základe sa dnes väčšina neurovedcov zhoduje na nasledovnom súbore faktov:

- Neurón je základná signálna jednotka mozgu.
- Rôzne časti mozgu majú rôzne funkcie (sluchová, vizuálna, ...).
- Mozgové oblasti sú poprepájané relatívne presne stanoveným spôsobom.
- Konkrétna oblasť má približne rovnakú funkciu u každého jednotlivca.
- Poškodenie určitej časti mozgu môže viesť k dlhodobej reorganizácii, pri ktorej funkciu danej časti preberie iná mozgová oblasť. Táto schopnosť je ale obmedzená.
- Na komplexných aktivitách (napr. jazyk) sa podieľa viacero mozgových oblastí súčasne.

Posledný bod sa dnes dá ľahko ilustrovať sledovaním mozgovej aktivity človeka pri vykonávaní rôznych úloh súvisiacich s jazykom, napr. použitím Pozitronovej emisnej tomografie. Pri takomto experimente Posner a kolegovia (Posner, Petersen et al. 1988) ukázali, že:

- ak sa dívame na slová, aktivuje sa primárne vizuálna kôra,
- ak počúvame tie isté slová, aktivuje sa primárne sluchová kôra (spomínaná Wernickeho oblasť),
- ak hovoríme slová, aktivuje sa hlavne Brocova oblasť,
- **ale**, ak o slovách rozmyšľáme, aktivujú sa mnohé mozgové oblasti, vrátane tých uvedených vyššie.

Zhrnutie

Tento súhrn historického vývoja teórie o fungovaní ľudského mozgu ukázal, ako sa vo vede často postupuje od extrémnych ale jednoduchých teórií k teóriám komplexnejším, ktoré sú nevyhnutné pre pochopenie komplexných systémov (akým mozog a myseľ určite je). Aj keď spomínané teórie boli nesprávne, v zásade boli užitočné, lebo ich overovaním sa vytvárali protiteórie, ktoré posúvali poznanie dopredu.

2.2 Prístupy k modelovaniu a umelá inteligencia

V odbore umelá inteligencia, ktorý sa dá chápať ako predchodca dnešnej výpočtovej neurovedy, historicky existovali dva hlavné prístupy k modelovaniu kognitívnych procesov. **Symbolický** prístup, ktorý používal ako základný nástroj program pozostávajúci z pravidiel pre manipuláciu so symbolmi, a **konekcionistický** prístup, pri ktorom veľké množstvo zložito prepojených jednoduchých výpočtových jednotiek napodobňuje určité aspekty neurálnej štruktúry mozgu.

Symbolický prístup, dnes nazývaný aj GOF AI (čiže "Good Old-Fashioned Artificial Intelligence", keďže je historicky starší, zhruba z r. 1960) sa snažil vysvetliť zložité kognitívne procesy ako súbor jednoduchých symbolických manipulácií, t.j. ako súbor pravidiel definovaných formálnou logikou, ktoré operujú nad premennými reprezentujúcimi symboly. Konkrétne sa často jednalo o súbor predikátov „ak X potom Y“, ktoré sa aktivujú podľa toho, kedy je symbolická klauzula X splnená.

Typickými kognitívnymi procesmi modelovanými týmto prístupom bolo riešenie úloh, porozumenie reči, či pamäťové procesy. Ambíciou výskumníkov používajúcich symbolický prístup bolo navrhnuť systémy, ktoré by sa vyznačovali všeobecnou inteligenciou a správaním sa podobnými človeku. Známym príkladom ilustrujúcim túto ambíciu je tzv. Turingov test navrhnutý v 50-tych rokoch Alanom Turingom (Turing 1950). Tento test zjednodušene hovorí, že mysliaci stroj by mal zvládnuť tzv. *imitačnú hru*. Pri nej v úlohe, ktorej zvládnutie človekom by sme považovali za prejav jeho inteligencie, by sa stroj vedel správať tak, že by sme nevedeli rozlíšiť, či úlohu vykonáva mysliaci stroj alebo človek.

Možno preto, že mala GOFAI takéto ambiciózne ciele, sa dnes rozvíja pomerne málo. Vo všeobecnosti má symbolický prístup niekoľko výhod. Napr., tieto systémy vedia pracovať aj s *meta* znalosťami (vedia o sebe „povedať“, či nejakú znalosť majú alebo nemajú, resp. vedia vysvetliť, ako došli k nejakej odpovedi na nejakú otázku). Podobne je u týchto systémov relatívne ľahké vložiť do znalostnej bázy nejakú *expertnú* znalosť či informáciu (je ľahké fixne špecifikovať nejaké pravidlo, ktoré je známe vo všeobecnosti, a to bez potreby, aby sa ho systém sám potenciálne nepresne naučil). Za nevýhody symbolických systémov sa dá považovať veľmi obmedzená (často žiadna) schopnosť učiť sa, žiadna podobnosť s nervovou štruktúrou mozgu, ako aj sekvenčná aktivita týchto systémov (vždy sa v nich sekvenčne vyhľadáva jedno pravidlo, ktoré sa v danom časovom kroku vykoná).

Konekcionistický prístup k modelovaniu, ktorého prvé idey sa objavili už v 50-tych rokoch (McCulloch and Pitts 1943), sa začal rozvíjať hlavne okolo r. 1980. Tento prístup nepoužíva reprezentáciu založenú na symboloch a na logických operáciách nad nimi. Miesto toho sa snaží mozgové procesy popísať použitím výpočtovej štruktúry, ktorá je bližšia tomu, ako sa informácie spracúvajú v mozgu. Konkrétne, konekcionistické systémy pozostávajú z veľkého množstva zložito poprepájaných jednoduchých výpočtových jednotiek, ktoré sú aktivované paralelne a ich výpočtová sila vyplýva zo silného paralelizmu a zo zložitosti siete, do ktorej sú tieto výpočtové jednotky pospájané. Čiže, konekcionistické modely používajú výpočtový mechanizmus, ktorý je kvalitatívne podobný neurónom a synaptickým prepojeniam v biologickom mozgu. Typickým cieľom konekcionistického modelovania je popísať nejaký konkrétny mozgový systém (napr. vnímanie, pamäť a učenie, jazyk, hierarchie kognitívnych procesov), nie navrhnuť všeobecný systém s inteligentným správaním. Medzi výhody konekcionizmu môžeme zaradiť to, že 1) výpočet je podobný tomu čo sa deje v mozgu, a 2) že učenie, t.j. schopnosť, ktorú má každý biologický neurálny systém, je špecifikované už ako súčasť daného modelu. Nevýhodou konekcionizmu je hlavne to, že neurobiologicky je konekcionistická neurónova sieť priveľmi zjednodušená štruktúra, ktorá má so svojou biologickou obdobou spoločný väčšinou len masívny paralelizmus jednoduchých výpočtových jednotiek. Ďalšou nevýhodou je, že matematicky je neurónová sieť priveľmi komplexná, pretože predstavuje dynamický model, pre ktorý je veľmi ťažké ukázať a dokázať vlastnosti ako stabilitu či konvergenciu. Z pohľadu GOFAI je zas slabinou konekcionizmu obmedzená symbolickosť reprezentácie u týchto modelov (v distribuovaných konekcionistických systémoch nie je zaručené, že ak vieme reprezentovať informáciu „John loves Mary“, že z danej reprezentácie budeme vedieť identifikovať aj jej komponenty „John“, „loves“ a „Mary“).

Oba zmienené prístupy, **symbolický** aj **konekcionistický**, sa používajú aj dnes. Ale v zásade sa dnes umelá inteligencia často pojíma skôr ako aplikovanie akýchkoľvek matematických a inforatických prostriedkov pre štúdium mozgu a mysle, a to ako pre modelovanie jednotlivých kognitívnych procesov, tak aj len pre analýzu dát o rôznych mozgových funkciách. Čiže zaradiť tu môžeme akékoľvek štatistické, stochastické, informačno-teoretické a iné výpočtové nástroje, ktoré môžu byť použité či už ako model správania sa človeka, alebo napr. na spracovanie a analýzu biomedicínskych obrazových dát popisujúcich aktivity mozgu. A práve pre túto všeobecnosť je dnes výstižnejšie hovoriť o výpočtovej neurovede, než o umelej inteligencii. My sa v najbližších kapitolách zameriame na niektoré konekcionistické modely. Ale väčšinu času strávime modelmi, ktoré do tradičnej UI nezapadajú, a vychádzajú z deskriptívnej štatistiky, diferenciálneho počtu a popisu dynamických systémov, a zo štatistickej teórie rozhodovania a detekovateľnosti.

2.3 Delenie modelov

Existuje mnoho pohľadov na to, ako sa dajú neurálne modely klasifikovať. Často miera všeobecnosti a prediktívna sila modelov súvisí s tým, ako detailne sú definované. Najmenej špecifické sú väčšinou **verbálne modely**, ktoré sú často len o málo viac informatívne, než všeobecný popis danej funkcie. Vyššiu úroveň predstavujú **blokové schémy**, ktoré často popisujú jednak súčasti danej nervovej štruktúry a jednak vzájomné prepojenia, závislosti a nadväznosti medzi nimi (slávnym príkladom toho, ako zložitá bloková schéma neurálnej funkcie môže byť, je schéma vizuálneho systému človeka na Obrázku 3 v článku (DeYoe and Van Essen 1988)). Najšpecifickejšie modely sú modely **kvantitatívne**, ktoré priamo definujú matematický vzťah medzi jednotlivými súčasťami danej neurálnej štruktúry, ako aj medzi jej vstupmi (napr. vnemovými podnetmi) a výstupmi (napr. behaviorálnymi reakciami človeka). My sa budeme prakticky zaoberať len týmto posledným typom modelov.

Dayan a Abbott (Dayan and Abbott 2001) navrhli základné delenie modelov, ktoré je užitočné pre mnohé neurálne funkcie. Modely rozdeľujú na:

- **deskriptívne**, ktoré kompaktné a presne popisujú veľké množstvo experimentálnych dát, a ktorých cieľom je popísať určité fenomény, ale nie vysvetliť, prečo k nim dochádza,
- **mechanistické**, ktoré dané fenomény vysvetľujú na základe známej anatómie, fyziológie a detailných popisov zodpovedajúcich neurálnych obvodov, a
- **interpretatívne**, ktoré používajú výpočtové a informačno-teoretické metódy na hodnotenie behaviorálneho a kognitívneho významu rôznych aspektov fungovania nervovej sústavy. Tieto modely sa snažia vysvetliť, prečo určitý neurálny proces funguje tak ako funguje.

Keďže pojem interpretatívnych modelov je asi najmenej intuitívny, uvedieme si jeden príklad, v ktorom autor na základe informačno-teoretickej analýzy dokázal vysvetliť konkrétny aspekt neurálneho procesu. Tento príklad je model kódovania sluchovej informácie Lewickeho (Lewicki 2002). Zjednodušene, Lewicki ukázal, že ak si dáme ako výpočtový cieľ čo najefektívnejšie zakódovať sluchovú informáciu

s použitím sústavy filtrov s rozdielnou frekvenčnou selektivitou, tak v prípade, že takáto sústava filtrov bude vystavená takým zvukom, s akými sa človek stretáva v každodennom prostredí, optimálny neurálny sluchový kód bude založený práve na periférnych filtroch s takými parametrami, aké sa nachádzajú v našom vnútornom uchu.

Modely je možné často rozdeliť tiež na základe stratégie, ktorú na popis experimentálnych dát používajú. Dve základné stratégie sú anglicky označované ako bottom-up (zdola hore) a top-down (zhora dole). Pri **bottom-up** modeloch sa vychádza z predpokladu, že na pochopenie vyšších kognitívnych funkcií alebo zložitejších neurálnych štruktúr je nevyhnutné plné poznanie fungovania daných funkcií na nižších úrovniach. Napríklad, na to, aby sme pochopili, čo a ako sa v mozgu deje pri počúvaní hudby, potrebujeme porozumieť, ako sa spracúva jednoduchá zvuková informácia, napr. ako sa extrahujú jednotlivé tóny či zvuky nástrojov. Na to potom musí nadväzovať popis reprezentácie melódií, a tak ďalej až k vzťahu hudby a emócií či pamäti. Táto stratégia je bežná hlavne v neurobiológii. Pri **top-down** modeloch sa vychádza z predpokladu, že štruktúra nervového systému nie je dôležitá pre poznanie perceptuálnych a kognitívnych procesov, a že stačí, keď model spĺňa relevantné psychologické a výpočtové ohraničenia. Túto stratégiu zvyčajne uprednostňujú informatici.

Ako príklad **top-down** chápania neurálnych modelov spomeňme rozdelenie úrovni výpočtových teórií podľa Davida Marra, informatika, ktorý v 80-tych rokoch 20. storočia dramaticky posunul naše poznanie vizuálneho systému človeka. Marr (Marr 1982) navrhol tri úrovne, na ktorých je možné porozumieť a študovať systémy spracúvajúce informácie (pričom Marr tu uvažoval hlavne o ľudskom videní):

1. **Výpočtová úroveň** predstavuje abstraktnú analýzu problému, ktorej výsledkom je stanovenie nevyhnutných výpočtových úkonov, ktoré musí mozog (alebo iný systém) vykonávať pri jeho riešení. Napr., takou analýzou môže byť uvedomenie si, že na to, aby náš vizuálny systém bol schopný rozoznať kruhy a štvorce, musíme byť schopní rozlíšiť rovné čiary, hrany a oblúky.
2. **Úroveň reprezentácii a algoritmov** je nižšia úroveň, pri ktorej sa už špecifikuje ako konkrétne sa jednotlivé informácie reprezentujú a aká je formálna procedúra, na základe ktorej sa dá pre daný vstup predikovať správny výstup.
3. **Implementačná/hardwarevá úroveň** predstavuje analýzu fyzickej implementácie výpočtov, napr. v neurónoch a neurálnych sieťach v ľudskom mozgu.

Marr veril, že vyššie úrovne chápania systémov pre spracovanie informácií sú viac-menej nezávislé na nižších úrovniach (t.j. že výpočtová úroveň je nezávislá na algoritmickej a algoritmická je nezávislá na implementačnej úrovni). Tento postoj sa dnes nazýva aj **Doktrína nezávislosti**.

2.4 Organizácia a modelovanie na rôznych úrovniach

Zatiaľ čo bottom-up a top-down prístupy hovoria o tom, ako poznanie neurálneho systému na jednej úrovni ovplyvňuje alebo neovplyvňuje jeho poznanie na

inej úrovni, nervová sústava sa dá modelovať na rôznych úrovniach detailu alebo v rôznych časopriestorových mierkach aj osobitne. Neurovedy popisujú správanie sa nervovej sústavy na priestorových mierkach v rozsahu od nanometrov až po metre. Konkrétne, v mierkach nanometrov sa sledujú procesy uvoľňovania jednotlivých neuroprenáščov (jednoduchých molekúl) pri synaptickom prenose vzruchov. Na úrovni mikrometrov sa študujú jednotlivé synapsie. Na úrovni stoviek mikrometrov popisujeme neuróny. Na úrovni centimetrov sa dívame na jednotlivé neurálne mozgové mapy (napr. mapa vizuálneho priestoru v primárnej vizuálnej kôre), na úrovni desiatok cm popisujeme jednotlivé mozgové systémy (napr. sluchovú dráhu), no a na úrovni metrov sa potrebujeme dívať, keď študujeme centrálnu nervovú sústavu ako celok.

Podobne je možné definovať aj temporálne úrovne organizácie nervovej sústavy, od najkratších udalostí trvajúcich len pár mikrosekúnd (otvorenie iónového kanálu) až po dni, týždne, mesiace a roky, zodpovedajúce zmenám v pamäti, napr. pri dlhodobej potenciácii synapsie alebo pri raste nových synaptických spojení.

Nasledujúce modely predstavujú príklady modelov na rôznej úrovni organizácie, ktoré sú podrobnejšie popísané v tejto knihe. Na najnižšej úrovni je napr. Hodgkin-Huxleyho model membránového potenciálu, ktorý popisuje otváranie jednotlivých iónových kanálov a časový priebeh aktivácie jedného neurónu. Na nasledujúcej úrovni je to napr. kompartmentálny model neurónu, ktorý zachytáva morfológiu neurónu a aj časové a elektrické vlastnosti šírenia elektrických signálov v neuróne. Ešte vyššie sú jednoduché modely neurálnych obvodov, ktoré vysvetľujú jednoduché neasociatívne typy učenia sa ako senzitáciu a habituáciu. Na úrovni máp je známym príkladom Jeffersov model nervových jadier, v ktorých sú zakódované informácie o priestorových charakteristikách počutých zvukov. Na záver, príkladom modelov najvyššej úrovne sú napr. konekcionistické modely jazyka, ktoré väčšinou zahŕňajú celé skupiny jadier, z ktorých pozostáva daný neurálny systém.

2.5 Zdroje informácií pri modelovaní a atribúty modelov

Základným zdrojom informácií pri modelovaní sú dáta získané z behaviorálnych, neuroanatomických a neurofyziologických experimentov. Ako bolo spomenuté skôr, v zásade platí, že modely môžeme hodnotiť na základe šírky a množstva dát, ktoré vedia popísať, a na základe ich prediktívnej sily. Okrem týchto dvoch je možné definovať niekoľko ďalších atribútov, na základe ktorých sa dá vyhodnotiť kvalita modelu. **Základná funkčnosť** je primárny atribút, ktorý hovorí, že modelom popísaný vzťah by mal súhlasiť s experimentálnymi dátami. Ďalší dôležitý aspekt je **súlad s experimentálnym protokolom**, či už verbálnym pri behaviorálnych alebo technickým pri fyziologických experimentoch. Tu sa má na mysli hlavne to, či model skutočne vychádza len z tých informácií, ktoré mali pri experimente k dispozícii subjekty. **Distribúcia chýb** je veľmi dôležitým, ale často opomínaným aspektom modelov. Hovorí o tom, že ak model správne popisuje nejakú kognitívnu funkciu, tak súlad by nemal byť len medzi tým, ako presne model predikuje priemernú odpoveď nejakého systému, ale aj či v prípade, že model robí chyby, sú tieto chyby podobné chybám, ktoré robí modelovaný systém. **Chronologický súlad** je atribút, ktorý zdôrazňuje, či model je schopný popísať aj to, že správanie sa modelovaného systému sa môže meniť v závislosti na dostupných dátach a predošliých skúsenostiach. S tým súvisí aj atribút **rýchlosti učenia sa**, ktorý

charakterizuje, ako rýchlo sa model v porovnaní s modelovanou neurálnou štruktúrou je schopný dané správanie naučiť. Záverečný atribút je **transfer učenia a zovšeobecňovanie**, t.j. schopnosť extrapolovať a interpolovať poznatky z dát na nové dáta.

Uvedený zoznam atribútov nie je vyčerpávajúci a neplatí, že každý model by mal mať všetky atribúty. Napr., ak dáta, ktoré sa snaží model popísať, nie sú dostatočne detailné na to, aby sme mohli spoľahlivo odhadnúť distribúciu chýb v nich, potom nie je dôvod snažiť sa distribúciu chýb modelovať. Tento zoznam by mal modelárom skôr poslúžiť na zváženie toho, aké aspekty ich model má, aké nemá a aké by od neho oni očakávali.

Kapitola 3 Lineárna regresia a neurálny kód

Jedným zo základných problémov neurovedy je pochopiť, ako funguje neurálny kód. Ak pozorujeme aktivitu nejakého neurónu, o ktorom vieme, že tento neurón spracováva vizuálnu informáciu, vznikajú hneď dve otázky:

1. Akú informáciu konkrétne daný neurón kóduje, či extrahuje? Keďže v našom príklade sa dívame na vizuálny neurón, je tento neurón citlivý na farbu objektu, jeho tvar, vzdialenosť, či na počet objektov, ktoré sa aktuálne vo vizuálnej scéne nachádzajú?
2. Ako je v aktivite daného neurónu zakódovaná táto informácia? Ak vieme, že sledujeme neurón, ktorý páli (čiže sa aktivuje, generuje impulzy, akčné potenciály), keď sa mení vzdialenosť nejakého objektu, bude sa so vzdialenosťou meniť frekvencia pálenia, výška jednotlivých impulzov, či opozdenie medzi nejakými dvomi konkrétnymi impulzmi?

Možných odpovedí na obe otázky je veľmi veľa. A to je jedným z dôvodov, prečo súčasné poznanie fungovania mozgu je ešte stále ďaleko od stavu, kedy by sme mohli povedať, že mu skutočne rozumieme.

Dobre známym príkladom neurofyziologického experimentu, ktorý k poznaniu neurálneho kódu nejakej funkcie prispel, je experiment Davida Hubela a Thorstena Wiesel (Hubel and Wiesel 1959), ktorí študovali aktiváciu neurónov vo vizuálnej kôre mačky (viď aj video v doplnkovom materiáli na web stránke knihy). V krátkosti popíšeme detaily tohto experimentu. Experimentálnym subjektom bola uspatá mačka, ktorá mala otvorené viečka a hlavu zafixovanú tak, aby sa dívala na plátno. Do primárnej vizuálnej kôry mačky bola zavedená wolframová elektróda, ktorá snímala aktivitu neurónov. Výstup tejto elektródy bol vyvedený do zosilňovača a napojený na reproduktor. Celý experiment prebiehal v zatemnenej miestnosti a experimentátor manipuloval, čo sa premieta na plátno, pričom z reproduktora zneli zvuky kliknutí, zodpovedajúce páleniam sledovaného neurónu. Hubel a Wiesel prezentovali v tomto experimente stimuly, ktoré boli jasne biele, a väčšinou to boli štvorce resp. obdĺžniky, u ktorých menili ich hrúbku (mohla to byť aj čiara), polohu (posúvali ich po plátno), a orientáciu (boli natočené vodorovne, zvislo, alebo šikmo). Stimuly sa behom experimentu menili, a na určenie toho, na aký stimul je neurón citlivý, stačilo počúvať, kedy neurón odpovedal. V experimente sa ukázalo, že neurón odpovedal, keď čiara resp. hrana objektu:

- bola na určitom mieste na plátno,
- mala určitú špecifickú orientáciu.

