

8. Hopfieldov model

8.1 Úvod

Pri štúdiu štatistických nelineárnych kooperatívnych systémov, akými sú napr. spinové sklá, prišli fyzici na myšlienku využiť mnohé ich zaujímavé vlastnosti na vytvorenie idealizovaných neurónových sietí, ktorých správanie sa možno interpretovať ako analógie rôznych mozgových (či psychických) funkcií, ako sú napr. asociatívne vyvolávanie z pamäti, spracovanie časových postupností stimulov, zabúdanie, a iné. Prvými, ktorí poukázali na analógiu medzi procesmi prechodu z neusporiadaných do usporiadaných stavov v magnetických látkach a procesmi, ktoré by mohli prebiehať v reálnych neurónových sieťach, boli Cragg a Temperley [17, 18] a Little [57]. V stabilnom stave by podľa nich priestorové usporiadanie domén (oblastí) atómových spinov s orientáciou "hore" a domén spinov orientovaných "dole" korešpondovalo s usporiadaním oblastí aktivovaných a neaktivovaných neurónov v sieti pri vnútornej reprezentácii nejakého zapamätaného vzoru. Po publikovaní článku Johna Hopfielda [40], ktorého model predstavuje interpretáciu Sherringtonovho-Kirkpatrickovho a Isingovho modelu magnetika [44, 48] v zmysle neurónovej siete, nastala v nasledujúcom desaťročí "explózia" štúdiá takto vytvorených modelov neurónových sietí. Vďaka vyvinutému fyzikálnemu aparátu patrí Hopfieldov model a jeho modifikácie medzi teoreticky najlepšie preštudované modely neurónových sietí.

Hopfieldove neurónové siete sa často označujú aj ako **atraktorové** alebo **autoasociatívne** neurónové siete. Patria do triedy tzv. celulárnych (bunečných) automatov, čo sú vo všeobecnosti dynamické systémy pozostávajúce z veľkého množstva dvojstavových (alebo viacstavových) prvkov, navzájom viazaných, s definovaným pravidlom na zmenu stavov prvkov, ktorých makroskopické správanie sa je popísané v ideálnom prípade analytickými rovnicami. Autoasociatívne siete sú iba jedným z mnohých prístupov k vytváraniu umelých neurónových sietí a k modelovaniu mozgových funkcií. Podľa nášho názoru netreba rôzne prístupy chápať ako vzájomne sa vylučujúce, ale skôr z toho hľadiska, že každý z nich lepšie vyjadruje či popisuje iné aspekty pestrej variety vlastností a činností, ktorými sa mozog vyznačuje. Na Hopfieldových autoasociatívnych neurónových sieťach nás zaujali najmä dve veci. Po prvé, ich výpočtové vlastnosti možno interpretovať v neurobiologických súvislostiach ako model predpokladaných procesov, ktoré prebiehajú v mozgu pri kognitívnom spracovávaní informácií (takýmto kognitívnym spracovávaním je napr. vnímanie) [6, 12, 13]. Po druhé, všetky ich výpočtové vlastnosti sa vynárajú (angl. *emerge*) ako dôsledok paralelnej činnosti veľkého množstva navzájom interagujúcich jednoduchých procesorov (modelových neurónov). Autoasociatívne siete môžu tvoriť most medzi "mikroskopickými" modelmi neurónov a ich sietí, založenými na káblovej teórii [45, 53], a "makroskopickými" modelmi, ktoré sú zamerané na hierarchickú organizáciu mozgovovej činnosti. Pritom sa javí atraktívnou možnosť, že mnohé javy

pozorované na vyšších úrovniach spracovávaní informácií v jednotlivých oblastiach mozgu by sa mohli dať pripísať emergentným (spontánne sa objavujúcim) kolektívnym vlastnostiam sietí zložených z veľkého počtu neurónov.

8.2 Základný popis

Základný popis Hopfieldovho modelu autoasociatívnej neurónovej siete (viď obr. 8.1) a dynamiky jej časového vývoja je založený na analógii s Isingovou magnetickou teóriou a Sherringtonovom-Kirkpatrickovom modeli spinového skla [9, 44, 48]:

(1) **Neurón** (analógia Isingovho spinu atómu) môže byť v jednom z dvoch stavov, t.j. $S_i \in \{-1, +1\}$. Nech N je celkový počet neurónov v sieti. Vstupný a zároveň aj výstupný stav (**konfigurácia aktivity**) siete je vyjadrený N -rozmer-ným binárnym vektorom $\mathbf{S}=(S_1, S_2, \dots, S_N)$.