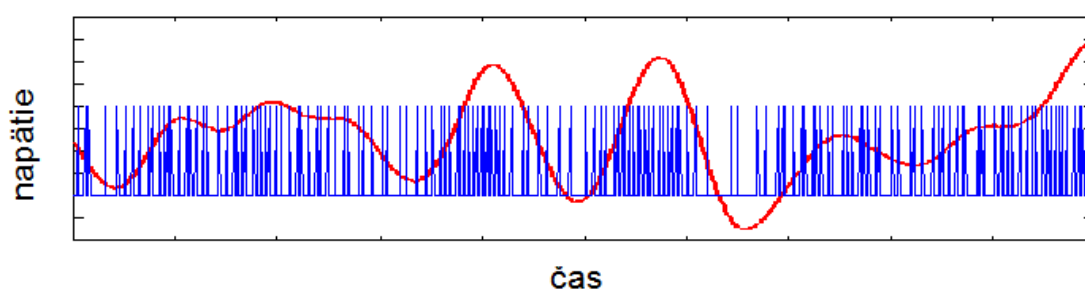
Keď experimentátori sledovali iné neuróny, pozorovali, že ich odpovede sa líšili len v tom, na akú polohu a na akú orientáciu boli citlivé. Na základe tohto pozorovania Hubel a Wiesel navrhli verbálny model funkcie neurónov v primárnej vizuálnej kôre (označovanej V1 alebo tiež ako Brodmannova oblasť 17), ktorý hovorí, že neuróny v tejto oblasti tvoria mapu, ktorá reprezentuje vizuálny priestor tak, že určuje, kde v priestore vidíme hrany akej orientácie. Navyiac pre tieto neuróny platilo, že sa menila frekvencia impulzov, ktoré neuróny generovali pre stimuly na rôznej polohe a s rôznou orientáciou. V zásade platilo, že frekvencia bola najvyššia

pre určitú polohu a orientáciu, a že keď sa poloha alebo orientácia postupne menila, frekvencia pálenia klesala. Čiže sa dá povedať, že charakteristikou pálenia, ktorú neuróny na kódovanie informácie o polohách a orientáciách hrán objektov používajú, je frekvencia pálenia.

3.1 Lineárny model pre popis jednodimenzionálnych neurofyzilogických dát

Ako formálne postupovať pri analýze neurálnych dát podobných tým z experimentu Hubela a Wiesela? Vo všeobecnosti existuje množstvo procedúr pre analýzu vzťahu medzi dvomi veličinami. Základnou štatistickou metódou, ktorá hľadá lineárny vzťah medzi dvomi premennými je *lineárna regresia*. Táto metóda má všeobecnú platnosť a používa sa bežne v mnohých vedeckých odboroch. V tejto kapitole ukážeme, ako sa lineárna regresia uplatňuje pre analýzu neurálnych dát, a definujeme aj matematické operácie *konvolúcie*, *korelácie*, a *Pearsonov korelačný koeficient*, ktoré sú pri popise experimentálnych dát dôležité.

Lineárna regresia predpokladá, že existuje jedna nezávislá premenná, ktorú môžeme voľne experimentálne meniť, a jedna závislá premenná, ktorá nadobúda rôzne hodnoty v závislosti na hodnotách nezávislej premennej (pričom, pre jednu hodnotu nezávislej premennej môžeme pri opakovanom meraní získať rôzne hodnoty závislej premennej). Keďže Hubel a Wiesel sledovali kódovanie neurónov, ktorých vstup sa menil vo viacerých parametroch (vertikálna poloha, horizontálna poloha, orientácia) na ilustráciu použijeme príklad Wessela a kol. (Wessel, Koch et al. 1996), ktorí skúmali kódovanie amplitúdovej modulácie v senzorickej neuróne elektrickej ryby latinsky nazývanej *Eigenmannia*. Táto ryba má špeciálny orgán, ktorým generuje slabé oscilujúce elektrické pole s frekvenciou niekoľko sto Hz, a zároveň má elektrosenzorický orgán, citlivý na takéto elektrické pole. *Eigenmannia* toto pole používa pre elektrolokalizáciu objektov vo svojom okolí a pre komunikáciu (doplňujúci materiál k experimentu Wessela a kol., vrátane tu analyzovaných stimulov, je k dispozícii na webstránke).



Obr. 3-1 Príklad časového priebehu stimulu (červená hladká čiara) a odozvy neurónu (modrá séria impulzov) v experimente Wessela a kol.

Wessel a kol. postupovali v experimente tak, že v blízkosti senzorickej orgánu ryby umelo generovali elektrické pole, pričom menili hĺbku jeho amplitúdovej modulácie. Súčasne zaznamenávali aktivitu neurónov v elektrosenzorickom orgáne ryby. Príklad stimulu (priebeh meniacej sa hĺbky modulácie elektrického poľa) a odozvy neurónu (impulzov vygenerovaných neurónom v neurosenzorickom orgáne) je uvedený na Obr. 3-1. Stimuly a odpovede sú v Obr. 3-1 značne odlišné. Môžeme si

v nich ale všimnúť jednu podobnosť: tam, kde je hĺbka modulácie stimulu veľká, tam je veľká aj hustota impulzov generovaných neurónom. To naznačuje, že sledovaný neurón je citlivý na hĺbku modulácie, a že túto hĺbku kóduje frekvenciou svojho pálenia. Aby sme toto pozorovanie overili, môžeme sériu impulzov zodpovedajúcich pálieniu neurónu previesť na pravdepodobnosť pálenia alebo na frekvenciu pálenia. Jednou z možností, ako túto transformáciu urobiť, je použitie *konvolúcie* so správne zvoleným konvolučným kernelom (jadrom).

Konvolúcia dvoch nekonečných časových radov g_i a h_i operácia je definovaná nasledovne:

$$(g * h)_i = \sum_{j=-\infty}^{\infty} g_{i-j} h_j$$

V prípade, že sú časové rady konečné, používa sa pri konvolúcii doplnenie nulami, alebo cyklické opakovanie radov. Po doplnení nulami sa rovnica zmení na:

$$(g * h)_i = \sum_{j=0}^{n-1} g_{i-j} h_j$$

Aj keď rady g a h sú v rovnici konvolúcie rovnocenné, typicky sú ich funkcie odlišné: jeden z nich je zvyčajne dlhý, a reprezentuje signál, zatiaľ čo druhý je krátky, koncentrovaný okolo $j=0$, a nazýva sa filter alebo konvolučný kernel (jadro, maska). Čo sa prakticky pri konvolúcii deje, je veľmi jednoduché. Časové rady g a h sa posúvajú „protismerne“ jeden proti druhému, a pre každý časový posun sa znásobia jednotlivé zarovnané členy a výsledok znásobenia sa sčíta. Napr., pre člen $(g * h)_0$ vyzerajú rozvinuté konečné rady g a h takto:

$$\begin{array}{ccccccccccc} \cdots & g_{m-1} & \cdots & g_1 & g_0 & 0 & 0 & 0 & 0 & \cdots \\ \cdots & 0 & \cdots & 0 & h_0 & h_1 & \cdots & h_{n-1} & 0 & \cdots \end{array}$$

a pre člen $(g * h)_1$ vyzerajú rozvinuté konečné rady g a h takto:

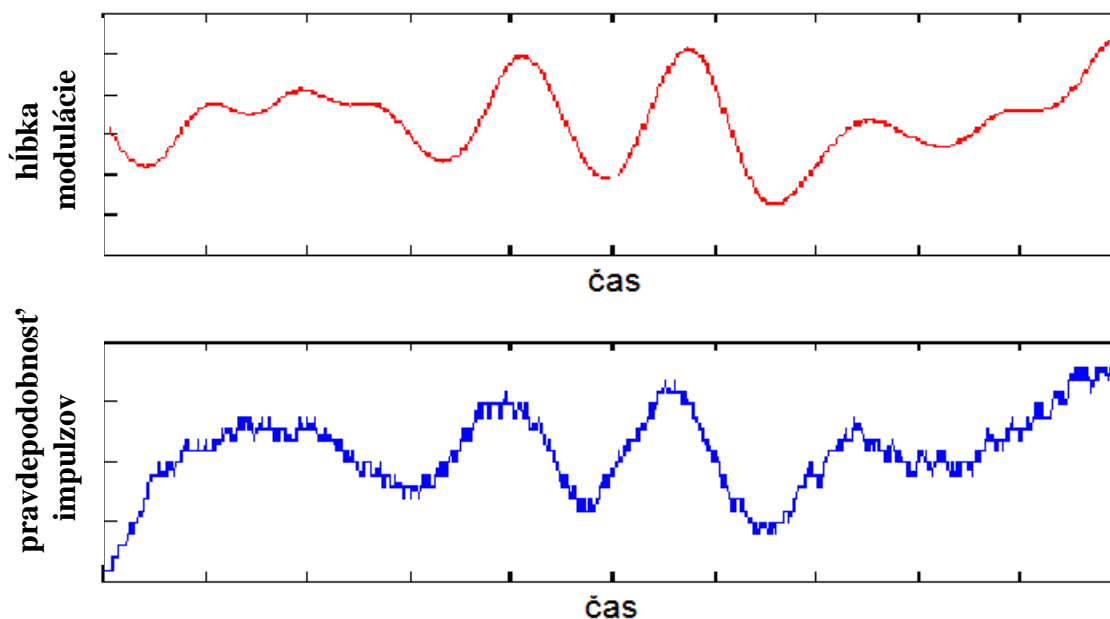
$$\begin{array}{ccccccccccc} \cdots & 0 & g_{m-1} & \cdots & g_1 & g_0 & 0 & 0 & 0 & \cdots \\ \cdots & 0 & \cdots & 0 & h_0 & h_1 & \cdots & h_{n-1} & 0 & \cdots \end{array}$$

pričom uvažujeme, že rad g_i má dĺžku m a rad h_i má dĺžku n . Pre každý časový posun sa v uvedených tabuľkách vynásobia členy nad sebou, a výsledný rad sa sčíta. Výsledok konvolúcie dvoch konečných radov je teda časový rad s dĺžkou $m+n-1$.

Konvolúcia je matematicky veľmi jednoduchá operácia. Ale má veľké množstvo užitočných vlastností, ktoré ju umožňujú použiť napr. na filtrovanie dát, ale aj na ich vyhladzovanie alebo na nájdenie pravdepodobnosti pálenia neurónu, čím sa teraz vrátíme k nášmu problému. Na výpočet pravdepodobnosti impulzov zo zaznamenaných sekvencie impulzov sa používa konvolúcia v tvare:

$$p_i = \sum_j \rho_{i-j} w_j$$

kde ρ je binárna sekvencia impulzov (sekvencia odpovedí neurónu z Obr. 3-1, zaznamenaná tak, že prítomnosť impulzu je reprezentovaná jednotkou a neprítomnosť nulou), w je jadro spĺňajúce podmienku, že $\text{suma}(w)=1$, a p je hľadaná pravdepodobnosť impulzov. Najjednoduchšie jadro, ktoré spĺňa túto podmienku, je pravouhlé jadro s dĺžkou n a hodnotou $w_j = 1/n$.



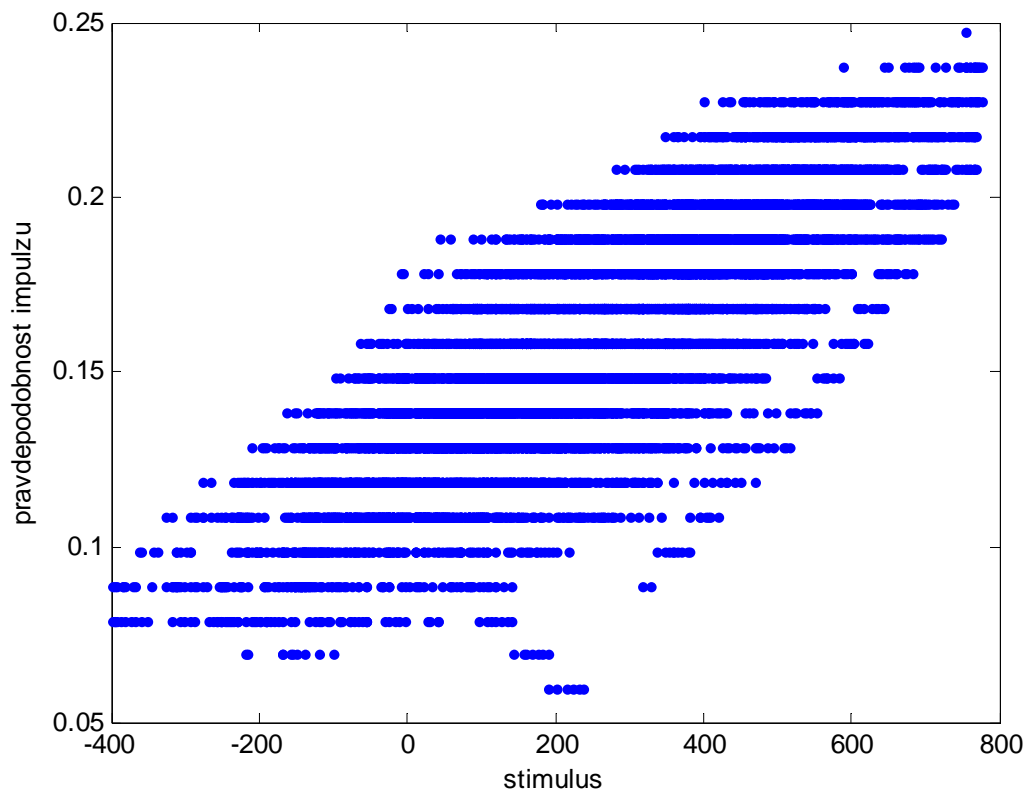
Obr. 3-2 Hĺbka modulácie, ktorou bol neurón stimulovaný (hore) a pravdepodobnosť vygenerovania impulzu neurónom v odozve na stimuláciu (dole).

Na Obr. 3-2 (dole) je vykreslený časový priebeh pravdepodobnosti generovania impulzu neurónom, získaný konvolúciou časového radu z Obr. 3-1 s jednoduchým pravouhlým kernelom. Z tohto obrázku už je zrejmé, že medzi stimulom (hĺbkou modulácie; Obr. 3-2 hore) a odozvou (pravdepodobnosťou impulzov zodpovedajúcou frekvencii pálenia; Obr. 3-2 dole) existuje nejaký vzťah. Tento vzťah sa dá vykresliť ešte jasnejšie, ak sa pre každý časový okamih vykreslí priamo stimul a odozva. Takto vykreslené dáta sú na Obr. 3-3. Z tohto obrázku je zrejmé, že závislosť medzi stimulom a odozvou existuje, ale že nie je ideálna. Konkrétne, ak by pre každú hĺbku modulácie neurón pálil jedinou presne špecifickou frekvenciou, na obrázku by bola iba jedna čiara. To, že na obrázku je zhuk bodov, nám hovorí, že závislosť medzi stimulom a odozvou je ovplyvnená aj inými faktormi, ktoré nateraz budeme považovať za šum, a ktoré budeme ignorovať.

Ďalšou dôležitou otázkou je, či je medzi stimulom a odozvou vzťah, ktorý je lineárny, exponenciálny, alebo ešte zložitejší (napr. polynomiálny). V princípe je možné vyšetřovať všetky tieto možnosti. My ale budeme uvažovať len o tej najjednoduchšej, budeme hľadať lineárny vzťah medzi stimulom a odozvou. A práve na určenie tohto lineárneho vzťahu slúži *lineárna regresia*, ktorá hľadá lineárnu závislosť popísanú modelom:

$$y_i \approx a + bx_i$$

kde y je odozva (pravdepodobnosť impulzu), x je stimul, a je priesečník priamky s osou Y , a b je smernica priamky. Premenné a a b sú hľadané parametre modelu, a našim cieľom je nájsť také ich hodnoty, ktoré najlepšie vyjadria vzťah medzi stimulom a odozvou. Ako sme už uviedli, tento popis nebude bezchybný, a to z dvoch hlavných príčin: a) prítomnosť šumu pri meraní a b) nelinearita dát.



Obr. 3-3 Závislosť medzi hĺbkou modulácie (stimulom) a pravdepodobnosťou vygenerovania impulzu neurónom (odzvou).

Štandardnou metódou hľadania optimálneho a a b je metóda najmenších štvorcov, ktorá minimalizuje chybu:

$$E = \sum_{i=1}^m \frac{1}{2} (a + bx_i - y_i)^2$$

Pre určenie optimálnych hodnôt a a b postupujeme tak, že definujeme podmienku, aby parciálne derivácie E podľa a a b boli rovné 0, čím vznikne systém dvoch lineárnych rovníc o dvoch neznámych

$$\begin{aligned}
 ma + b \sum_i x_i &= \sum_i y_i \\
 a \sum_i x_i + b \sum_i x_i^2 &= \sum_i y_i x_i
 \end{aligned}$$

Substitúciou strednej hodnoty $\langle x \rangle$, druhého momentu $\langle x^2 \rangle$, a korelácie $\langle xy \rangle$ definovaných nasledovnými vzorcami

$$\begin{aligned}
 \langle x \rangle &= \frac{1}{m} \sum_{i=1}^m x_i & \langle x^2 \rangle &= \frac{1}{m} \sum_{i=1}^m x_i^2 \\
 \langle y \rangle &= \frac{1}{m} \sum_{i=1}^m y_i & \langle xy \rangle &= \frac{1}{m} \sum_{i=1}^m x_i y_i
 \end{aligned}$$

sa naša sústava rovníc zjednoduší na

$$\begin{aligned}
 a + b \langle x \rangle &= \langle y \rangle \\
 a \langle x \rangle + b \langle x^2 \rangle &= \langle xy \rangle
 \end{aligned}$$

Túto sústavu vyriešime vyjadrením a z prvej rovnice

$$a = \langle y \rangle - b \langle x \rangle$$

čím získame riešenie pre b v tvare

$$b = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\langle x^2 \rangle - \langle x \rangle^2}$$

Menovateľ tohto zlomku sa nazýva variácia (rozptyl) a udáva, ako veľmi sa menia hodnoty x . Jeho čitateľ sa nazýva kovariancia, a udáva vzájomnú závislosť medzi x a y (z vyššie uvedenej definície korelácie si môžeme všimnúť, že táto operácia je takmer identická s operáciou *konvolúcie*, až na to, že korelované časové rady sa neposúvajú „proti sebe“, ale „rovnakým smerom“).

Čiže na to, aby sme našli optimálny lineárny vzťah dvoch premenných (nezávislej a závislej premennej), stačí nám vydeliť ich kovarianciu variáciou nezávislej premennej x . Ďalší krok je už len dosadenie do rovnice pre výpočet a (nájdenie konkrétnych hodnôt a a b pre náš príklad je úlohou jedného zo zadaní, popísaného v doplnkových materiáloch na webstránke knihy).

Poznámka: Keďže regresná analýza je určená pre vyhodnotenie závislosti závislej premennej na nezávislej premennej, výsledok analýzy sa zmení, pokiaľ tieto premenné jednoducho zameníme. A takáto zámena nie je správna. Taktiež nie je správne používať túto metódu pre analýzu optimálneho lineárneho vzťahu medzi dvomi náhodnými premennými.

Pearsonov korelačný koeficient

Posledným dôležitým konceptom regresnej analýzy je štatistická miera nazývaná *Pearsonov korelačný koeficient*. Tento koeficient sa označuje r a používa sa na vyhodnotenie miery lineárnej závislosti v dátach. Je definovaný ako:

$$r = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\sqrt{\langle x^2 \rangle - \langle x \rangle^2} \sqrt{\langle y^2 \rangle - \langle y \rangle^2}}$$

a jeho význam je taký, že r blízke ± 1 hovorí, že lineárny model je dobrý (resp., že lineárna závislosť medzi dátami x a y je silná), zatiaľ čo r blízke 0 hovorí, že nie. Keďže význam jednotlivých členov rovnice pre r sme definovali na predošlej strane, vidíme, že korelačný koeficient vyjadruje kovarianciu medzi dátami normalizovanú súčinom štandardných odchýlok v dátach.

Často sa na vyhodnotenie miery lineárnej závislosti v dátach (alebo lineárneho regresného modelu) používa r^2 , ktoré hovorí, „akú časť variancie v dátach je daný lineárny model schopný vysvetliť“. Na to, aby sme videli, ako korelačný koeficient súvisí s variáciou, a teda čo predošlá veta vlastne hovorí, najprv pripomeňme, že náš lineárny model je optimálny v zmysle metódy najmenších štvorcov, t.j. že chybu sme v ňom definovali ako variáciu odchýlky dát od priamky $bx+a$:

$$E = \sum_{i=1}^m \frac{1}{2} (a + bx_i - y_i)^2$$

Táto variácia (čiže táto chyba) je nulová, len keď je model dokonalý. Význam korelačného koeficientu sa dá chápať tak, že tento koeficient meria pokles vo variácii (v chybe), ktorý získame použitím lineárneho modelu dát (model prvého rádu) oproti použitiu modelu založeného len na jednej konštante (model $y=a$, kde a je napr. priemerná hodnota, čo by bol model nultého rádu). Aby sme našli najlepší model nultého rádu, stačí, ak v predošlej rovnici dáme $b=0$. E je v tomto prípade minimálne pre $a=\langle y \rangle$ (stredná hodnota y). Konkrétne bude hodnota E_0 (index znamená rád modelu) priamo úmerná variácii y . Pri hľadaní najlepšieho modelu prvého rádu hľadáme hodnoty dvoch premenných a a b tak, aby sme minimalizovali E_1 . Hodnota E_1 sa v porovnaní s modelom nultého rádu zníži. A ak vypočítame pomer tejto novej hodnoty E_1 k tej starej hodnote E_0 , zistíme, že tento pomer je rovný $1-r^2$. To znamená, že r^2 vyjadruje aká časť variancie v y (našej chyby E) bola vysvetlená pridaním lineárneho člena b do nášho modelu nultého rádu.

3.2 Zhrnutie a záverečné poznámky

Na jednoduchom príklade sme ilustrovali *lineárnu regresiu*, základnú štatistickú metódu pre popis lineárnej závislosti neurálnych (a iných) dát. Základnými rozšíreniami tohto modelu, ktoré ho robia oveľa praktickejšim, je uvažovanie o viacerých nezávislých premenných, alebo uvažovanie o nelineárnych závislostiach medzi premennými. Rozšírenie na viacero nezávislých premenných je relatívne jednoduché, a nepridáva takmer žiadne matematické komplikácie. Podobne je tento model ľahko možné použiť na analýzu nelineárnej závislosti pokiaľ je táto závislosť monotónna (napr. exponenciálna). Vtedy stačí dáta pred analýzou pretransformovať

na jednotky, v ktorých je závislosť lineárna (pre exponenciálnu závislosť by to bola logaritmickej transformácia). Pre zložitejšie modely odkazujeme čitateľa na literatúru o deskriptívnej a inferenčnej štatistike [odporúčame (Devore and Peck 1994; Mann and Lacke 2010)].

Kapitola 4 Výpočtové modely aktivácie neurónu

V tejto kapitole popisujeme elektrické správanie sa neurónov, ktoré predstavuje elementárny princíp spracovania informácií, komunikácie medzi neurónmi a šírenia signálov v mozgu. Potom podávame stručný prehľad najdôležitejších modelov aktivácie neurónu, od najjednoduchších po najzložitejšie.

4.1 Elektrické charakteristiky neurónu a akčný potenciál

V ľudskom mozgu je približne 10^{11} neurónov, ktoré sú poprepájané rádovo 10^{15} synaptickými spojeniami. Základné rozdelenie mozgových buniek je na:

- neuróny: hlavné signálne bunky, ktoré spracúvajú vnemy, riadia motorickú činnosť, podieľajú sa na učení, pamäťových a všetkých ostatných kognitívnych funkciách,
- gliové (podporné) bunky: ktorých informačná úloha je nejasná. Vie sa, že napr. spevňujú štruktúru mozgu, elektricky izolujú časti nervových vlákien, čistia medzibunkové prostredie, atď.

V zjednodušenom ponímaní sa dá povedať, že sa neurón skladá z troch základných častí: tela (nazývaného aj soma), dendritického stromu, ktorý tvorí vstup, cez ktorý sa šíria signály do neurónu, a z axónu, ktorý predstavuje výstup neurónu, ktorým sa šíria elektrické signály k ďalším neurónom. Prepojenie medzi axónom jedného neurónu a dendritickým stromom ďalšieho neurónu sa väčšinou deje prostredníctvom synaptických spojení, relatívne zložitým elektrochemickým procesom.

V dendritickom strome sa signály z predchádzajúcich neurónov pasívne šíria, kombinujú a integrujú. Pre aktiváciu neurónu je najdôležitejším miestom axónový výbežok, rozhranie medzi telom bunky a axónom, kde sa v závislosti na napätí, šírenom a kombinovanom v dendritickom strome, generujú impulzy, ktoré sa potom aktívne šíria do synaptických spojení a k ďalším neurónom.

V tejto časti sa zameriame na popis generovania impulzov. Platí, že drvivá väčšina buniek v ľudskom mozgu (s výnimkou niektorých sensorických neurónov, napr. tyčínok a piestikov v oku), impulzy generuje, a že ich generuje veľmi stereotypickým spôsobom. Takže poznanie toho, ako sa impulzy generujú, znamená poznanie základného princípu fungovania mozgu.

Dôležitou charakteristikou, ktorou sa neuróny odlišujú od ostatných buniek v našom tele je to, že na svojom povrchu majú dráždivú membránu. Táto membrána má niekoľko špecifických vlastností. Jednou je tá, že táto membrána si udržiava pokojové elektrické napätie, ktoré môžeme odmerať, ak porovnáme vnútro- a mimobunkový potenciál. Druhou je, že elektrické vlastnosti membrány neurónu (vodivosť alebo napätie na membráne) sa môžu meniť v závislosti na elektrických a chemických vlastnostiach jej okolia. A pre túto schopnosť zmeny elektrických vlastností sa neurón nazýva excitovateľnou (vybuditeľnou či dráždivou) bunkou.

Neurónová membrána je približne 3-4 nm hrubá, a tvoria ju tri dôležité časti:

- **Lipidická dvojvrstva** predstavuje súvislý plášť neurónu. Má hydrofóbické zakončenia, ktoré vytvárajú nepolarizovanú oblasť v jej vnútri. Táto dvojvrstva sa správa ako elektrický izolátor.
- **Iónové kanály** sú bielkoviny v lipidickej dvojvrstve, prepájajúce vonkajšiu stranu neurónu s vnútornou. Vedia sa otvárať a zatvárať, čím umožňujú riadené presúvanie iónov z a do neurónu. (Okrem iónových kanálov sú v membráne aj rôzne iné druhy bielkovín bez jasnej funkcie.)
- **Iónové pumpy**, ktoré udržujú rozdielnu koncentráciu iónov vnútri a mimo bunky.

Iónové kanály sú maličké (len 10 nm hrubé), a ich základná vlastnosť je, že pokiaľ sú otvorené a prepúšťajú ióny, znižujú odpor membrány až 10000-násobne. Existujú rôzne typy kanálov, niektoré sú veľmi selektívne, niektoré prepúšťajú rôzne ióny. Dôležité ale je, čo určuje, či je kanál otvorený alebo nie. Z tohto hľadiska riadenia (alebo „bránovania“) kanálov existujú dva základné typy:

- **napät'ovo riadené** kanály sa otvárajú a zatvárajú v závislosti na membránovom potenciáli,
- **chemický riadené** kanály sa otvárajú alebo zatvárajú v závislosti na vnútrobunkovej koncentrácii niektorých prvkov (napr. Ca^{2+}), alebo na mimobunkovej koncentrácii neuroprenášačov (neurotransmitters) v synapsii.

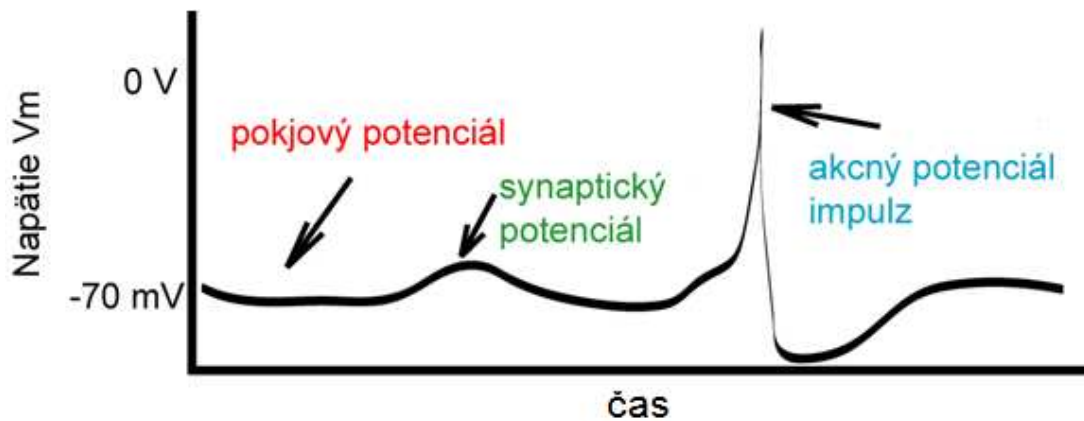
4.1.1 Elektrické vlastnosti neurónovej membrány

Na sledovanie aktivity nervového tkaniva sa používajú dva typy merania, vnútro- a mimobunkové. In vivo merania tkaniva neuspaných zvierat sa robia použitím mimobunkového (extracelulárneho) merania aktivity, pri ktorom sa vlastne sleduje akýsi súčet zmien náboja na membránach neurónov v okolí elektródy, ktorá je v mimobunkovom prostredí. Preto je citlivosť tohto merania veľmi malá, a nameraná amplitúda dosahuje len asi 100 μV . Tieto merania sú vo všeobecnosti ťažké, a je prakticky nemožné určiť, ktorý z neurónov v okolí elektródy sa aktivoval.

Na detailné pochopenie správania sa jednotlivého neurónu potrebujeme vnútrobunkové (intracelulárne) meranie. Toto umožňuje priamo merať rozdiel potenciálov medzi vnútrom a vonkajškom bunky. Jeho amplitúda dosahuje približne 100 mV, a väčšinou sa robí in vitro (v preparáte, v ktorom je z mozgu oddelená tenká vrstva tkaniva, ktoré sa rozpustí tak, aby sa jednotlivé neuróny uvoľnili). V nasledujúcej časti podrobne popíšeme správanie sa neurónu pri takomto meraní.

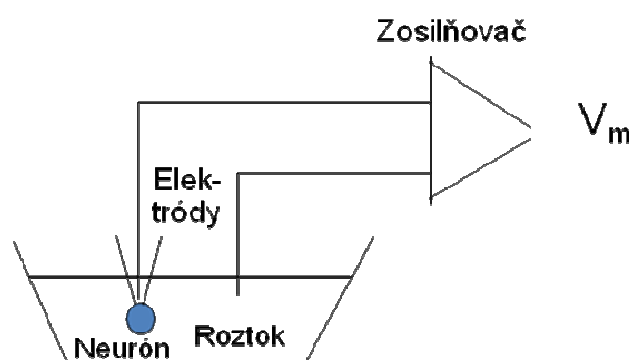
Obr. 4-1 ukazuje príklad merania membránového napätia na tele neurónu. Bez stimulácie z presynaptických neurónov je neurón v pokoji, v ktorom si udržiava **pokojevý potenciál** približne -70mV. Ak je neurón budený inými presynaptickými neurónmi, jeho membránové napätie sa mení. Rádovo sa impulz šírený z jedného presynaptického neurónu prejaví tak, že membránové napätie sa krátkodobo zmení v jednotkách milivoltov (toto ale silne závisí na mnohých faktoroch, ako napr. na tom, ako je aktivovaná synapsia ďaleko od tela neurónu). Ak sa impulzy z viacerých presynaptických neurónov objavia v rovnakom čase, synaptický potenciál z nich sa v dendritickom strome neurónu sčítava, čoho výsledkom môže byť relatívne veľká

zmena membránového napätia. Tu ale ešte stále platí, že neurón sa správa približne lineárne.



Obr. 4-1 Príklad elektrického správania sa neurónu: pokojový, synaptický a akčný potenciál, meraný vnútrobunkovo na membráne somy neurónu.

To, že neuróny sú excitovateľné (dráždivé), sa prejaví, ak presynaptická aktivita spôsobí zvýšenie synaptického potenciálu nad prah s hodnotou približne -50mV . V tom prípade neurón vygeneruje impulz (nazývaný aj akčný potenciál, AP, pálenie), ktorý trvá cca. 1 ms a ktorý pozostáva z depolarizácie (prudkého zvýšenia membránového napätia voči jeho pokojovej hodnote) a následnej hyperpolarizácie (prudkého zníženia napätia až na úroveň nižšiu než pokojový potenciál). Až po nej sa neurón dostáva do pokojového stavu, v ktorom môže generovať nové impulzy.



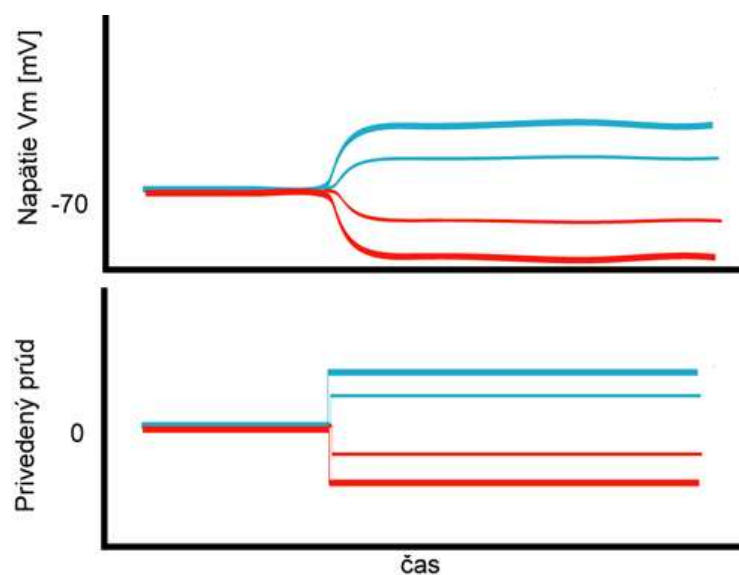
Obr. 4-2 Preparát pre štúdium elektrických vlastností neurónu.

4.1.2 Pasívny elektrický obvod ako model pokojových vlastností membrány neurónu

Teraz sa pozrime podrobnejšie na to, čím je v neuróne spôsobený pokojový potenciál. Uvažujme, že máme in vitro preparát, v ktorom môžeme sledovať pokojové

napätie neurónu. Zostava na takéto meranie je načrtnutá na Obr. 4-2. Neurón je ponorený vo výživnom roztoku, v ktorom je aj jedna z elektród. Druhá elektróda je napichnutá do vnútra neurónu. Obe elektródy sú vyvedené do zosilňovača, ktorý je schopný riadiť množstvo prúdu, pretekajúceho cez obvod. Uvažujeme, že meranie robíme v „prúdovej svorke“, t.j., že zosilňovačom vieme zafixovať, aký konštantný prúd tečie cez neurón nezávisle na tom, či neurón mení svoje elektrické vlastnosti. Zároveň meriame rozdiel v potenciáloch medzi elektródami, ktorý udáva membránové napätie V_m .

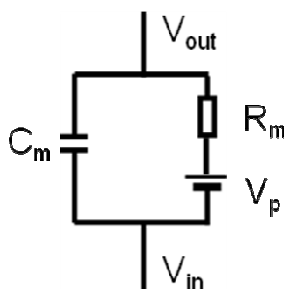
Výsledok takéhoto merania pre štyri úrovne privedeného prúdu je ilustrovaný na Obr. 4-3. Aj keď je neurón vo všeobecnosti nelineárny, ak do neho privedieme len malé množstvo prúdu, správa sa približne lineárne. Konkrétne, pre malú skokovú zmenu prúdu sa membránové napätie trocha zmení, a to tak, že sa exponenciálne približuje asymptotickej hodnote, ktorá je priamo úmerná veľkosti privedeného prúdu. Takéto elektrické správanie má obvod, ktorý pozostáva z RC člena so zdrojom napätia, ilustrovaný na Obr. 4-4. V tomto RC člene má kondenzátor náboj $Q=C_m V_m$ a tečie cez neho prúd $I=C \, dV/dt$. Rezistor má elektrické vlastnosti popísané rovnicou $V=IR_m$, a zdroj určuje pokojové membránové napätie V_p . Konštanta $R_m C_m$ sa nazýva aj časovou konštantou τ , pretože určuje začiatkový sklon krivky a tým aj ako rýchlo sa membrána dostane k svojej novej pokojovej hodnote.



Obr. 4-3 Odozva neurónu (zmena membránového napätia, obr. hore) na skokovú zmenu privedeného prúdu pri malých množstvách privedeného prúdu (obr. dole). V každom obrázku hrúbka čiar udáva zodpovedajúcu silu prúdu a napätia.

Model z Obr. 4-4 je v mnohom zjednodušený. Základný rozdiel v porovnaní s biologickým neurónom je, že v reálnom neuróne sú napätie a prúd v rôznych častiach neurónu rôzne, zatiaľ čo tu uvažujeme o jednej hodnote pre celý neurón. Typické hodnoty premenných modelu sa líšia medzi rôznymi druhmi zvierat a aj medzi neurónmi. Približne sa pohybujú v rozsahoch $V_p = -70\text{mV}$, $R_m = 10$ až $100 \text{ M}\Omega$, $C_m = 0.1$ až 1 nF , $\tau = R_m C_m = 10 - 100 \text{ ms}$.

Diskrétné elektrické prvky tohto modelu sú v membráne neurónu realizované nasledovne: kapacita C_m reprezentuje izolačné vlastnosti lipidickej dvojvrstvy membrány a odpor R_m reprezentuje priepustnosť iónových kanálov. Situácia je zložitejšia pre zdroj napätia V_p . Ten je vytvorený aktívne udržiavaným rozdielom v iónových koncentráciách medzi vnútro a vonkajškom bunky, na mechanizmus ktorého sa teraz pozrieme.



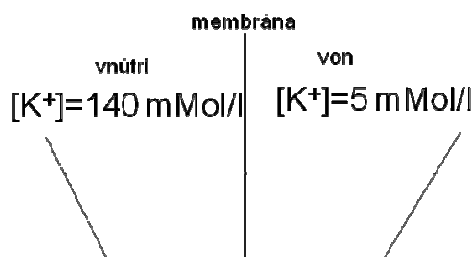
Obr. 4-4 Ekvivalentný elektrický obvod popisujúci pasívne elektrické vlastnosti neurónovej membrány.

Základom pre porozumenie zdroju napätia je poznanie, že koncentrácia rôznych iónov je rozdielna vnútri a mimo bunky. Na elektrosignalizácii sa podieľajú štyri rôzne ióny, ktorých koncentrácie sú uvedené v Tab. 1 (údaje pre neuróny cicavcov pri 37°C):

Ión	Koncentrácia (miliMol/liter)	
	Vnútoraná	Vonkajšia
Draslík K^+	140	5
Sodík Na^+	5-15	145
Chlór Cl^-	4	110
Vápnik Ca^{2+}	1-2	2-2.5

Tab. 1 Koncentrácie iónov podieľajúcich sa na signálnej funkcii neurónov.

Pretože vnútorané a vonkajšie koncentrácie sú rozdielne, potrebujeme uvažovať aj o difúzných silách, ktoré sa snažia tieto koncentrácie vyrovnávať. Pre vysvetlenie ich účinku uvažujme len o draslíku, a o modelovej situácii, v ktorej máme nádobu s dvoma oddielmi (zodpovedajúcimi vnútru a vonkajšku bunky), predelenými membránou, ktorá prepúšťa len draslík. Koncentrácie draslíku sú tak ako v Tab. 1, 140 mMol/l vnútri proti 5 mMol/l mimo bunky. Ďalej, nech v oddieloch je aj nejaký anión, ktorý vyrovnáva (neutralizuje) náboj v každej časti, ktorý ale cez membránu nemôže prejsť. Čo sa bude v experimente diať?



Obr. 4-5 Príkladový experiment pre vysvetlenie vzájomného pôsobenia difúzných a elektrických síl, ktoré spôsobujú vytvorenie náboja na membráne neurónu.

Na začiatku sú elektrické sily vyvážené, keďže v oboch častiach nádoby je rovnaké množstvo katiónov a aniónov. Difúzna sila ale spôsobí, že sa koncentrácie K^+ začnú vyrovnávať, čiže počet katiónov draslíka v pravej časti sa bude zvyšovať, zatiaľ čo množstvo kladného náboja v ľavej časti sa bude znižovať. Tým sa ale membrána polarizuje, pretože na jej pravej strane bude viac pozitívneho náboja než na jej ľavej strane. V dôsledku polarizácie začnú pôsobiť elektrické sily, ktoré budú nútiť draslík vrátiť sa do ľavého oddielu, a to dovtedy, kým sa nedosiahne stabilný bod, pri ktorom sa elektrické a difúzne sily navzájom vyrušia. Keďže na to, aby sa tento náboj na membráne vytvoril, potrebovalo cez ňu prejsť relatívne málo katiónov, tento stabilný stav sa dosiahne bez zmeny koncentrácií iónov. Na membráne sa ale vytvorí napätie, ktorého veľkosť popisuje Nernstova rovnica:

$$E = V_{dnu} - V_{von} = \frac{RT}{zF} \ln \frac{C_{von}}{C_{dnu}}$$

v ktorej T je absolútna teplota, R je plynová konštanta, F je Faradayova konštanta a z je valencia. Na základe tejto rovnice vieme určiť Nernstovo napätie pre všetky ióny. Konkrétne hodnoty sú uvedené v Tab. 2.