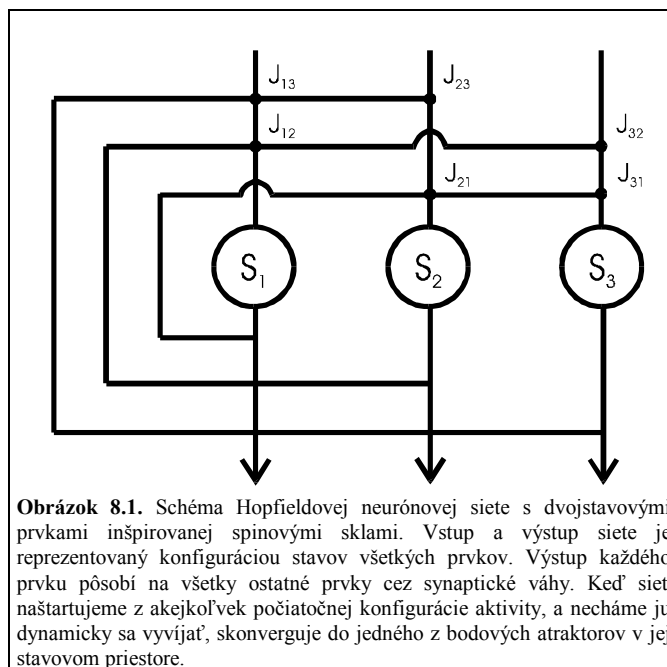
(2) J_{ij} (angl. *junction*) je **váha synapsy** tvorenej j -tym neurónom na i -tom neuróne (analógia interakčnej konštanty medzi i -tym a j -tym spinom). Pre excitačnú synapsu platí $J_{ij} > 0$ a pre inhibičnú synapsu $J_{ij} < 0$. Neurón na sebe samom synapsy netvorí, t.j. $J_{ii} = 0$. Suma príspevkov od jednotlivých neurónov S_j , pričom $j=1, \dots, N$, váhovaná prostredníctvom synaptických váh J_{ij} , vyjadruje postsynaptický potenciál, ktorý je mierou excitácie i -teho neurónu. **Postsynaptický potenciál** sa zvykne značiť ako h_i^{int} (analógia vnútorného magnetického poľa).

$$h_i^{\text{int}} = \sum_{j=1}^N J_{ij} S_j \quad (8.1)$$

(3) Neurón sa aktivuje, t.j. generuje na svojom výstupe akčný potenciál, ak postsynaptický potenciál h_i^{int} prekročí istú hodnotu prahového napätia, tzv. **prah excitácie** neurónu. Táto veličina sa tiež zvykne značiť h_i^{ext} ako jej analógia, vonkajšie pole pôsobiace na spin v magnetiku. Celkový **efektívny postsynaptický potenciál** neurónu je potom $h_i = h_i^{\text{int}} - h_i^{\text{ext}}$.

(4) **Deterministické prechodové pravidlo** pre zmenu stavu i -teho neurónu je dané týmto predpisom

$$S_i \rightarrow S'_i = \text{sign}(h_i) = \text{sign}\left(\sum_{j=1}^N J_{ij} S_j - h_i^{\text{ext}}\right) \quad (8.2)$$



pričom funkcia $\text{sign}(x)$ je definovaná takto:

$$\text{sign}(x) = \begin{cases} +1 & \text{pre } x > 0, \\ -1 & \text{pre } x < 0. \end{cases} \quad (8.3)$$

Stavy neurónov môžu byť $S_i \in \{-1, +1\}$, vonkajšie pole kladieme obyčajne rovné nule, takže $h_i \neq 0$. Aktualizovanie stavu neurónov podľa vzťahu (8.2) môže prebiehať dvoma spôsobmi. Prvým je **synchronná (paralelná) dynamika**, keď všetky neuróny menia svoj stav naraz, t.j. v čase t platí

$$S_i(t) = \text{sign} \left(\sum_{\substack{j=1 \\ j \neq i}}^N J_{ij} S_j(t-1) - h_i^{\text{ext}}(t) \right), \quad \text{pre } i = 1, \dots, N. \quad (8.4)$$

Jeden cyklus relaxácie (prechod) odpovedá aktualizácii stavu všetkých N neurónov.

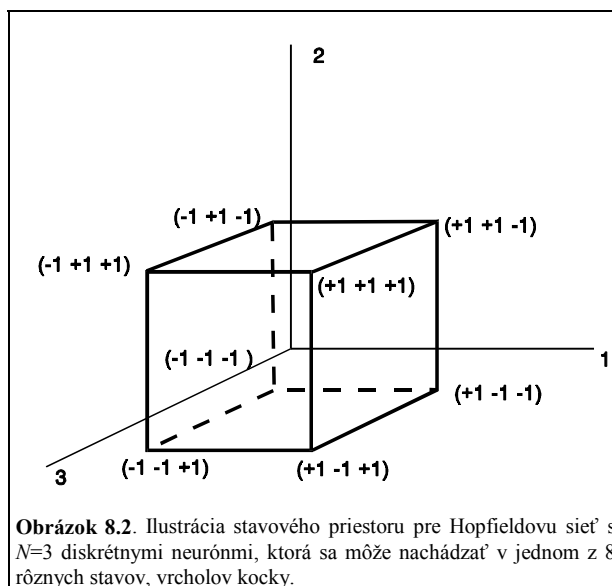
Druhou možnosťou je **asynchronná (sekvenčná) dynamika**, keď v každom časovom momente t mení svoj stav len jeden náhodne vybraný neurón i . To znamená, že v N krokoch za sebou aktualizujeme vždy znova náhodne vybraný neurón podľa vzťahu