Ión	Koncentrácia (miliMol/liter)		Nernstovo napätie(mV)
	Vnútoraná	Vonkajšia	
Draslík K^+	140	5	-89.7
Sodík Na^+	5-15	145	90.7 – 61.1
Chlór Cl^-	4	110	-89
Vápnik Ca^{2+}	1-2	2-2.5	136 - 145

Tab. 2 Nernstovo napätie spôsobené na neurónovej membráne rozdielmi v koncentracii iónov

Keďže tieto ióny nepôsobia na membráne samostatne, potrebujeme určiť, ako Nernstove napätia jednotlivých iónov prispievajú k membránovému napätiu. To popisuje Goldmanova rovnica:

$$E = V_{in} - V_{out} = \frac{RT}{F} \ln \frac{P_K C_{K^+out} + P_{Na} C_{Na^+out} + P_{Cl} C_{Cl^-in}}{P_K C_{K^+in} + P_{Na} C_{Na^+in} + P_{Cl} C_{Cl^-out}}$$

ktorá definuje pokojový potenciál, ak je membrána priepustná pre viacej rôznych iónov. V tejto rovnici predstavuje P permeabilitu membrány pre daný ión. Typický pomer pre neurónovú membránu je $P_K : P_{Na} : P_{Cl} = 1 : 0.04 : 0.45$. A po zadaní týchto hodnôt, získame pokojový membránový potenciál E rovný -70 mV, t.j. rovný membránovému potenciálu neurónu, a zároveň napätiu V_p modelového zdroja napätia v ekvivalentnom elektrickom obvode (Obr. 4-4).

4.1.3 Akčný potenciál

Priebeh akčného potenciálu (AP, tiež nazývaného vygenerovanie impulzu, alebo pálenie neurónu) je načrtnutý na Obr. 4-1. Ako bolo spomenuté pri popise obrázku, impulzy z presynaptických neurónov spôsobujú **depolarizáciu membrány** (synaptický potenciál). Ak je táto depolarizácia dostatočná, čiže ak membránové napätie prekročí **prah/spúšť'aciú úroveň** (-50 mV), neurón **vygeneruje impulz** rôznej sily. Nasleduje **repolarizácia** (návrat k pokojovému potenciálu) a **hyperpolarizácia** (dočasný pokles potenciálu pod pokojovú hodnotu). Od vygenerovania impulzu až po návrat k pokojovej hodnote po hyperpolarizácii je neurón refraktérny (je preň ťažšie vygenerovať nový impulz). Toto **refraktérne obdobie** je na začiatku (do 1/3 repolarizácie) **absolútne**, nový impulz sa nevygeneruje v žiadnom prípade. Zvyšok času do ukončenia priebehu akčného potenciálu (do konca hyperpolarizácie) sa nazýva **relatívna refraktérna doba**, keďže v tomto čase neurón môže odpáliť, pokiaľ je stimulovaný dostatočne silne (silnejšie, než v kľudovom stave). Posledným dôležitým konceptom pri akčnom potenciáli je **latencia**, opozdenie medzi časom, keď sa vygeneruje impulz a časom, keď sa objaví podnet (keď neurón začne byť stimulovaný). Táto hodnota sa mení v závislosti na sile vstupného stimulu.

Čo a ako túto zložitú dynamiku spôsobuje? Odpoveď je čiastočne aj odpoveďou na otázku, prečo sú pre signálnu funkciu neurónu potrebné až 4 ióny, ako je uvedené v Tab. 1. Pre vygenerovanie akčného potenciálu sú dôležité katióny Na^+ a K^+ , ktoré behom potenciálu vtekajú a vytekajú z neurónu vďaka tomu, že sa pri akčnom potenciáli otvárajú a zatvárajú **napät'ovo riadené iónové kanály**, prepúšť'ajúce tieto dva ióny. Keď sa pri depolarizácii prekročí spúšť'ací prah, začnú sa rýchlo otvárať Na^+ kanály, čím sa depolarizácia ďalej zrýchľuje (vytvára sa pozitívna spätná väzba). A to až do doby, kým sa nezačnú otvárať aj pomalšie K^+ kanály, ktoré spôsobia, že sa membránové napätie začne zas znižovať. Zároveň sa začnú spontánne zatvárať Na^+ kanály, čím sa neurón repolarizuje. Keďže K^+ kanály sa zatvárajú len pomaly, vzniká na konci akčného potenciálu hyperpolarizácia, a až keď sa zavrú aj K^+ kanály, dostane sa neurón zas do pokojového stavu. Tým sa ukončí jeden cyklus vygenerovania impulzu. Ak je ale membrána presynapticky stimulovaná aj po ukončení akčného potenciálu, vygeneruje sa nový AP, atď.

Popísaná dynamika akčného potenciálu je zložitá a je ťažké vysvetliť, prečo takto komplikovaný a nelineárny systém je najvhodnejším základným signálnym správaním sa neurónu. Navyše, akčný potenciál má veľa charakteristík, ktoré sa menia v závislosti na podnete, a to frekvenciu impulzov, latenciu, výšku impulzov, či relatívnu veľkosť impulzov. Všetky tieto charakteristiky môžu kódovať relevantné informácie, a ako sme uviedli v predošlej kapitole, pre rôzne skupiny neurónov môže byť dôležitý iný aspekt pálenia.

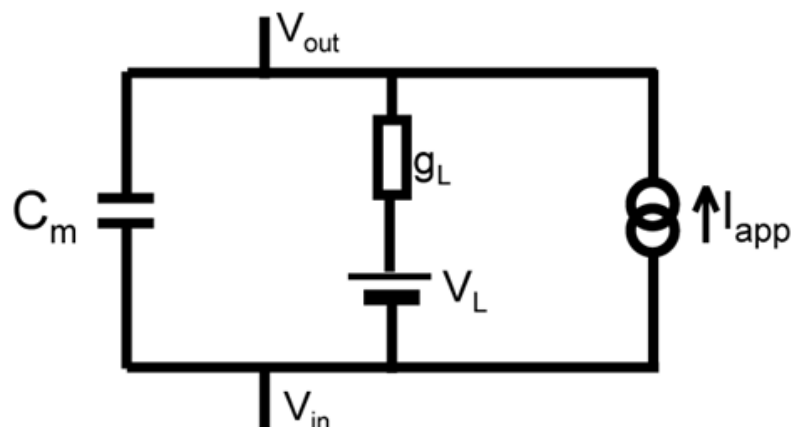
4.2 Modely aktivácie neurónu

Z hľadiska výpočtových potrieb neurálnych modelov nie je vždy nutné, aby modelový neurón mal všetky charakteristiky zhodné s veľmi komplikovanou dynamikou reálneho neurónu, tak ako sme ju popísali. Vo všeobecnosti sa vo výpočtovej neurovede používajú modely neurónov, ktoré majú rôznu stupeň komplexnosti a tým aj zhody s biologickým neurónom. Tu popíšeme 4 takéto modely. Najjednoduchší je **perceptrón** (ekvivalentný s McCulloch-Pittsovým neurónom). Na vyššej úrovni je to potom **Grossbergov model** (popísaný v nasledujúcej kapitole), aktivita ktorého je definovaná v spojitom čase. Keďže zámerom tejto kapitoly je popísať skutočný neurón, predošlé dva modely si len zmienime a zameriame sa len na zvyšné dva stupne. Prvým je **Integrate-and-fire** model, ktorý má kontinuálnu aktiváciu a zároveň generuje aj akčné potenciály. A na najvyššej úrovni je to potom **Hodgkin-Huxlyho** detailný model neurónu, ktorý popisuje presne dynamiku akčných potenciálov, ako sme si ju ilustrovali na Obr. 4-1.

4.2.1 Štandardný konekcionistický model neurónu a Grossbergov model neurónu

Štandardné neurónové siete používajú konekcionistický model neurónu, popísaný v nasledujúcej kapitole. Tento model, navrhnutý McCullochom a Pittsom (McCulloch and Pitts 1943), predpokladá jednoduchú konštantnú závislosť aktivity neurónu na vstupe. Uvažuje, že v diskretnom čase aktivita neurónu je funkciou váženej sumy vstupných aktivít z presynaptických neurónov. Vôbec sa teda nesnaží postihnúť komplexitu akčného potenciálu, ani časopriestorové závislosti šírenia signálov v neuróne a v synapsiách. Typicky je aktivačná funkcia sigmoidálna, skoková, lineárna, alebo po častiach lineárna. Často sa uvádza, že číselná hodnota aktivácie takéhoto typu neurónov zodpovedá frekvencii impulzov reálneho neurónu, čo nemusí byť zlé priblíženie, ako sme videli v predošlej kapitole.

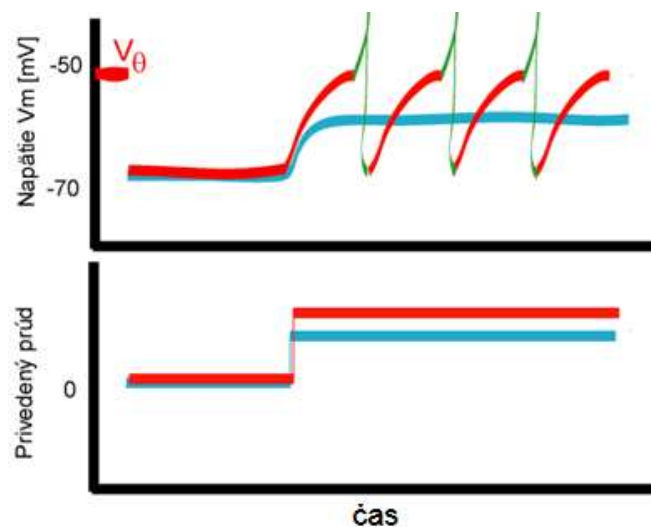
Grossbergov konekcionistický model, popísaný v nasledovnej kapitole, sa líši od najjednoduchšieho modelu vlastne len tým, že popis aktivity je dynamický, t.j. je definovaný diferenciálnymi rovnicami v čase. Tento neurón v čase integruje excitačné a inhibičné podnety zo vstupov a bez vstupov jeho aktivita klesá k nule. Stále ale ešte nemá žiaden ekvivalent akčného potenciálu, a je výrazne zjednodušený aj v iných aspektoch.



Obr. 4-6 Schéma Integrate-and-fire modelu aktivácie neurónu.

4.2.2 Integrate-and-fire model neurónu

Integrate-and-fire (integruj-a-pál') model neurónu vychádza z ekvivalentného elektrického obvodu pre popis membrány, ktorý sme uviedli na Obr. 4-4, a ktorý hovorí, že pre množstvo prúdu tečúceho cez membránu platí $I = -g_L(V - V_L)$. V tomto vzorci sme membránové napätie V_m z predošlého výkladu nahradili napätím na vetve V_L , a odpor R_m sme nahradili vodivosťou vo vetve g_L . K tomuto obvodu pridávame zdroj prúdu, ktorý reprezentuje vstupný signál, a výsledkom je model uvedený na Obr. 4-6, ktorý popisuje rovnica $C_m dV/dt = -g_L(V - V_L) + I_{app}$. Ďalej tento model zavádza prahovú hodnotu napätia, V_θ , prekročenie ktorej spôsobí vygenerovanie impulzu (viď Obr. 4-7).



Obr. 4-7 Príklad aktivácie Integrate-and-fire neurónu pri prezentovaní podprahového stimulu (modrá) a stimulu, ktorý vyvolá prekročenie prahu (červená). Dole: privedený vstupný prúd. Hore: Odozva neurónu.

Integrate-and-fire model ale nepopisuje explicitne to, ako sa impulz vygeneruje (impulzy v obrázku vyznačené zelenou si vlastne len domýšľame). Jednoducho predpokladá, že impulz sa vygeneruje a po jeho vygenerovaní sa neurón vráti to východzieho stavu (t.j. že nastane nejaký reset).

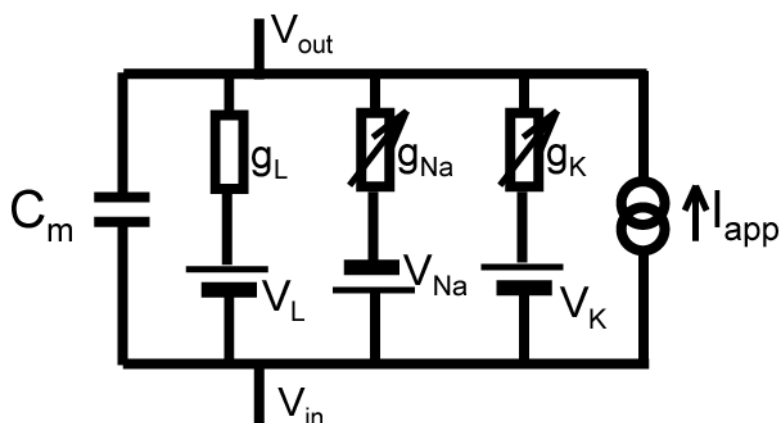
Tento model správne postihuje niekoľko charakteristík biologického neurónu. V prvom rade je to existencia prahu pre vznik impulzu. Ďalej, samotný fakt, že generuje impulzy (aj keď len virtuálne), a že frekvencia impulzov je priamo úmerná privedenému prúdu (viď Obr. 4-8). Základné obmedzenie Integrate-and-fire modelu je, že frekvencia pálenia by sa zvyšovala do nekonečna (aj keď reálny neurón má obmedzenú maximálnu frekvenciu pálenia). Táto limitácia je daná hlavne tým, že tento neurón nijak nemodeluje refraktérne doby, a teda, v momente odpálenia je pripravený páliť znovu. Ďalším obmedzením tohto modelu je, že keďže v skutočnosti negeneruje impulzy, nie je možné ním modelovať tie neuróny, ktoré výšku (príp. šírku) impulzov používajú na zakódovanie informácie. Aj napriek týmto obmedzeniam, je toto najjednoduchší model neurónu, ktorý generuje impulzy. A preto sa často používa napr. pri simulácii veľkých sietí reálnych neurónov.



Obr. 4-8 Frekvencia pálenia Integrate-and-fire neurónu pri postupnom zvyšovaní sily vstupného stimulu (prúdu).

4.2.3 Hodgkin-Huxleyho model neurónu

Hodgkin a Huxley (Hodgkin and Huxley 1952) v 40-50. rokoch 20. storočia študovali šírenie vzruchov v axóne druhu sépie (anglicky squid giant axon), pre ktorú je charakteristické, že jej axón je veľmi veľký (0.5 mm x niekoľko cm). To, že táto sépia má takýto veľký axón, je evolučným riešením zabezpečujúcim, že sa cez axón môžu rýchlo šíriť vzruchy, a teda, že v prípade potreby môže sépia rýchlo uniknúť pred predátormi. Hodgkin a Huxley matematicky detailne popísali podmienky vzniku akčného potenciálu. A až na výnimky je tento model platný pre väčšinu neurónov živých organizmov, takže je to jeden z najdôležitejších výsledkov modelovania v neurovedách. (Ďalším dôležitým počinom Hodgkina a Huxleyho bolo, že vyvinuli techniku “voltage clamp”, napäťová svorka, ktorou boli schopní udržiavať membránové napätie konštantné. My sme v minulých častiach uvažovali, že sledujeme správanie sa membrány v prúdovej svorke.)



Obr. 4-9 Schéma Hodgkin-Huxleyho modelu aktivácie neurónu.

Pri odvodení modelu Hodgkin a Huxley vychádzali z pasívneho modelu membrány, ktorý upravili tak ako v Integrate-and-Fire modeli (Obr. 4-6) pridaním zdroju prúdu. Navyše pridali dve samostatné vetvy, jednu pre K^+ a jednu pre Na^+ iónové kanály (viď. Obr. 4-9). Vodivosť týchto kanálov je závislá na membránovom napätí tak, ako je to aj pri skutočnom akčnom potenciáli. Všetky ostatné pasívne

kanály (vrátane toho pre Cl⁻) sú vo vetve "L". V pokojovom stave tečie cez jednotlivé vetvy prúd, preto model obsahuje aj zdroje, ktoré predstavujú iónové pumpy.

Matematicky sa obvod z Obr. 4-9 dá popísať jednoducho:

$$C_m \frac{dV}{dt} = -g_K(V - V_K) - g_{Na}(V - V_{Na}) - g_L(V - V_L) + I_{app}$$

Hlavným príspevkom Hodgkina a Huxleyho bolo, že definovali dynamiku otvárania a zatvárania iónových kanálov pre K⁺ a Na⁺. Na to slúžia nasledovné dve rovnice:

$$g_K = \bar{g}_K n^4$$

$$g_{Na} = \bar{g}_{Na} m^3 h$$

Z nich prvá popisuje dynamiku draslíkových a druhá dynamiku sodíkových kanálov. Premenné m, n, h sú bránové (gating) premenné. Tieto premenné určujú závislosť vodivosti kanálov na membránovom prúde V a sú definované ako

$$\frac{dm}{dt} = \alpha_m(V)(1-m) - \beta_m(V)m$$

$$\frac{dh}{dt} = \alpha_h(V)(1-h) - \beta_h(V)h$$

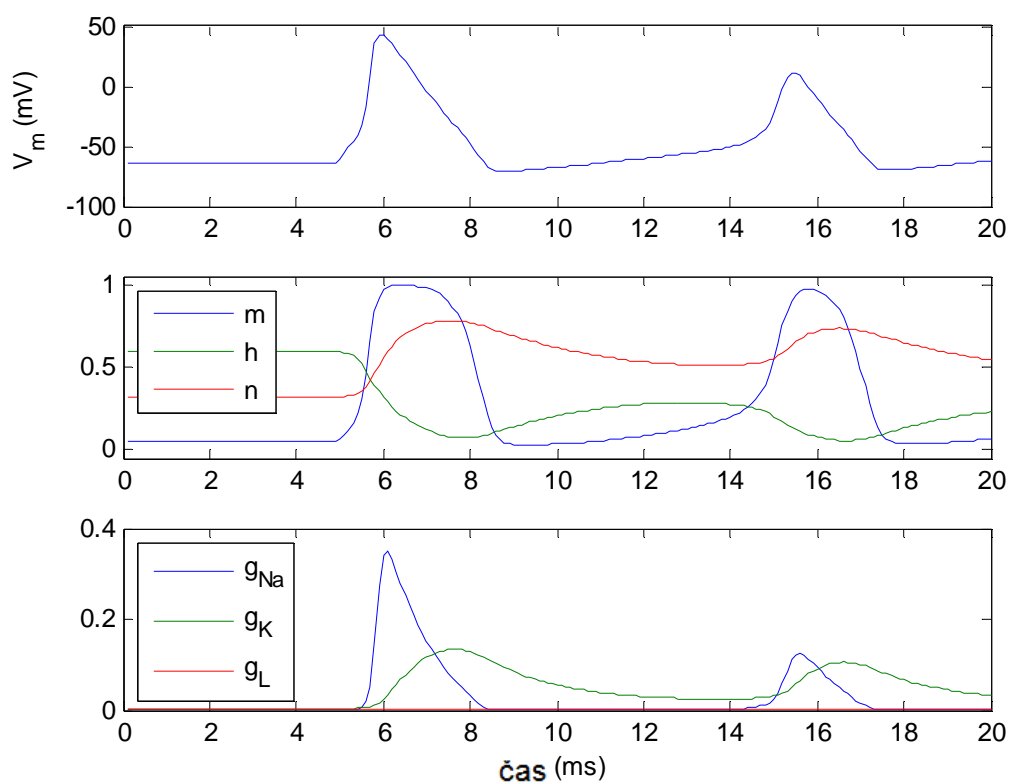
$$\frac{dn}{dt} = \alpha_n(V)(1-n) - \beta_n(V)n$$

$$\alpha_n = \frac{.01(V + 55)}{1 - \exp(-.1(V + 55))} \quad \beta_n = 0.125 \exp(-0.0125(V + 65))$$

$$\alpha_m = \frac{.1(V + 40)}{1 - \exp(-.1(V + 40))} \quad \beta_m = 4 \exp(-.0556(V + 65))$$

$$\alpha_h = .07 \exp(-.05(V + 65)) \quad \beta_h = 1/(1 + \exp(-.1(V + 35)))$$

Hodgkin-Huxleyho model je sústavou diferenciálnych rovníc 4. rádu. Zvyčajne je cieľom modelovania navrhnuť čím jednoduchší model, ktorý by bol schopný popísať čím zložitejšie dáta. Tu máme model, ktorý je natoľko zložitý, že porozumieť samotnému modelu je náročné (napr., pri pohľade na model nie je vôbec jednoduché odpovedať na otázku, ako a kde je v ňom implementovaný prah pálenia, aj keď tento model prah pálenia má).



Obr. 4-10 Príklad dynamiky aktivácie Hodgkin-Huxleyho neurónu pri privedení skokového prúdu do neurónu v čase 5 ms. Hore: odpoveď neurónu. Stred: dynamika otvárania a zatvárania sa váhových premenných. Dole: priebeh zmien vodivosti iónových kanálov.

Napriek obmedzenej možnosti analytického vniknutia do dynamiky Hodgkin-Huxleyho modelu, je tento model možné analyzovať simulovaním jeho správania pre rôzne hodnoty parametrov. Príklad takejto simulácie je na Obr. 4-10, ktorý ukazuje časový priebeh generovania akčných potenciálov v odozve na privedenie skokového vstupného prúdu, dostatočne silného na to, aby membránové napätie neurónu prekročilo prah pálenia. V hornom paneli obrázku je vykreslený priebeh odpovede neurónu. Vidíme, že táto odpoveď má všetky charakteristiky, ktoré sme uviedli v teoretickom popise akčného potenciálu: prah pálenia, depolarizácia, repolarizácia, hyperpolarizácia, latencia, zmena výšky impulzov, opakované pálenie po dosiahnutí pokojového potenciálu atď. Podobne, dolný panel ukazuje, ako sa otvárajú a zatvárajú sodíkové a draslíkové kanály. Opäť vidíme, že v súlade s teoretickým popisom sa sodíkový kanál pri generovaní impulzu rýchlo otvorí a potom samovoľne začne zatvárať. Draslíkový kanál sa otvára oveľa pomalšie a jeho zatváranie je spôsobené samovoľne vlastne len tým, že dôsledkom jeho otvorenia je hyperpolarizácia, ktorá spôsobí zavretie tohto elektricky riadeného kanálu.

Pri pohľade na Hodgkin-Huxleyho model vyvstáva množstvo otázok. Napr., prečo je bránová premenná m umocnená na 3-tiu a n na 4-tú? Hodgkin a Huxley sledovali tvar, akým sa mení g_K a g_{Na} v napäťovej slučke, čiže pri udržiavaní konštantného membránového napätia. Pritom na oddelenie vplyvu jednotlivých

kanálov ponárali sépiu do roztoku, z ktorého vylúčili K^+ alebo Na^+ , čím dosiahli že správanie sa týchto kanálov neovplyvnilo celkovú vodivosť. Potom už len potrebovali matematicky popísať dynamiku zmien kanálu, ktorý nebol blokovaný. (Pozn.: dnes už nie je potrebné vylučovať z prostredia jednotlivé katióny, ako to robili H&H - poznáme blokovače, neurotoxické látky, ktoré blokujú špecifické iónové kanály. Napr. jed tetrodotoxín TTX známy z Voodoo praktík, blokuje sodík a Tetraethylammonium TEA blokuje draslík.)

4.3 Zhrnutie

Dnes je známy relatívne detailný popis elektrických vlastností dráždivej membrány neurónov. Správanie sa neurónu je ale veľmi komplexné, a to aj v tom najjednoduchšom prípade bodového neurónu, o ktorom sme uvažovali v tejto kapitole. Tento model sa nazýva bodový, pretože sme v popise vôbec neuvažovali o tvare neurónu ani o elektrických a časopriestorových charakteristikách šírenia sa signálov v neuróne. Na takéto morfológicky detailnejší popis neurónov sa používa kombinácia Hodgkin-Huxleyho modelu s tzv. kompartmentálnym popisom anatómie neurónu, ktorým môžeme popísať všetky dôležité elektrické vlastnosti akéhokoľvek neurónu. Ďalším veľmi komplikovaným procesom, ktorý táto kapitola neobsahuje, je elektrochemický proces synaptického prenosu vzruchov medzi neurónmi. Tieto dôležité témy prekračujú zameranie predloženej učebnice. V doplnkových materiáloch na webstránke knihy je viac informácií o týchto témach, ako aj MATLABové skripty pre simulovanie neurónu, a zadanie pre cvičenie na tému Hodgkin-Huxleyho model.

Kapitola 5 Konekcionistické modelovanie

5.1 Neurovedecká inšpirácia konekcionizmu

Ako sme už spomenuli, základom konekcionistického modelovania je predpoklad, že neurálne spracovanie informácií a všetky kognitívne mozgové funkcie sa dejú v sieťach neurónov. V ľudskom mozgu je rádovo desať miliárd neurónov, pričom každý neurón je v priemere spojený s tisícom ďalších neurónov. Túto komplexnosť jednotlivých neurónov ani neurónových sietí dnes nie je jednoduché zmysluplne postihnúť. A tak sa používajú konekcionistické modely, ktoré sú zásadne zjednodušené, a ktoré s reálnymi neurálnymi štruktúrami zdieľajú dve základné charakteristiky:

- používajú jednoduchú výpočtovú jednotku, nazývanú **neurón**,
- základom komplexných výpočtov je **distribúovaný paralelizmus** neurónovej siete pozostávajúcej z množstva neurónov.

Pre popis rozdielov medzi biologickým a konekcionistickým neurónom sa zameriame na McCulloch-Pittsov (MCP) neurón, ktorý je aj dnes základom mnohých modelov. Nasledujúce zoznamy zhŕňajú základné charakteristiky biologických a MCP neurónov.

Biologický neurón

- je ich veľa a je medzi nimi veľa spojení,
- sú ich rôzne typy, napr. unipolárne, bipolárne, pyramídálne, atď.,
- podnety z okolia sa časopriestorovo integrujú v dendritickom strome, čo je silne nelineárna operácia,
- pri prekročení prahu v tele neurónu (some) vygeneruje neurón impulz (AP alebo pálenie, anglicky aj spike),
- impulz sa šíri axónom a synapsiami k ďalším neurónom,
- má refraktérne doby, t.j. po určitý čas po vygenerovaní impulzu sa nový impulz nemôže vygenerovať, a
- informácia je kódovaná rôzne v rôznych neurónoch a rôznych situáciách, napr. výškou impulzov, frekvenciou impulzov, časovaním impulzov, atď.

McCulloch-Pittsov neurón

- Výstupom je prahovaná suma vážených vstupov

$$n_i(t+1) = \Theta\left(\sum_j w_{ij} n_j(t) - \mu_i\right)$$

kde t reprezentuje diskretný čas, n_i predstavuje aktivitu sledovaného neurónu, n_j predstavuje aktivitu j -teho vstupného neurónu, w_{ij} sú váhy synaptických spojení medzi i -tým a j -tým neurónom, a μ_i je prah, ktorý musí suma prevážených vstupných aktivít prekročiť, aby sa i -tý neurón aktivoval

- pracuje v diskretnom čase, t.j. signály sa od jedného neurónu k ďalšiemu šíria v konštantných diskretných krokoch,
- je deterministický, čiže ak je vážená suma vstupov väčšia než prah, neurón sa vždy aktivuje,
- v princípe je schopný akéhokoľvek výpočtu (viď neskôr),
- informácia je kódovaná diskretné, keďže výstup každého neurónu je len 1 alebo 0.

Zo zoznamov charakteristík je jasné, že základné rozdiely medzi biologickým a McCulloch-Pittsovým (MCP) neurónom sú:

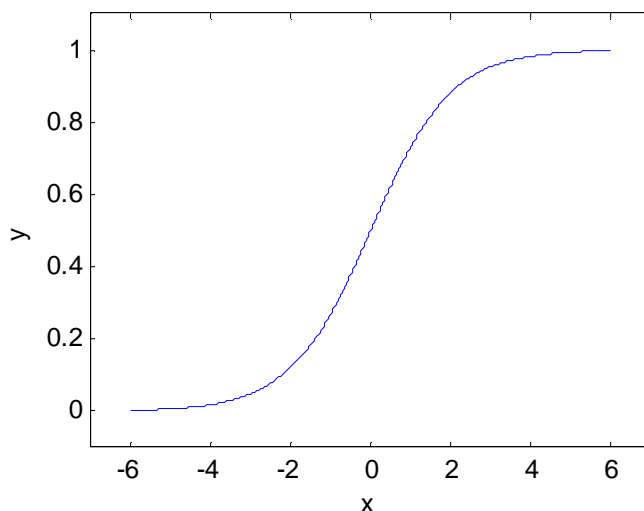
- biologický neurón má kontinuálny výstup, zatiaľ čo MCP má diskretný výstup,
- biologický neurón je silne nelineárny, zatiaľ čo v MCP je nelinearitou práve len diskretný prah,
- v biologickom neuróne sa už v dendritickom strome deje výpočet, ktorý môže byť veľmi zložitý, v MCP neuróne sa v dendritickom strome robí len lineárny vážený súčet,
- synapsia biologického neurónu je stochastická (zatiaľ čo u MCP je spojenie fixné a deterministické), čo znamená, že na základe pozorovania jedného neurónu nikdy nevieme s istotou povedať, ako sa správa nejaký iný neurón,
- v biologickom neuróne nie je jasné, aká charakteristika série impulzov charakterizuje aktivitu neurónu (najčastejšie sa uvažuje, že je to frekvencia pálenia), zatiaľ čo MCP má jasný binárny výstup,
- v MCP neuróne je časovanie synchronné, zatiaľ čo biologické neuróny sa aktivujú kontinuálne a asynchronne.

Najčastejším vylepšením MCP neurónov je náhrada prahovej aktivačnej funkcie funkciou spojitou, pričom najbežnejšia je sigmoidálna funkcia (viď Obr. 5-1). Veľa psychometrických funkcií (funkcie popisujúce vzťah medzi veľkosťou nejakého fyzikálneho podnetu a psychologickou odozvou naň, meranou napr. pravdepodobnosťou správneho rozpoznanie podnetu) aj veľa neurometrických funkcií (funkcie popisujúce závislosť medzi silou podnetu a silou neurálnej odozvy, vyjadrenou napr. frekvenciou pálenia neurónov) má práve tvar podobný sigmoide. Preto použitie sigmoidálnej funkcie v konekcionistických modeloch je významným krokom k ich priblíženiu k biologickej realite.

5.2 História konekcionistického modelovania

Historicky bol prvým krokom vo vývoji konekcionistického modelovania vyššie spomínaný článok McCullocha a Pittsa (McCulloch and Pitts 1943). Autori v ňom definovali jednoduchú neurónovú sieť, ktorá mala slúžiť na rozpoznávanie vizuálnych objektov. Základom tu bol prahový MCP neurón, ktorý podľa McCullocha a Pittsa mal byť teoreticky schopný akéhokoľvek logického výpočtu. Tento článok bol dôležitý aj tým, že v ňom autori formulovali ideu, ktorá sa neskôr hojne používala v konekcionistických sieťach, a to, že v distribuovanom systéme malých

stochastických elementov by malo byť možné vytvoriť spoľahlivú sieť, ktorá by sa skladala z nespoľahlivých elementov. Túto ideu ďalej rozpracovali napr. Winograd a Cowan (Winograd and Cowan 1963).



Obr. 5-1 Príklad sigmoidálnej funkcie definovanej ako $y = 1/(1+e^{-x})$.

V protiklade k sieťam diskretných MCP neurónov, ktoré sa snažili modelovať logické operácie nad symbolmi, sa od štyridsiatych rokov 20. storočia začali rozvíjať aj *teórie založené na kontinuálnom popise neurálnej vrstvy* (napr. neurodynamika alebo teórie neurálnych polí, angl. Neural field theories). V týchto teóriách sa neprikladá taký dôraz jednotlivým neurónom, ale neurálna štruktúra je popisovaná ako kontinuálna vrstva neurónov, ktorú môžeme chápať ako mapu, ktorej aktivácia odráža transformáciu vstupných aktivít na základe nejakého výpočtu. Tieto teórie boli motivované fyzikálnymi odbormi ako mechanika, ktorých matematický aparát aj používali. Čiže, aktivita neurálnej mapy bola často popísaná podobne, ako sa popisujú dynamické vlastnosti hmoty. Príklady prvých takýchto štúdií zhrnul napr. Rashevsky (Rashevsky 1938).

V šesťdesiatych rokoch sa výskum zamerl na neuróny podobné tým od MCP, nazývané *perceptróny*. Hlavným predstaviteľom tohto výskumu bol Frank Rosenblatt (Rosenblatt 1958; Rosenblatt 1962). Posun proti MCP bol hlavne v tom, že zatiaľ čo v MCP neuróne boli váhy spojení medzi neurónmi nastavené fixne, výskum perceptrónov bol zameraný práve na to, ako stabilne nastaviť váhy spojení tak, aby neurónová sieť nadobudla určité správanie. Výsledkom práce na perceptróne bolo vytvorenie prvých učiacich algoritmov, postupov pre iteratívne upravovanie váh, ktoré práve k takémuto výslednému nastaveniu váh viedli.

Po desaťročiach rýchleho vývoja koneccionistických modelov nastal zlom v roku 1969, kedy Minsky a Papert publikovali o perceptrónoch kritickú knihu (Minsky and Papert 1969). V nej poukázali na obmedzené schopnosti perceptrónov, a to hlavne na ich neschopnosť naučiť sa určité nelineárne problémy, pokiaľ uvažujeme o sieti perceptrónov s jednou vrstvou. Najjednoduchším príkladom problému, ktorý perceptrón nedokáže vyriešiť, je logická operácia „exclusive or“ (XOR), pri ktorej je pre dva binárne vstupy definovaný binárny výstup, a to taký, že

výstup je „1“ len v prípade, že vstupy sú rôzne „0“ a „1“ (alebo „1“ a „0“), ale výstup je „0“ v prípade že sú vstupy rovnaké „1“ a „1“ (alebo „0“ a „0“). Keďže jediným riešením pre takéto problémy by bola sieť s viacerými vrstvami perceptrónov, a keďže pre takéto siete neexistoval učiaci algoritmus, ktorý by vedel váhy spojení medzi vrstvami vhodne nastaviť, Minsky a Papert došli k záveru, že konekcionistické modely nikdy nebudú použiteľné pre zložitejšie problémy. Výsledkom publikácie tejto knihy bolo, že vývoj neurónových sietí sa na nasledujúcich 20 rokov prakticky zastavil.

Pomalé oživovanie konekcionistického modelovania nastalo koncom sedemdesiatych a v priebehu osemdesiatych rokov. Hlavný predmet štúdia v tom čase bol vývoj tzv. asociačnej obsahom adresovanej pamäti (angl. „content-addressable memory“; viď napr. (Grossberg 1967; Marr 1969; von der Malsburg 1973; Kohonen, Lehtio et al. 1974)). Základná idea obsahom adresovateľnej pamäti je, že na vyvolanie nejakej informácie z pamäti nepotrebujeme poznať jej adresu, tak ako je to u počítačov, ale že prezentovaním danej informácie sa priamo aktivuje tá oblasť pamäti, ktorá ju reprezentuje. Čiže napr., na to, aby sme rozpoznali štvorec medzi geometrickými útvarmi, nebudeme sekvenčne prehľadávať a porovnávať náš obrázok štvorca s nejakými vzormi, ktoré máme uložené v adresovateľnej pamäti, ale prezentácia štvorca nám priamo aktivuje tú oblasť, v ktorej informácie o štvorci uložené sú.

Ďalším novým konceptom bolo zavedenie energetickej funkcie na sledovanie správania sa siete a na riadenie procesu učenia sa siete (Hopfield 1982). Hlavným prínosom Hopfielda bolo, že s konceptom energetickej stability dokázal, že o pamäti môžeme uvažovať ako o poli atraktorov, ktoré sú stabilné v dynamickom poli nestabilných neurónov.

Ďalším významným krokom bolo predefinovanie McCulloch-Pittsovhovho neurónu použitím pojmov štatistickej fyziky (v ktorej sa napr. sleduje a popisuje orientácia spinov atómov feromagnetických materiálov v magnetickom poli). To umožnilo sledovanie dynamických vlastností neurónových sietí použitím analytických nástrojov vyvinutých pre štatistickú mechaniku (Caianiello 1961). A výsledkom boli modely, ktorých učenie malo charakteristicky asociatívneho Hebbovho učenia, známeho z psychológie.

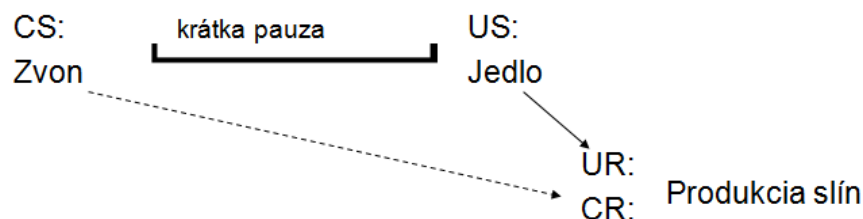
Ale najvýraznejší zlom nastal, keď nezávisle od seba niekoľko výskumníkov, napr. (Rumelhart, McClelland et al. 1986; Werbos 1994) a ďalší, navrhlo metódu učenia perceptrónov s použitím spätného šírenia chyby (známu jednoducho ako *backpropagation*), ktorá umožnila zmysluplný návrh a realizáciu viacvrstvových sietí perceptrónov (angl. multi-layer perceptron). Tieto konekcionistické siete sa vedeli naučiť aj komplexné funkcie (napr. aj XOR problém). A aj keď to ešte stále neboli úplne všeobecné siete, ktoré by vedeli aproximovať akúkoľvek funkciu, pre svoju jednoduchosť a relatívnu všeobecnosť sa stali extrémne populárnymi. Do značnej miery spôsobili obnovený vedecký záujem o konekcionistické modelovanie, ktorý, ako sme už spomínali, sa dramaticky znížil po publikovaní knižky Minskeho a Paperta. [Pozn.: V reakcii na túto obnovenú popularitu, a na žiadosti mnohých výskumníkov, sa v r. 1988 Minsky a Papert rozhodli vydať novú edíciu svojej knižky (Minsky and Papert 1988). V nej ale došli k záveru, že multi-layer perceptron a backpropagation nepriniesli žiadnu zásadnú zmenu v konekcionistickom

modelovaní, a reedícia bola len okomentovaná verziou ich pôvodnej knihy z r. 1969].

Od osemdesiatych rokov bolo publikovaných mnoho variantov neurónových sietí a konekcionistických modelov. Užitočné zhrnutie tých najzákladnejších článkov je k dispozícii v prehľadovej knižke Andersona a Rosenfelda (Anderson, Rosenfeld et al. 1988). A pre záujemcov o históriu, neformálny pohľad na konekcionistické modelovanie prezentovaný formou rozhovorov so samotnými výskumníkmi, je k dispozícii v ďalšej Andersonovej a Rosenfeldovej knihe (Anderson and Rosenfeld 1998). Užitočný úvod do matematických základov konekcionistického modelovania poskytuje (Hertz, Krogh et al. 1991).

5.3 Konekcionistické modelovanie a Teória adaptívnej rezonancie (ART)

Z veľkého množstva konekcionistických modelov, uvedených v predošlej kapitole, sa tu zameriame na modely vyvinuté Stevom Grossbergom a Gail Carpenter v Centre pre adaptívne systémy a na Katedre kognitívnych a neurálnych systémov Bostonskej univerzity. Tieto modely boli aplikované na popis širokého množstva neurokognitívnych dát, a Teória adaptívnej rezonancie, na ktorej sú založené, je popri Backpropagation jedným zo štandardov konekcionistického modelovania. Tu ilustrujeme jeden model, ktorý relatívne jednoducho dokáže vysvetliť mnohé aspekty klasického podmieneného reflexu.



Obr. 5-2 Časová následnosť pri vytváraní klasického podmieneného reflexu.

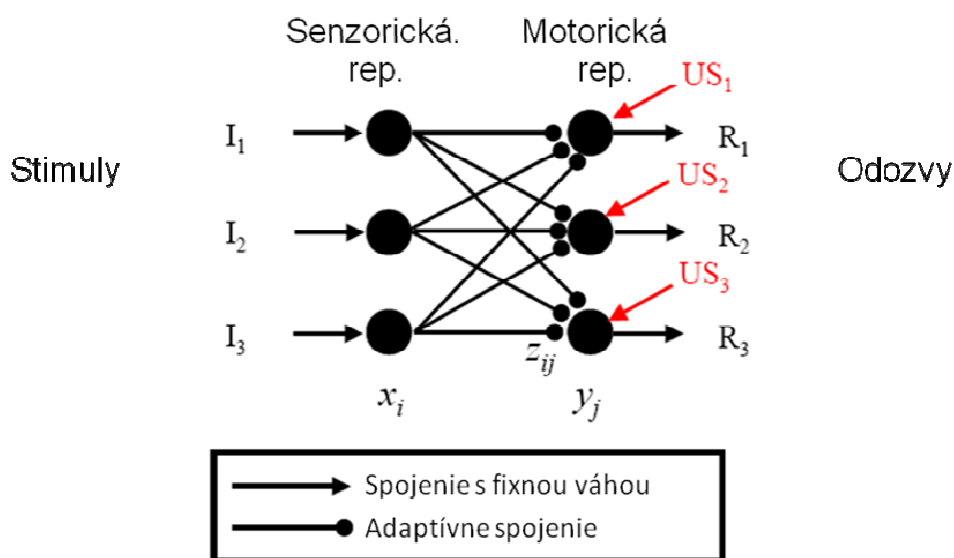
5.3.1 Interakcie medzi krátkodobou a dlhodobou pamäťou v jednoduchom neurálnom modeli klasického podmieneného reflexu

V tejto kapitole ilustrujeme vytvorenie jednoduchého dynamického modelu klasického podmieneného reflexu. Popísaný model vychádza z práce Grossberga (Grossberg 1976). Klasické podmienené reflexy (CC, z anglického *classical conditioning*), ktoré prvý popísal Pavlov (Pavlov and Anrep 1927), je základným typom asociatívneho učenia v živých organizmoch.

Pavlov pozoroval učenie, ktoré bolo výsledkom opakovanej sekvencie udalostí, ilustrovanej na Obr. 5-2. Pri pavlovovskom klasickom podmienení sa vychádza z predpokladu, že v neurálnej štruktúre organizmu už existuje asociácia medzi nepodmieneným stimulom (US) a nepodmieneným reflexom (UR). A organizmus sa musí naučiť robiť predikciu, či, a aký, nepodmienený stimul (US) bude nasledovať po prezentovaní podmieneného stimulu (CS). O schopnosti naučiť sa predikciu svedčí to, že sa objaví podmienená odozva (reflex, CR), podobná nepodmienenému reflexu (UR). Typicky sa pri klasickom podmienení dodržiava nasledovný experimentálny protokol:

- 1) **Učenie / tréning**, pri ktorom sa opakovane prezentuje podmienený stimul CS, nasledovaný nepodmieneným stimulom US, pričom medzi týmito párami stimulov sú zvyčajne dlhšie prestávky.
- 2) **Uchovanie v pamäti** sa zabezpečí tým, že necháme uplynúť dlhšie obdobie, počas ktorého sa zvierá zamestná aktivitou nesúvisiacou s danou štúdiou.
- 3) **Vyvolanie z pamäti / testovanie**, pri ktorom sa zas prezentuje podmienený stimul CS, ale bez nepodmieneného stimulu, a sleduje sa, či zvierá reaguje podmienenou odozvou CR.

Výsledkom tréningu je, že podmienená pravdepodobnosť $p(\text{CR}|\text{CS})$ (pravdepodobnosť, že na podmienený stimul zvierá odpovie podmienenou odozvou), vzrastie. Z matematického hľadiska môžeme povedať, že klasické podmienenie je nestacionárny proces, ktorého výsledkom je dlhodobá pretrvávajúca zmena, ktorú nazývame vytvorenie **dlhodobej pamäte** (vytvorenie LTM).



Obr. 5-3 Jednoduchý neurálny model klasického podmienenia

Teraz navrhujeme jednoduchý konekcionistický model, ktorý umožňuje popísať viacero aspektov klasického podmienenia. Definujeme jednoduchú neurónovú sieť, ktorá zahŕňa senzorickú a motorickú reprezentáciu (viď Obr. 5-3). Na klasickom podmienení si môžeme všimnúť, že podmieneným stimulom CS je možné podmieniť aj udalosť (US), ktorá sa uskutoční až keď CS zmizol. To znamená, že reprezentácia podnetu v neurálnej štruktúre pretrváva aj keď podnet samotný už zanikol (napr., ak zvonček už doznel). Okrem tejto vlastnosti klasického podmienenia si môžeme všimnúť ešte dve ďalšie:

1. US, ktorý nasleduje krátko po CS zvyčajne podmieni silnejšie a rýchlejšie než US nasledujúci po dlhšej pauze.
2. Ak US nasleduje súčasne s (alebo pred) CS, podmienenie sa neuskutoční.

Z behaviorálneho hľadiska sú tieto vlastnosti dané nasledovnými faktormi:

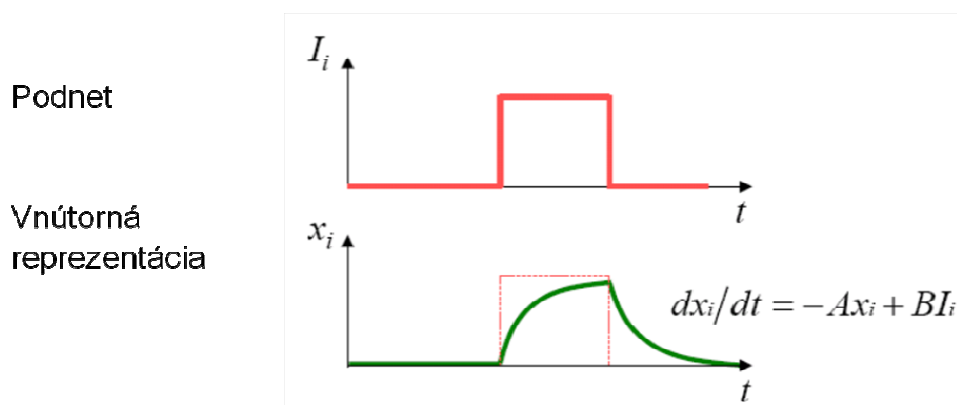
1. Podnet v reálnom svete, ktorý bezprostredne predchádza významnej udalosti je spoľahlivejším prediktorom danej udalosti, než podnet, ktorý sa objaví s väčším predstihom.

2. Podnet z okolia, ktorý sa objaví po alebo súčasne s dôležitou udalosťou, má nulovú užitočnosť ako prediktor danej udalosti.

V našom modeli hľadáme senzorickú reprezentáciu, ktorá bude mať všetky tri uvedené vlastnosti, a to, že a) reprezentácia podnetu pretrváva aj keď podnet už zanikol, b) US, ktorý nasleduje krátko po CS zvyčajne podmieňuje silnejšie a rýchlejšie než US nasledujúci po dlhšej pauze, a c) ak US nasleduje súčasne s alebo pred CS, podmienenie sa neuskutoční.

Pre vstupný signál I_i a aktivitu neurónu x_i , zabezpečí všetky tieto vlastnosti reprezentácia aktivácie neurónu popísaná rovnicou

$$dx_i/dt = -Ax_i + BI_i$$



Obr. 5-4 Priebeh reprezentácie vstupného signálu v modeli CC.

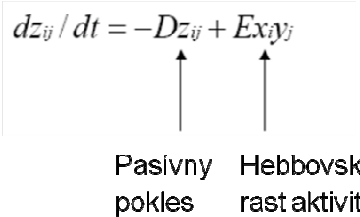
Táto diferenciálna rovnica popisuje okrem iného aj dynamiku naplňovania sa a vyprázňovania zatekajúcej nádoby (alebo zatekajúceho integrátora z anglického názvu „leaky integrator“). V tejto rovnici to, aké veľké časové opozdenie medzi CS a US vie táto sieť reprezentovať, závisí na **časovej konštante** zatekajúceho integrátora, ktorá súvisí s parametrom A . Tento parameter má ale ešte jednu, dôležitejšiu funkciu, a to, že zabezpečuje spontánny pokles aktivity v prípade, že na neurón nie je privedený žiaden vstup. Spontánny pokles aktivity je jedným z najprirodzenejších predpokladov, keď sa dívame na fungovanie nervovej sústavy: po každej excitácii sa membránový potenciál neurónov vracia do svojho pokojového stavu. Bez spontánneho pasívneho poklesu aktivity by aktivita neurónu neobmedzene rástla v prípade, že by bunka nemala žiadne inhibičné vstupy. (Pozn.: Niektoré neuróny, alebo ich zložky, čiastočne vykazujú takéto správanie, t.j. fungujú ako integrátory v prítomnosti modulátorov ako serotonín.)

Podobne, ako sme definovali dynamickú reprezentáciu stimulov, môžeme definovať aj reprezentáciu motorických odoziev y_j , ako:

$$dy_j/dt = -Ay_j + B \sum_i x_i z_{ij}$$

Táto rovnica hovorí, že vstupný signál do každého motorického neurónu závisí na súčte prevážených aktivít vstupných neurónov.

Keďže v našom modeli sa aktivity x_i a y_i menia relatívne rýchlo, o týchto premenných budeme hovoriť ako o reprezentácii krátkodobej pamäte (STM). Na vysvetlenie klasického podmieňovania potrebuje naša sieť aj váhy, ktoré by kódovali naučené asociácie medzi STM reprezentáciami senzorických podnetov a motorických odoziev. Napr., ak v modeli z Obr. 5-3 senzorický vstup I_1 je opakovane párovaný s US , ktorý spôsobuje motorickú odozvu R_2 , a žiadne iné stimuly spárované neboli, potom môžeme očakávať, že sila synaptického spojenia medzi týmito dvomi neurónmi bude vysoká. Váha, čiže sila synaptických spojení medzi jednotlivými neurónmi, je v našom modeli reprezentovaná premennou z_{ij} . Matematicky sa dá vzťah medzi vstupným a výstupným neurónom, ktorý by zabezpečil takúto požadované správanie nášho neurálneho modelu, popísať rovnicou:

$$\frac{dz_{ij}}{dt} = -Dz_{ij} + Ex_i y_j$$


Pasivny pokles Hebbovsky rast aktivity

V tejto rovnici, prvý člen reprezentuje pasívny pokles hodnoty, známy už z predošlých rovníc popisujúcich STM. Druhý člen hovorí, že sila spojenia má rásť v prípade, že obe STM aktivity sú pozitívne súčasne. To je v súlade s Hebbovým pravidlom (Hebb 1949), ktoré popisuje neurobiologickú bázu učenia, a ktoré hovorí:

“When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased.” čiže „Ak axón bunky A je dostatočne blízko k bunke B na to, aby ju aktivoval, a ak často vyvoláva jej pálenie, nastane v jednej z týchto buniek, alebo v oboch, rastový alebo metabolický proces, v ktorého dôsledku sa zvýši účinnosť, s akou je bunka A schopná vyvolať pálenie bunky B.“

Predchádzajúcou rovnicou sme definovali učiace pravidlo nášho jednoduchého neurálneho modelu. Toto pravidlo má ale jeden potenciálny problém: Ak sa asociácia pravidelne netrénuje, postupne sa stratí (čo je spôsobené pasívnym poklesom sily spojenia). V podmieňovacích experimentoch sa to ale väčšinou nestane: silne naučené asociácie pretrvávajú dlho, pokiaľ zvieru neabsolvuje extinkčný tréning, pri ktorom sa konkrétna naučená asociácia potláča. Ako môžeme upraviť naše učiace pravidlo, aby sme doň túto vlastnosť zapracovali? Odpoveďou môže byť *presynapticky bránovaný pokles*, definovaný ako

$$dz_{ij}/dt = -Dx_i z_{ij} + Ex_i y_j$$

resp.

$$dz_{ij} / dt = x_i \cdot [-Dz_{ij} + Ey_j]$$

V tejto rovnici sa už váhy znižujú len keď je prítomný CS. Názov *presynapticky bránovaný pokles* sa používa pre to, že táto rovnica vlastne obsahuje bránu, reprezentovanú presynaptickou aktiváciou, ktorá riadi, či sa učenie uskutočňuje, alebo nie. Keď $x_i=0$, brána je zavretá a daná synapsia sa neučí.

Čo by sa stalo, ak by sme použili postsynapticky bránovaný pokles? Uvažujme o učiacom pravidle:

$$dz_{ij} / dt = -Dy_j z_{ij} + Ex_i y_j$$

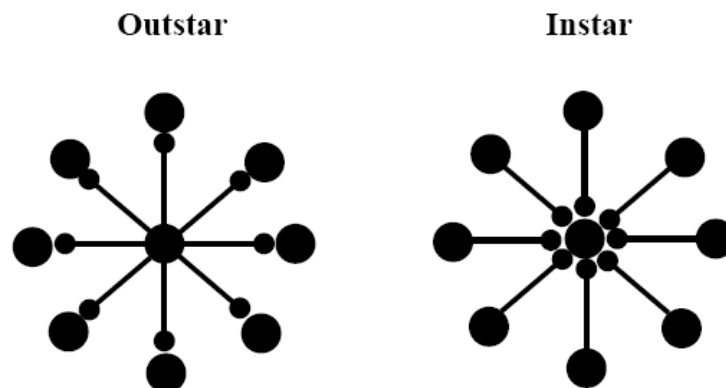
resp.

$$dz_{ij} / dt = y_j \cdot [-Dz_{ij} + Ex_i]$$

V tomto prípade prítomnosť odozvy bez CS povedie k poklesu váhy synapsie, zatiaľ čo prítomnosť CS bez prítomnosti odozvy nespôsobí žiadnu zmenu váhy.

V porovnaní s *presynapticky bránovaným poklesom* je toto správanie v menšom súlade s behaviorálnymi dátami o klasickom podmieňovaní, pretože prítomnosť odozvy bez prítomnosti CS (napr. tvorby slín u Pavlovových psov) zvyčajne nevedie k potlačenej asociácie.

Predchádzajúcimi dvomi rovnicami sme popísali dva základné typy učenia navrhnuté Grossbergom, a to učenie typu Instar a Outstar, schématicky zobrazené na Obr. 5-5. Pri učení Outstar, synapsie vedúce Z daného neurónu sa menia, keď je neurón aktívny. T.j. synapsie sa naučia „pattern“ výstupných aktivácií. Pri učení Instar, synapsie vedúce DO daného neurónu sa menia, keď je neurón aktívny. T.j. synapsie sa naučia odpovedať na daný vstupný „pattern“.



Obr. 5-5 Schématické zakreslenie typickej skupiny neurónov, pre ktorú sa využíva učenie Instar a Outstar.

Matematicky sú tieto dva typy učenia rovnako platné. Ale v neurofyziologických experimentoch, sledujúcich zmeny sily synaptických väzieb, sa častejšie pozoruje postsynaptické bránovanie poklesu, čiže učenie **Instar**.

Keď nejaký vstupný stimul vedie k motorickej odozve (reakcii zvierat'a), ktorá vyvolá negatívnu spätnú väzbu z prostredia (napr. ak sa po zaznení zvončeka dotkneme elektrickej zástrčky), synaptické spojenie medzi reprezentáciou daného vstupu a odozvou (akciou), ktorú vyvolal, by sa malo zoslabiť. Tomuto procesu sa hovorí **negatívne posilňovanie** alebo **negatívny reinforcement**. Pravidlá Outstar ani Instar zatiaľ možnosť správne reagovať na takúto situáciu nemali. Možnosťou, ako ich upraviť, je rozšíriť učiace pravidlo o člen popisujúci **chybový signál**, ktorý spôsobí zmenu v správaní sa učiaceho pravidla v prípade, že aktuálna reakcia nebola správna. Čiže rovnicu pre **Outstar** môžeme rozšíriť na:

$$dz_{ij} / dt = -Dx_i z_{ij} + Ex_i y_j - F \cdot error$$

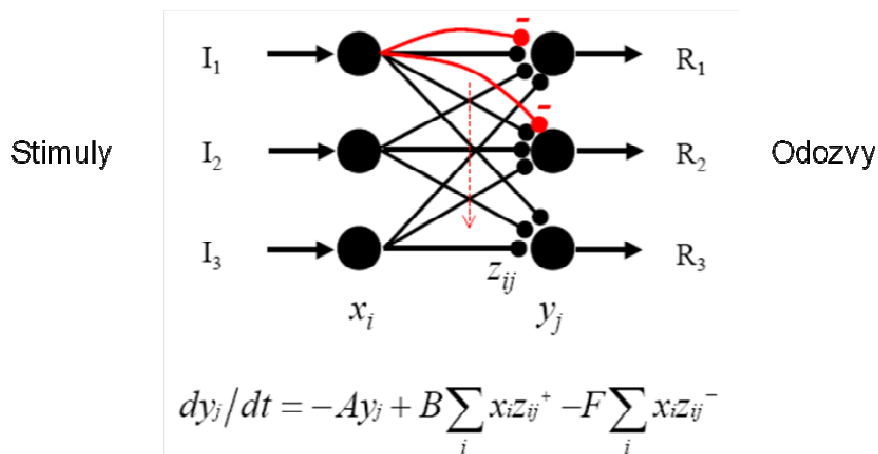
kde F je ďalší parameter a *error* predstavuje chybový signál. Táto rovnica má ale problém, že výskyt chyby by spôsobil zoslabenie všetkých asociácií, vrátane tých, ktoré sú správne a ktoré nijak nesúvisia s aktuálnou chybou. Možnosťou, ako to ošetriť je rozšíriť chybový člen nasledovne:

$$dz_{ij} / dt = -Dx_i z_{ij} + Ex_i y_j - F \cdot error \cdot x_i y_j$$

Negatívne posilňovanie vykazuje v mnohých experimentoch charakteristiku, že stačí jeden pokus s negatívnym posilnením na to, aby sa dlhodobo potlačila chybná odozva. V našom modeli sa táto vlastnosť dá dosiahnuť napr. nastavením parametra F na hodnoty výrazne vyššie než hodnoty parametrov D a E .

Príkladom neurálnej štruktúry, kde takéto negatívne posilňovanie prebieha, je mozoček (cerebellum) počas motorického učenia. Tu chybná akcia (pohyb) spôsobí aktiváciu vzostupných vlákien. Aktivácia vzostupných vlákien spôsobí vygenerovanie „komplexného pálenia“ v Purkyňovej bunke, ktorý spôsobí pokles v sile synaptického spojenia z aktívnych paralelných vlákien (dlhodobá depresia, LTD).

Toto chybové učenie pozorované v mozočku je pomalšie než nami navrhnuté učenie. Preto môže byť v našom modeli vhodnejšie uvažovať nad alternatívnym spôsobom zapracovania funkčnosti negatívneho posilňovania, a to pridaním inhibičných spojení, ktoré budú kódovať *negatívne asociácie* (viď Obr. 5-6).



Obr. 5-6 Rozšírenie modelu CC o inhibičné spojenia medzi STM neurónmi.

V tomto prípade sa môžeme vrátiť k rovniciam pre excitačné váhy bez chybového člena

$$dz_{ij}^+ / dt = -Dx_i z_{ij} + Ex_i y_j$$

No a inhibičné synapsie, ktoré by sa mali posilňovať v prípade chyby, a ktoré by zároveň mali mať schopnosť spontánne sa zoslabovať v prípade, že žiadna chyba už dlho nenastala, môžu byť popísané rovnicou

$$dz_{ij}^- / dt = -Gx_i y_j z_{ij} + H \cdot error \cdot x_i y_j$$

Týmto sme dospeli k dvom alternatívnym modelom klasického podmieňovania, jedným s inhibičnými spojeniami a jedným bez nich. Pri rozhodovaní, ktorý z nich je vhodnejší, sa môžeme pozrieť na to, či existujú mozgové štruktúry, ktoré kódujú pozitívne/negatívne posilňovanie. V niekoľkých mozgových oblastiach sa nachádzajú neuróny, ktorých aktivita reprezentuje rozdiel medzi očakávanou odmenou a skutočnou odmenou. K takýmto oblastiam patria amygdala a bazálne gangliá. Tiež Rescorla-Wagnerov (Rescorla and Wagner 1972) ako aj Sutton-Bartov (Sutton and Barto 1981) model klasického podmieňovania predpokladajú, že neuróny s takouto odozvou existujú.

Dolad'ovanie siete

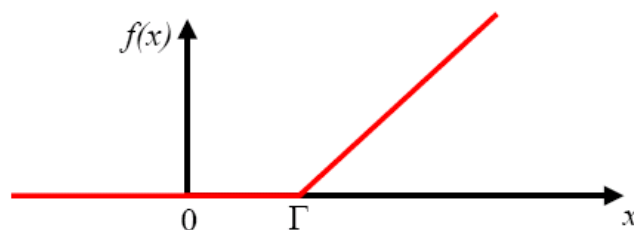
Predstavme si, že simulujeme našu sieť (tú bez inhibičných spojení) na počítači a zistíme, že pri určitej náhodnej kombinácii stimulu a odozvy sa udeje priveľká zmena. T.j. v porovnaní s behaviorálnymi dátami má prezentácia stimulu, ktorý bol len málokedy spárovaný s odozvou, prisilnú tendenciu vygenerovať odozvu. Čo môžeme spraviť s našou sieťou, aby sme to napravili?

Možné riešenia sú napr.:

- zvýšime parameter rýchlosti poklesu aktivácie (D) v učiacom pravidle:

$$dz_{ij} / dt = -Dx_i z_{ij} + Ex_i y_j - F \cdot error \cdot x_i y_j$$

- znížime parameter E ,
- motorickým neurónom priradíme *prah-lineárnu signálnu funkciu*. Miesto toho, aby sa odozva R_i rovnala aktivácii neurónu y_i , použijeme rovnicu $R_i = f(y_i)$, kde $f()$ predstavuje prah-lineárnu závislosť, graficky zobrazenú na Obr. 5-7.



Obr. 5-7 Prah-lineárna signálna funkcia

Matematicky je prah-lineárna signálna funkcia definovaná ako

$$\begin{aligned} f(x) &= 0 && \text{pre } x \leq \Gamma \\ f(x) &= x - \Gamma && \text{pre } x > \Gamma \end{aligned}$$

a v ďalšom takýto typ funkcie budeme zapisovať ako $[x - \Gamma]^+$, kde operácia $[]^+$ označuje *usmernenie* (ako dióda).

Poznámka: Prah-lineárnu funkciu by bolo možné použiť aj na popis aktivácie sensorických neurónov. Nechávame na zváženie čitateľa zamyslieť sa nad tým, aké dôsledky by to malo pre dynamiku našej siete.

Ďalšou vlastnosťou, ktorú by model klasického podmieňovania mohol mať, by bola schopnosť postihnúť *vrodené rozdiely v učení*, ktoré sa často pri klasickom podmieňovaní vyskytujú. Napr., nevoľnosť je zvyčajne spôsobená jedlom, len zriedka zvončkami alebo svetlom, a zvieratá majú tendenciu veľmi rýchlo vytvoriť negatívnu asociáciu s jedlom, ktoré nevoľnosť spôsobilo (Garciov efekt), ale nie asociáciu medzi svetlom a nevoľnosťou. Čo by sme mali zmeniť v našich rovniciach, aby to dokázali obsiahnuť? Jednou z možností je zvoliť rôzne parametre pre rôzne spojenia:

$$\begin{aligned} dy_j / dt &= -Ay_j + \sum_i B_{ij} x_i z_{ij} \\ dz_{ij} / dt &= -Dx_i z_{ij} + E_{ij} x_i y_j - F \cdot \text{error} \cdot x_i y_j \end{aligned}$$

Ak má niektorá synapsia vysoké hodnoty B_{ij} a E_{ij} , potom sa táto synapsia učí rýchlejšie (a rovnaké množstvo učenia u nej vedie k silnejšej odozve) než u ostatných synapsí. T.j. ak sa má vytvoriť asociácia medzi podnetom a odozvou zodpovedajúcim danej synapsii, táto asociácia sa vytvorí ľahšie.

Poznámka: Jednou z často diskutovaných tém v neurovede a v neurofilozofii je otázka o „nature vs. nurture“, čiže o tom, čo je vrodené a čo je naučené; vid' napr. (von Melchner, Pallas et al. 2000). Naš model vie prednastavením parametrov B_{ij} a E_{ij} popísať rôzne vzťahy medzi vrodenými a naučenými vlastnosťami.

Na záver je dôležité pripomenúť, že navrhnutý model klasického podmieňovania je veľmi jednoduchý. Preto má aj množstvo slabín, na ktoré je pri modelovaní vždy dôležité myslieť. Napr., v tomto modeli sensorické vstupy nie sú distribuované (jedna vstupná bunka reprezentuje jeden vstup), čo určite nie je v súlade s poznatkami z neurobiológie. Podobne, nie sú distribuované ani motorické odozvy. Ďalej, modelové bunky sú len čiastočne podobné neurónom v biologickom nervovom systéme, napr. negenerujú impulzy, nemajú žiaden reálny tvar, atď. Na skutočné vyhodnotenie modelu je potrebné porovnať počítačovou simuláciou jeho správanie s behaviorálnymi dátami. Pri tom napr. môžeme zistiť, že učiace krivky získané z behaviorálnych dát nemusia zodpovedať krivkám generovaným modelom. Konkrétne, jedným z charakteristických znakov psychologického učenia sa a zapamätávania si je, že množstvo zapamätaných informácií v závislosti na čase od ich prezentácie má tvar podobný písmenu U. Teda pamätáme si veľa toho, čo sme počuli na začiatku a na konci, a najmenej si zapamätáme z toho, čo zaznie niekde v strede. Je relatívne jasné, že parameter časovej konštanty nášho zatekajúceho integrátora pravdepodobne nestačí na to, aby tento model mohol vykazovať učiacu krivku tvaru U. Napriek týmto

nedostatkom ale náš model dokáže popísať množstvo jednoduchých učiacich procesov.

5.3.2 Zhrnutie

Tu popísaný model klasického podmieňovania je najjednoduchším v plejáde Grossbergových modelov neurokognitívnych procesov. Je to dopredný dynamický model, ktorý dokáže popísať len relatívne jednoduchšie interakcie. Na jeho základe boli neskôr navrhnuté zložitejšie nerekurentné siete, rekurentné siete, a nakoniec aj komplexné siete implementujúce všetky princípy teórie ART, ktoré predstavujú funkčný model mnohých aspektov učenia, vnímania, rozpoznávania, pozornostných mechanizmov atď. Časť z týchto modelov je popísaná v doplnkových materiáloch ako aj v zadaniach k predmetu Výpočtová a kognitívna neuroveda.

Kapitola 6 Teória optimálneho rozhodovania a modelovanie psychofyzikálnych dát

6.1 Psychofyzika

Na štúdium vnemových systémov človeka sa často používajú psychofyzikálne experimenty, čiže experimenty, pri ktorých sa hľadá vzťah medzi fyzikálnymi charakteristikami nejakého stimulu a psychologickými veličinami, t.j. charakteristikami vnemov, stavov vedomia, pocitov, či správania sa, ktoré tieto fyzikálne podnety vyvolali. Zakladateľom tohto vedného odboru je Gustav Fechner (Fechner 1860).

Psychofyzika sa ďalej delí podľa toho, na aký vnem sa zameriava. Napr., vedný odbor zaoberajúci sa štúdiom sluchového vnímania sa nazýva **psychoakustika** alebo **sluchová psychofyzika** (Yost 1994), a hľadá vzťahy medzi akustickými charakteristikami a psychologickými a fyziologickými odozvami vyvolanými zvukom (vrátane napr. hudby a reči). Príklady fyzikálnych a psychologických veličín, ktoré sa študujú v psychoakustike, sú uvedené v Tab. 3).

Fyzikálne	Psychologické
<i>Zvuk</i>	<i>Sluchový vnem</i>
Intenzita	Hlasitosť
Frekvencia	Výška tónu
Spektrum a časové charakteristiky	Farba zvuku (timbre)

Tab. 3 Fyzikálne veličiny a im zodpovedajúce psychologické veličiny, tak ako sa študujú v sluchovej psychofyzike.

Primárne sa psychofyzikálne metódy rozdeľujú na subjektívne a objektívne. Subjektívne metódy sa pýtajú na subjektívny vnem, zážitok, uvedomenie si, či reakciu experimentálneho subjektu, a v zásade pre ne platí, že kladú otázky, na ktoré neexistuje správna odpoveď nezávislá od experimentálneho subjektu. V psychoakustike sú takými metódami napríklad metódy pre odhadovanie hlasitosti, hľadanie tónu zhodnej výšky, hodnotenie „kvality zvuku“, atď. Pri nich sa vyžaduje introspekcia a tieto metódy majú väčšinou základný matematický nedostatok, a to že meranie je pri nich ovplyvnené nezmyslovými faktormi (napr. náladou alebo zameraním pozornosti poslucháča).

Objektívne metódy sú charakteristické jednak tým, že pre každý stimul existuje objektívne správna odpoveď (napr. ak sa pýtame na frekvenciu stimulu, nie na vnímanú výšku tónu). Typickými experimentálnymi metódami sú detekcia prítomnosti stimulu alebo diskriminácia stimulov (čiže rozlišovanie medzi dvomi stimulmi, ktoré sa líšia len v jednej charakteristike, povedzme vo svojej intenzite, a úlohou subjektu je určiť, ktorý z dvoch prezentovaných stimulov je intenzívnejší, hlasnejší, svetlejší atď.). V týchto experimentoch sa meria presnosť odpovedí a vždy existuje objektívne správna odpoveď. Naviac, objektívne psychofyzikálne metódy sa

snažia minimalizovať vplyv subjektívnych faktorov odpovedí subjektu, ako napr. toho, že keď má subjekt rozhodnúť, či je hlasnejší prvý alebo druhý tón, subjekt odpovedá, že druhý len preto, že si nie je istý, a viac sa mu pozdáva číslo 2. Takéto subjektívne faktory spôsobujú systematické chyby (angl. biasy) v experimentálnych dátach, ktoré je možné odstrániť použitím matematických metód vychádzajúcich z Teórie detekcie signálov, angl. Signal Detection Theory (Green and Swets 1974; Macmillan and Creelman 2005) a z bayesovskej Teórie optimálneho rozhodovania (Duda, Hart et al. 2001). Keďže dáta získané v psychofyzikálnych experimentoch majú určité matematické vlastnosti, je vhodné používať na nich aj modely, ktoré vychádzajú z podobných vlastností. V nasledujúcej sekcii jeden taký model popíšeme.

6.2 Model rozhodovania pre psychofyzikálne dáta

Nathaniel Durlach v roku 1968 popísal model rozhodovania, ktorý sa často používa pre popis psychofyzikálnych dát (Durlach and Braida 1969; Macmillan, Braida et al. 1989). Tento model je založený na princípoch Pravidla komparatívneho rozhodovania (Law of comparative judgment, niekedy tiež označovaného ako Law of categorical judgment), navrhnutého L. L. Thurstoneom (Thurstone and Chave 1929; Torgerson 1958), a súvisí tiež s Weber-Fechnerovým zákonom o vzťahu psychologických a fyzikálnych veličín (Masin, Zudini et al. 2009), a so Shannonovou matematickou Teóriou komunikácie (Shannon 1997).

Základné vlastnosti tohto modelu sú:

- Je pravdepodobnostný a tak zohľadňuje fakt, že subjekt (napr. poslucháč) neodpovie pri opakovanom prezentovaní toho istého stimulu vždy rovnako.
- Explicitne zahŕňa vplyvy vonkajších faktorov na rozhodovanie (napr., závislosť odpovedí na inštrukciách a na informáciách, ktoré poslucháč dostane od experimentátora, a na rôznych osobných preferenciách subjektu).
- Je možné ho použiť v širokom spektre experimentálnych paradigiem a je ním možné popísať množstvo experimentálnych výsledkov.
- Veľmi často návrh behaviorálnych experimentov priamo vychádza z tohto modelu.
- V experimentálnej psychológii a behaviorálnych štúdiách sa používa natoľko často, že niekedy je bez toho, aby sme modelu rozumeli, ťažké porozumieť publikovaným článkom.

6.2.1 Definícia modelu

Pri ďalšom popise budeme uvažovať, že model aplikujeme na sluchové behaviorálne dáta. Konkrétne model definujeme na experimentálne dáta, získané v experimentoch vychádzajúcich z paradigmy v angličtine nazývanej „one-interval, two-alternative-forced-choice“ („jeden interval, dve alternatívy a vynútená odpoveď“, skratka 1I-2AFC). V tejto paradigme pri každom pokuse v experimente môže byť subjektu stimul prezentovaný len v jednom intervale (v jednom čase), v ktorom stimul buď zaznie alebo nie, a subjekt musí odpovedať jednou z dvoch predpísaných

odpovedí (napr. „stimul zaznel“ alebo „stimul nezaznel“), a nemôže odpovedať „neviem“, alebo „nie som si istý“). Takže, uvažujeme o experimente, v ktorom:

- Môžu zaznieť dva stimuly, S_1 a S_2 .
- Pripúšťajú sa len dve odpovede, R_1 a R_2 .
- Pri každom pokuse experimentátor náhodne vyberie, či prezentuje stimul S_1 alebo S_2 , pričom pravdepodobnosti výberu stimulov sú $P(S_1)$ a $P(S_2) = 1 - P(S_1)$.
- Subjekt (poslucháč) dostal inštrukcie, že má odpovedať R_1 v prípade, že počul stimul S_1 , a R_2 v prípade, že počul stimul S_2 .
- Experimentátor „odmení“ subjekt odmenou $V(R_i | S_j)$, ak subjekt na stimul S_j zadal odpoveď R_i . Pričom $V(R_i | S_j) > 0$ (t.j. je to odmena) pre $i = j$, a $V(R_i | S_j) < 0$ (t.j. je to trest) pre $i \neq j$.

V experimentoch, ktoré sa v psychofyzike označujú ako *diskriminačné* (keď subjekt má odlíšiť dva rôzne stimuly) budú stimuly S_1 a S_2 predstavovať dva rozdielne nenulové signály. V *detekčných* experimentoch (keď sa meria, do akej miery je subjekt schopný detekovať prítomnosť nejakého stimulu), S_1 bude označovať nulový signál (keď žiaden stimul nezaznie) a S_2 bude označovať stimul obsahujúci signál, ktorý má subjekt detekovať. Pri experimentoch, v ktorých sa meria schopnosť človeka detekovať nejaký zvuk (tón, reč...) maskovaný šumom, sa cieľový zvuk zvyčajne označuje ako S a maskovací rušivý šum ako N. Potom S_1 bude označovať prezentovanie samotného šumu N a S_2 bude označovať stimul obsahujúci signál aj šum S + N.

Ak predpokladáme, že jednotlivé pokusy v našom experimente sú od seba štatisticky nezávislé, a že jediná charakteristika odpovedí, na ktorú sa dívame, je *aká* odpoveď to bola (t.j. nezaujímajú nás iné vlastnosti odpovede, ako napr. jej latencia), potom výsledky takéhoto experimentu môžeme zhrnúť do matice 2×2 , ktorej prvky f_{ij} vyjadrujú relatívne početnosti odpovedí R_i keď je prezentovaný stimul S_j :

$$f_{ij} = \frac{N(R_i | S_j)}{N(R_1 | S_j) + N(R_2 | S_j)}$$

kde $N(R_i | S_j)$ udáva koľkokrát subjekt odpovedal R_i na S_j a $N(R_1 | S_j) + N(R_2 | S_j)$ udáva koľkokrát bol stimul S_j prezentovaný. Navyše, keďže platí, že $f_{1j} + f_{2j} = 1$, je matica plne definovaná, ak poznáme akékoľvek jej dva prvky:

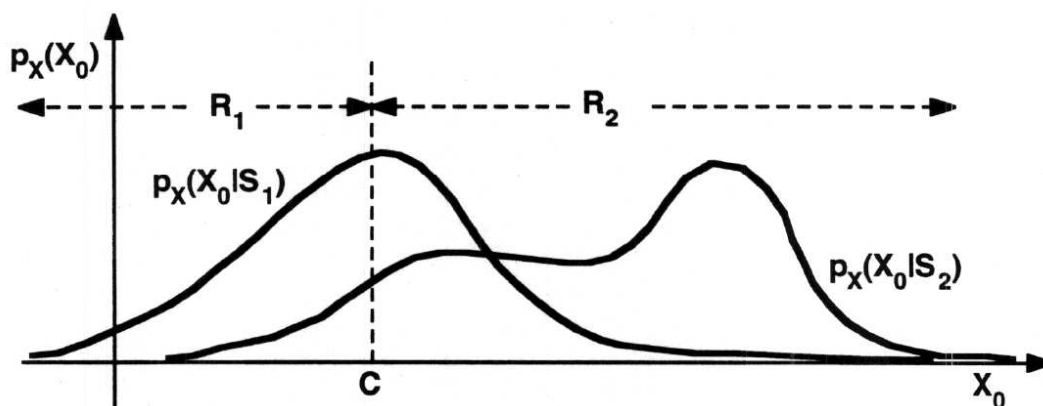
$$\begin{array}{cc} [f(R_2 | S_2), f(R_2 | S_1)] & [f(R_1 | S_2), f(R_1 | S_1)] \\ [f(R_2 | S_2), f(R_1 | S_1)] & [f(R_1 | S_2), f(R_2 | S_1)]. \end{array}$$

Pre zjednodušenie, budeme v ďalšom uvažovať len o páre prvkov $[f(R_2 | S_2), f(R_2 | S_1)]$. A vychádzajúc zo zjednodušenia, keď sa uvažuje o detekcii, budeme $f(R_2 | S_2)$ nazývať anglicky *hit rate* (niekedy sa tiež používa *detection rate*, alebo slovensky relatívna početnosť výskytu správnej detekcie), a $f(R_2 | S_1)$ budeme označovať ako *false alarm rate* (čo by slovensky šlo preložiť ako relatívna početnosť výskytu nesprávnej detekcie).

Axiómy rozhodovacieho modelu pre 1I-2AFC experimentálnu paradigmu sú:

1. Existuje reálna náhodná premenná X (nazveme ju „rozhodovacia os“, alebo „decision axis“), ktorá má tú vlastnosť, že pri každej prezentácii signálu nadobudne premenná X nejakú hodnotu.
2. Na osi X existuje nejaká kritériálna hodnota C (nazývaná „kritérium“).
3. Správanie subjektu je také, že odpovedá vždy R_1 v prípade, že $X < C$, a že odpovedá vždy R_2 v prípade, že $X \geq C$.
4. Štatistické charakteristiky premennej X sú nezávislé od všetkých aspektov daného experimentu s výnimkou stimulov S_1 a S_2 , a tieto charakteristiky sú plne popísané funkciami podmienenej hustoty pravdepodobnosti $p_X(X_0|S_1)$ a $p_X(X_0|S_2)$. Hlavne platí, že štatistické charakteristiky premennej X sú nezávislé na apriórnych pravdepodobnostiach odmiern, a že jednotlivé pokusy experimentu sú na sebe štatisticky nezávislé.

Výsledný model správania sa poslucháča je ilustrovaný na Obr. 6-1.



Obr. 6-1 Model rozhodovania pre experimentálnu paradigmu 1I-2AFC.

V tomto modeli sú podmienené pravdepodobnosti odpovedí $P(R_i | S_j)$ definované ako:

$$\begin{aligned}
 P(R_1 | S_1) &= \int_{-\infty}^C p_X(X_0 | S_1) dX_0 & P(R_2 | S_1) &= \int_C^{+\infty} p_X(X_0 | S_1) dX_0 \\
 P(R_1 | S_2) &= \int_{-\infty}^C p_X(X_0 | S_2) dX_0 & P(R_2 | S_2) &= \int_C^{+\infty} p_X(X_0 | S_2) dX_0
 \end{aligned}$$

a pre tieto pravdepodobnosti platí, že $P(R_1 | S_j) + P(R_2 | S_j) = 1$, a že *detection rate* $P_D = P(R_2 | S_2)$ a *false alarm rate* $P_F = P(R_2 | S_1)$.

6.2.2 Všeobecné vlastnosti modelu

Podľa tohto modelu je schopnosť človeka odlíšiť stimuly (tiež nazývaná senzitivitou) daná mierou, do akej sa pravdepodobnostné hustoty $p_X(X_0|S_1)$ a $p_X(X_0|S_2)$ neprekrývajú. A parameter C vyjadruje mieru, do akej má človek tendenciu preferovať jednu z dvoch možných odpovedí, tiež nazývanú *systematická chyba* (angl. *bias*, my budeme používať aj slovo *strannosť*). Vo všeobecnosti sa predpokladá, že hodnota C je určená apriórными pravdepodobnosťami $P(S_j)$ a hodnotami odmeny $V(R_i | S_j)$. Napr., ak poslucháč vie, že $P(S_1) \gg P(S_2)$ a preto očakáva S_1 oveľa častejšie ako

S_2 , je pravdepodobné, že v jeho odpovediach bude systematická chyba zodpovedajúca častejším odpovediam R_1 než R_2 , čiže hodnota C bude u neho vysoká. Podobne, ak $V(R_2 | S_2) \gg V(R_1 | S_1)$ (t.j. ak odmena za odpoveď R_2 na stimul S_2 je oveľa väčšia než odmena za odpoveď R_1 na stimul S_1) a zároveň ak $-V(R_1 | S_2) \gg -V(R_2 | S_1)$ (t.j. ak trest za odpoveď R_1 na stimul S_2 je oveľa väčší než trest za odpoveď R_2 na stimul S_1), potom v odpovediach subjektu bude pravdepodobne systematická chyba zodpovedajúca častejším odpovediam R_2 než R_1 , čiže hodnota C bude u neho nízka. Opäť, na základe motivácie z detekčných experimentov (v ktorých S_2 zodpovedá stimulu obsahujúcemu signál, ktorý máme detekovať) sa v prípade veľkých hodnôt C hovorí o „konzervatívnom“ umiestnení kritéria, zatiaľ čo, ak je C malé, hovorí sa, že človek používa „liberálne“ kritérium. Čím je kritérium liberálnejšie, tým vyššie sú hodnoty P_D aj P_F .

Ďalšie dôležité vlastnosti popísaného modelu rozhodovania sú:

1. Vplyv stimulov na presnosť odpovedí [charakterizovaný vlastnosťami pravdepodobnostných hustôt $p_x(X_0|S_1)$ a $p_x(X_0|S_2)$] je úplne separovateľný od vplyvu apriórnych pravdepodobností stimulov a odmien (charakterizovaný hodnotou parametra C). To znamená, že zmena v apriórnych pravdepodobnostiach alebo odmenách nespôsobí žiadnu zmenu v pravdepodobnostných hustotách.
2. Pri každom jednotlivom pokuse je vplyv prezentácie stimulu na odpoveď poslucháča úplne charakterizovaný hodnotou jedinej reálnej premennej. Čiže, predpokladá sa, že priestor rozhodovania X je jednodimenzionálny.
3. Rozhodovacie regióny na osi X pozostávajú z jednotných intervalov $X < C$ a $X \geq C$, ktoré sa počas experimentu nemenia. To znamená, že na základe definovaných axiém sa v tomto modeli nepripúšťa možnosť, že by existovali dve kritéria C_1 a C_2 , a že odpoveď R_1 by zodpovedala napr. oblasti $C_1 < X < C_2$, zatiaľ čo odpoveď R_2 by zodpovedala napr. oblasti $(X \leq C_1) \cup (X < C_2)$. Taktiež je axiomatically vylúčené, že kritérium C by bola náhodná premenná, ktorá by sa menila od pokusu k pokusu, tak ako rozhodovacia premenná X .

Na aplikovanie tohto modelu na konkrétne dáta je v modeli potrebné došpecifikovať model senzitivity sluchového systému, čiže toho, ako sa v sluchovej dráhe prevádza zvuk na pravdepodobnostnú distribúciu $p_x(X_0|S_i)$. Taktiež je potrebný model “biasu” (model toho, kam subjekt umiestnil kritérium C), to ale nateraz ignorujeme, a pozrime sa podrobnejšie na špecifikáciu p_x .

6.2.3 ROC krivka pre II-2AFC experimenty

Definujme si teraz na rovnice, ktoré popisujú vzťah pravdepodobností *detection rate* P_D a *false alarm rate* P_F k pravdepodobnostným hustotám $p_x(X_0|S_1)$ a $p_x(X_0|S_2)$, a ku kritériu C :

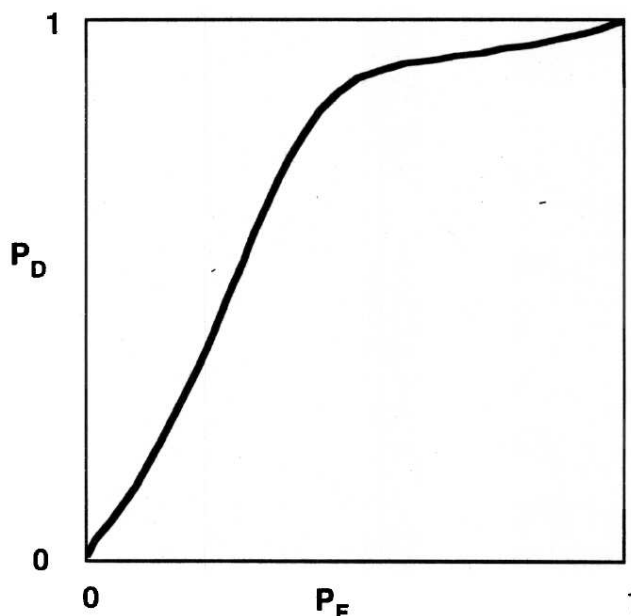
$$P_D = \int_C^{+\infty} p_x(X_0 | S_2) dX_0 \quad P_F = \int_C^{+\infty} p_x(X_0 | S_1) dX_0$$

Ak poznáme p_x a C , môžeme podľa týchto rovníc priamo vypočítať P_D a P_F . My ale potrebujeme vyriešiť opačnú úlohu: experimentálne nameriame P_D a P_F a chceme určiť $p_x(X_0|S_1)$ a $p_x(X_0|S_2)$. Toto je, bohužiaľ, vo všeobecnosti neriešiteľný

problém. Reálne sa k správne mu riešeniu dá aspoň teoreticky dostať nasledovným postupom:

1. Budeme meniť hodnoty parametra C a pre každú hodnotu odmeriame zodpovedajúce P_D a P_F .
2. Na základe takéhoto súboru hodnôt P_D a P_F ešte stále nemôžeme odvodiť samotné distribúcie $p_x(X_0|S_1)$ a $p_x(X_0|S_2)$, pretože pre žiadnu z nich nepoznáme zodpovedajúce C (C totiž môžeme len nepriamo ovplyvňovať, nedokážeme ho priamo nastaviť).
3. C ale môžeme z rovníc eliminovať a miesto dvoch rovníc môžeme sledovať len závislosť hodnôt P_D ako funkciu P_F .

Takáto krivka, ktorá vykresľuje závislosť hodnôt P_D ako funkciu P_F , sa anglicky nazýva „Receiver Operating Characteristic“ (ROC) alebo krivka izosezitivity. Príklad ROC krivky je ukázaný na Obr. 6-2.



Obr. 6-2 Krivka izosezitivity ROC pre experimentálnu paradigmu 1I-2AFC.

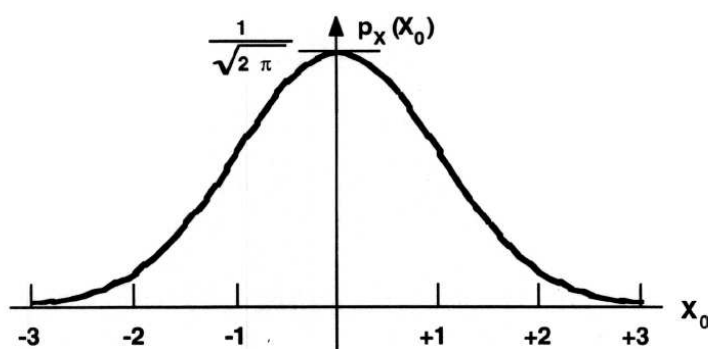
Táto krivka má niekoľko všeobecných vlastností. Napr., je neklesajúca a vždy prechádza bodmi $[0,0]$ a $[1,1]$. Ďalej pre ňu v zásade platí, že všetky jej body ležia nad diagonálou. A vo všeobecnosti platí, že čím bližšie k bodu $[0,1]$ prechádza, tým je schopnosť modelovaného človeka rozlišovať stimuly lepšia. Ešte stále nám táto krivka ale neumožňuje nájsť pravdepodobnostné distribúcie $p_x(X_0|S_1)$ a $p_x(X_0|S_2)$, na základe ktorých sa subjekt rozhodoval.

6.2.4 Index citlivosti a bias pre špeciálny prípad Gaussovských distribúcií a rovnakých variancií

Vo všeobecnosti nie je možné určiť tvar kriviek $p_x(X_0|S_1)$ a $p_x(X_0|S_2)$. Pri určitých zjednodušeníach sa k tomu ale môžeme priblížiť. Prvým zjednodušením, nad ktorým budeme uvažovať, je že tieto distribúcie sú Gaussove:

$$p_X(X_0 | S_j) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(X_0 - M_j)^2}{2\sigma^2}}$$

Pre tieto distribúcie sa dá ľahko dokázať, že $p_X(X_0 | S_j)$ je skutočne hustota pravdepodobnosti, a že pre ňu platí, že $E[X | S_j] = M_j$ a $E[(X - M_j)^2 | S_j] = \sigma^2$, čiže že priemerná hodnota X podmienená výskytom stimulu S_j je M_j a že variancia X podmienená stimulu S_j je σ^2 (pre obe j). Toto zjednodušenie je prípustné, ak uvažujeme, že X je súčtom veľkého množstva nezávislých neurónov a ak platí Centrálna limitná teoréma. Príklad Gaussovej pravdepodobnostnej distribúcie p_X je na Obr. 6-3.



Obr. 6-3 Gaussova distribúcia pravdepodobnosti.

Ak našu situáciu ďalej zjednodušíme tým, že budeme predpokladať, že obe naše Gaussove distribúcie $p_X(X_0 | S_1)$ a $p_X(X_0 | S_2)$ majú rovnakú varianciu σ^2 , potom pre tieto distribúcie platí

$$p_X\left(\frac{M_1 + M_2}{2} + K | S_1\right) = p_X\left(\frac{M_1 + M_2}{2} - K | S_2\right),$$

čiže tieto distribúcie sú identické až na to, že sú od seba navzájom posunuté. Teraz zavedieme dve nové premenné:

$$d' = \frac{M_2 - M_1}{\sigma} \quad \beta = \frac{1}{\sigma} \left[C - \frac{M_2 + M_1}{2} \right]$$

Parameter d' (anglicky sa píše aj d -prime) je nezávislý na kritériu C a vlastne meria ako veľmi sú od seba distribúcie $p_X(X_0 | S_1)$ a $p_X(X_0 | S_2)$ vzdialené (špecificky, d' udáva vzdialenosť stredov distribúcií v jednotkách štandardných odchýlok). Inými slovami, d' vyjadruje mieru, do akej je v mozgu poslucháča stimul S_1 odlišiteľný od stimulu S_2 . Táto miera sa nazýva aj „citlivosť“ alebo „index citlivosti“ (angl. *sensitivity* alebo *sensitivity index*) a je veľmi dôležitá, pretože nám hovorí, s akou presnosťou sa dané stimuly mentálne reprezentujú, a to nezávisle na tom, ako subjektívny alebo tendenčný môže byť poslucháč pri svojom odpovedaní. Čiže, d' je **objektívnou** mierou citlivosti.

Parameter β , je vlastne len prenормovaním hodnoty kritéria C , a teda meria tendenciu človeka odpovedať skôr R_1 než R_2 , nech už je táto tendencia spôsobená čímkoľvek. Anglicky sa tento parameter nazýva *bias*, my ho budeme nazývať aj

systematická chyba alebo *strannosť* (ako opak nestrannosti). Preškáľovanie hodnoty β má význam, keďže po ňom pre β platí, že $\beta = 0$ zodpovedá situácii, kedy človek nemá žiadnu tendenciu preferovať niektorú z odpovedí, čiže keď je nestranný.

6.2.5 Transformácia na jednotky z-scores

Uvažujme teraz o transformácii súradnicového systému $[P_D, P_F]$ na súradnicový systém $[Z_D, Z_F]$, a to použitím integrálových rovníc:

$$P_D = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Z_D} e^{-\frac{x^2}{2}} dx \quad P_F = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Z_F} e^{-\frac{x^2}{2}} dx$$

ktoré predstavujú normálne kumulatívne distribučné funkcie. Hodnoty Z_D a Z_F sa nazývajú normálnymi odchýlkami (angl. *normal deviates*) hodnôt P_D a P_F . Súradnicový systém definovaný párom $[Z_D, Z_F]$ sa potom nazýva „súradnicový systém normálnych odchýliek“, zatiaľ čo súradnicový systém $[P_D, P_F]$, v ktorom je definovaná štandardná ROC krivka, je „lineárnym súradnicovým systémom“. Čo konkrétne hodnoty Z_D a Z_F vyjadrujú, sa dá ľahko vidieť, keď dosadením do posledných rovníc získame vzťahy:

$$Z_F = \frac{M_1 - C}{\sigma} \quad Z_D = \frac{M_2 - C}{\sigma}$$

Tieto rovnice hovoria, že Z_F nám udáva, ako ďaleko od kritéria C v jednotkách štandardných odchýlok je stred distribúcie pre S_1 , a Z_D nám udáva, ako ďaleko je stred distribúcie pre S_2 .

Vo všeobecnosti sa transformovanie dát do jednotiek „z“ nazýva aj vyhodnocovaním použitím jednotiek „z-scores“. Táto transformácia je užitočná napr. pre to, že v nej je veľmi jednoduché určiť parametre d' a β , a to konkrétne týmito rovnicami:

$$\begin{aligned} d' &= Z_D - Z_F & \beta &= -\frac{Z_D - Z_F}{2} \\ Z_D &= \frac{d'}{2} - \beta & Z_F &= -\frac{d'}{2} - \beta \end{aligned}$$

Čiže, na výpočet indexu citlivosti nejakého poslucháča v behaviorálnom experimente nám stačí získať P_D a P_F . Potom, výpočtom inverznej kumulatívnej distribučnej funkcie (alebo použitím tabuliek) tieto hodnoty transformujeme na Z_D a Z_F . A index citlivosti d' je priamo daný rozdielom týchto dvoch hodnôt.

Poznámka: Transformácia na z-scores má aj množstvo ďalších užitočných vlastností. Napr., ROC krivka vykreslená na základe týchto hodnôt má vlastnosť, že krivky konštantných citlivostí sú pre každú hodnotu citlivosti priamkami so sklonom 1. To isté platí aj pre krivky konštantných stranností, ktoré sú na krivky konštantných citlivostí kolmé (majú sklon -1).

6.2.6 Zhrnutie a ďalšie modely

V tejto kapitole sme odvodili najjednoduchšiu formu modelu rozhodovania sa v experimentálnej paradigme jedného intervalu, dvoch alternatív a vynútenej odpovede (1I-2AFC). Vychádzajúc tohto modelu a z predpokladu, že sú distribúcie

vnemov gaussovské s rovnakou varianciou, vieme na základe meraní presne určiť index citlivosti (d') pre daný podnet, ako aj sklonnosť (bias, tendenciu subjektu) odpovedať jednostranne (β). Tieto dva parametre vieme určiť nezávisle od seba, čo je užitočné lebo často nás zaujíma iba jeden z nich (najčastejšie d'). Konkrétne príklady postupov pri použití tohto modelu sú uvedené v zadaní dostupnom na webstránke knihy.

Náš model by sa dal relatívne ľahko rozšíriť na zložitejšie situácie (napr. na nerovnaké variancie) a na zložitejšie experimentálne paradigmy (napr. na 2I-2AFC experimenty). Takéto rozšírenie by už ale presahovalo rozsah tejto knihy. Zájemcovia môžu štúdium začať knihami (Green and Swets 1974; Macmillan and Creelman 2005).

Kapitola 7 Záver

Účelom tohto textu bolo podať základný prehľad o výpočtovej neurovede a neurálnom modelovaní, a na štyroch vybraných modeloch popísať, akú hĺbku matematických poznatkov je potrebné obsiahnuť pre správne porozumenie dátam a teóriám v neurovede. Keďže podaný prehľad výpočtovej neurovedy nie je ani zďaleka kompletný, tu sa zmieňujeme aspoň o najdôležitejších ďalších konceptoch, ktoré sa v tomto odbore používajú, a ktoré môžu čitateľa zaujať. Prvým takýmto konceptom je použitie bayesovsových techník alebo Shannonovej informačnej teórie pre analýzu dát (Rieke 1997). Ďalším je použitie celého spektra metód pre spracovanie a analýzu signálov a obrazov (Oweiss 2010), s ktorým sa v neurovedách bežne stretávame. Veľké množstvo každodenných analýz používa štandardné ale aj nové metódy deskriptívnej a inferenčnej štatistiky, ako napr. Analýzu variancie (ANOVA), ktorá bola vždy východnou metódou pre analýzu behaviorálnych dát (Winer, Brown et al. 1991), ale dnes sa modifikuje pre analýzu dát získaných zobrazovacími metódami ako funkčná magnetická rezonancia (angl. functional Magnetic Resonance Imaging, fMRI) (Ashby 2011). Ale aj vývoj ďalších nových techník pre sledovanie aktivity mozgu vedie k novým výpočtovým metódam. Napr., na analýzu elektroencefalografických dát sa používajú výpočtové techniky „analýzy nezávislých komponentov“ (angl. Independent Component Analysis, ICA) a „slepejšie separácie zdrojov“ [angl. Blind Source Separation, BBS, (Comon 2010)].

Ambíciou tejto knihy nebolo poskytnúť všeobecný úvod do kognitívnych a neurálnych vied. Za zmienku ale stojí aspoň usmerniť čitateľa k najdôležitejším prehľadovým zdrojom v týchto odboroch. Štandardnými úvodnými textmi v angličtine sú napr. (Gazzaniga, Ivry et al. 2009), (Kandel, Schwartz et al. 2000) a (Stillings 1995). Ale aj v slovenčine a češtine sú už k dispozícii úvodné prehľadové texty, napr. (Rybar, Benusková et al. 2002), (Benuskova, Kvasnicka et al. 2000), (Dobes 2005), (Novak, Faber et al. 1993), či (Rybar, Kvasnicka et al. 2005). Asi najúplnejší výčet slovenských a českých neurokognitívnych zdrojov, ako aj kontakty na vedeckú komunitu a oznamy o konferenciách, sú na webe <http://cogscience.cuni.cz>. Nech aj táto kniha, ktorú ste práve dočítali, prispeje k rozvoju neurovedy, kognitívnej vedy a ich medicínskych a technických aplikácií u nás.

Literatúra

- Anderson, J. A. and E. Rosenfeld (1998). Talking nets : an oral history of neural networks. Cambridge, Mass., MIT Press.
- Anderson, J. A., E. Rosenfeld, et al. (1988). Neurocomputing. Cambridge, Mass., MIT Press.
- Ashby, F. G. (2011). Statistical analysis of fMRI data. Cambridge, Mass., MIT Press.
- Benuskova, L., V. Kvasnicka, et al., Eds. (2000). Hľadanie spoločného jazyka v kognitívnych vedach. Bratislava, Iris.
- Caianiello, E. R. (1961). "Outline of a theory of thought-processes and thinking machines." J Theor Biol **1**: 204-235.
- Comon, P. (2010). Handbook of blind source separation : independent component analysis and blind deconvolution. Boston, MA, Elsevier.
- Daugman, J. (2001). Brain metaphor and brain theory. Philosophy and the Neurosciences. W. e. a. Bechtel. Oxford, Blackwell Publishers.
- Dayan, P. and L. F. Abbott (2001). Theoretical Neuroscience. Cambridge, MA, MIT Press.
- Devore, J. L. and R. Peck (1994). Introductory statistics. Minneapolis/St. Paul, West Pub.
- DeYoe, E. A. and D. C. Van Essen (1988). "Concurrent processing streams in monkey visual cortex." Trends Neurosci **11**(5): 219-226.
- Dobes, M. (2005). Zaklady neuropsychologie. Kosice, Spolocenskovedny ustav SAV.
- Duda, R. O., P. E. Hart, et al. (2001). Pattern Classification, Wiley-Interscience.
- Durlach, N. I. and L. D. Braida (1969). "Intensity perception. I. Preliminary theory of intensity resolution." Journal of the Acoustical Society of America **46**(2): 372-383.
- Fechner, G. T. (1860). Elemente der psychophysik. Leipzig,, Breitkopf und Härtel.
- Freud, S. (1913). The interpretation of dreams. New York,, The Macmillan company.
- Gazzaniga, M. S., R. B. Ivry, et al. (2009). Cognitive neuroscience : the biology of the mind. New York, W.W. Norton.
- Green, D. M. and J. A. Swets (1974). Signal detection theory and psychophysics. Huntington, N.Y., R. E. Krieger Pub. Co.
- Grossberg, S. (1967). "Nonlinear difference-differential equations in prediction and learning theory." Proc Natl Acad Sci U S A **58**(4): 1329-1334.
- Grossberg, S. (1976). "Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors." Biol Cybern **23**(3): 121-134.
- Hebb, D. O. (1949). The organization of behavior. New York, Wiley & Sons.

- Helmholtz, H. v. and A. J. Ellis (1875). On the sensations of tone as a physiological basis for the theory of music. London, Longmans, Green, and Co.
- Hertz, J., A. Krogh, et al. (1991). Introduction to the theory of neural computation. Redwood City, Calif., Addison-Wesley Pub. Co.
- Hodgkin, A. L. and A. F. Huxley (1952). "A quantitative description of membrane current and its application to conduction and excitation in nerve." J Physiol **117**(4): 500-544.
- Hopfield, J. J. (1982). "Neural networks and physical systems with emergent collective computational abilities." Proc Natl Acad Sci U S A **79**(8): 2554-2558.
- Hubel, D. H. and T. N. Wiesel (1959). "Receptive fields of single neurones in the cat's striate cortex." J Physiol **148**: 574-591.
- Kandel, E. R., J. H. Schwartz, et al. (2000). Principles of neural science. New York, McGraw-Hill, Health Professions Division.
- Kohonen, T., P. Lehtio, et al. (1974). "Modelling of neural associative memory." Ann Acad Sci Fenn A(167): 1-18.
- La Mettrie, J. O. d. (1748). L'homme machine. Paris,, Nord-Sud.
- Lewicki, M. S. (2002). "Efficient coding of natural sounds." Nature Neuroscience **5**(4): 356-363.
- Macmillan, N. A., L. D. Braida, et al. (1989). "Psychophysical laws: A call for deregulation." Behavioral and Brain Sciences **12**: 251-320.
- Macmillan, N. A. and C. D. Creelman (2005). Detection theory : a user's guide. Mahwah, N.J., Lawrence Erlbaum Associates.
- Mann, P. S. and C. J. Lacke (2010). Introductory statistics. Hoboken, NJ, John Wiley & Sons.
- Marr, D. (1969). "A theory of cerebellar cortex." J Physiol **202**(2): 437-470.
- Marr, D. (1982). Vision: A computational investigation into the human representation and processing of visual information, W.H. Freeman publ.
- Masin, S. C., V. Zudini, et al. (2009). "Early alternative derivations of Fechner's law." J Hist Behav Sci **45**(1): 56-65.
- McCulloch, W. and W. Pitts (1943). "A Logical Calculus of Ideas Immanent in Nervous Activity." Bulletin of Mathematical Biophysics **5**: 115-133.
- Minsky, M. L. and S. Papert (1969). Perceptrons; an introduction to computational geometry. Cambridge, Mass., MIT Press.
- Minsky, M. L. and S. Papert (1988). Perceptrons : an introduction to computational geometry. Cambridge, Mass., MIT Press.
- Novak, M., J. Faber, et al., Eds. (1993). Neuronove site a informacni systemy zivych organismu. Praha, Grada.

- Oweiss, K. G. (2010). Statistical signal processing for neuroscience and neurotechnology. Burlington, MA, Academic Press/Elsevier.
- Pavlov, I. P. and G. V. Anrep (1927). Conditioned reflexes; an investigation of the physiological activity of the cerebral cortex. London, Oxford University Press: Humphrey Milford.
- Posner, M. I., S. E. Petersen, et al. (1988). "Localization of cognitive operations in the human brain." Science **240**(4859): 1627-1631.
- Rashevsky, N. (1938). Mathematical Biophysics, The University of Chicago Press.
- Rescorla, R. A. and A. R. Wagner (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. Classical Conditioning II. A.H. Black & W.F. Prokasy, Eds.: 64–99.
- Rieke, F. (1997). Spikes : exploring the neural code. Cambridge, Mass., MIT Press.
- Rosenblatt, F. (1958). "The perceptron: a probabilistic model for information storage and organization in the brain." Psychol Rev **65**(6): 386-408.
- Rosenblatt, F. (1962). Principles of Neurodynamics. Washington, DC, Spartan Books.
- Rumelhart, D. E., J. L. McClelland, et al. (1986). Parallel distributed processing : explorations in the microstructure of cognition. Cambridge, Mass., MIT Press.
- Rybar, J., L. Benusková, et al., Eds. (2002). Kognitívne vedy. Bratislava, Kalligram.
- Rybar, J., V. Kvasnicka, et al., Eds. (2005). Jazyk a kognícia. Bratislava, Kalligram.
- Shannon, C. E. (1997). "The mathematical theory of communication. 1963." MD Comput **14**(4): 306-317.
- Stillings, N. A. (1995). Cognitive science : an introduction. Cambridge, Mass., MIT Press.
- Sutton, R. S. and A. G. Barto (1981). "Toward a modern theory of adaptive networks: expectation and prediction." Psychol Rev **88**(2): 135-170.
- Thurstone, L. L. and E. J. Chave (1929). The measurement of attitude; a psychophysical method and some experiments with a scale for measuring attitude toward the church. Chicago, Ill., The University of Chicago press.
- Torgerson, W. S. (1958). Theory and methods of scaling. New York., Wiley.
- Turing, A. (1950). "Computing Machinery and Intelligence." Mind **LIX**(236): 433–460.
- von der Malsburg, C. (1973). "Self-organization of orientation sensitive cells in the striate cortex." Kybernetik **14**(2): 85-100.
- von Melchner, L., S. L. Pallas, et al. (2000). "Visual behaviour mediated by retinal projections directed to the auditory pathway." Nature **404**(6780): 871-876.
- Werbos, P. J. (1994). The roots of backpropagation : from ordered derivatives to neural networks and political forecasting. New York, Wiley.
- Wessel, R., C. Koch, et al. (1996). "Coding of time-varying electric field amplitude modulations in a wave-type electric fish." J Neurophysiol **75**(6): 2280-2293.

- Winer, B. J., D. R. Brown, et al. (1991). Statistical principles in experimental design. New York, McGraw-Hill.
- Winograd, S. and J. D. Cowan (1963). Reliable Computation in the Presence of Noise, The M. I. T. Press.
- Yost, W. A. (1994). Fundamentals of Hearing: An Introduction (3e). San Diego, CA, Academic Press.