$$S_i(t) = \text{sign} \left(\sum_{\substack{j=1 \\ j \neq i}}^N J_{ij} S_j(t) - h_i^{\text{ext}}(t) \right). \quad (8.5)$$

Časový vývoj Hopfieldovej siete s N neurónmi, čiže sekvenciu stavov siete $\mathbf{S}(t)=(S_1(t), S_2(t), \dots, S_N(t))$ v čase, možno chápať ako trajektóriu idúcu cez vrcholy N -rozmernej hyperkocky, ktorá má 2^N možných vrcholov (obr. 8.2).

V prípade asynchrónnej dynamiky sú v jednom časovom kroku povolené iba prechody pozdĺž hrán do najbližších vrcholov hyperkocky prislúchajúcich stavom, ktoré sa od pôvodného líšia len hodnotou 1 spinu. Jeden **cyklus relaxácie (prechod)** siete odpovedá aktualizácii stavu neurónov v N krokoch.

(5) **Energia** danej konfigurácie aktivity \mathbf{S} (hamiltonián systému) je definovaná ako



$$E(\mathbf{S}) = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N J_{ij} S_i S_j - \sum_{i=1}^N S_i h_i^{ext} \quad (8.6)$$

8.3 Spontánna evolúcia Hopfieldovej siete

Výsledky pozorovaní **spontánnej evolúcie (relaxácie)** siete na základe vlastných počítačových simulácií a na základe výsledkov uvedených v [9, 40] zhrnieme do nasledujúcich bodov. Sieť štartovala z náhodného počiatočného stavu, t.j. počiatočné stavy neurónov S_i boli neurónom priradené náhodne, pričom $S_i = \pm 1$ s pravdepodobnosťou rovnou 0,5. Prahy excitácie h_i^{ext} boli buď všetky položené rovné nule alebo boli vygenerované ako malé náhodné čísla z intervalu $(-1,+1)$. Synaptické váhy J_{ij} boli tiež vygenerované ako náhodné čísla z intervalu $(-1,+1)$, a potom sa sieť nechala vyvíjať podľa vzťahu (8.4) alebo (8.5).

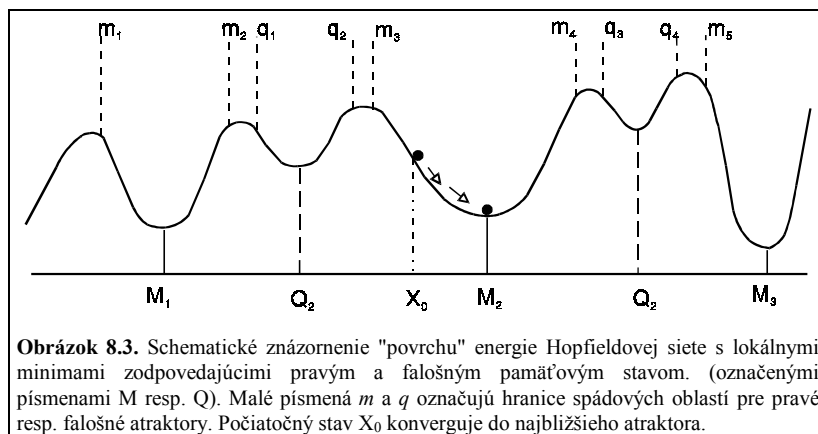
Evolúcia siete v priestore stavov silno závisí na tom, či je matica váh \mathbf{J} symetrická ($J_{ij} = J_{ji}$) alebo asymetrická ($J_{ij} \neq J_{ji}$), na konkrétnych hodnotách J_{ij} a na tom, či je prechodová dynamika synchronná (rovnica 8.4) alebo asynchronná (rovnica 8.5). Tá istá sieť (s tou istou maticou váh \mathbf{J} a tými istými prahmi h_i^{ext}) má úplne inú evolúciu vtedy, keď má synchronnú dynamiku, ako vtedy, keď má asynchronnú dynamiku.

Pre asynchronnú a synchronnú prechodovú dynamiku môžeme rozlíšiť tri typy asymptotického (dlhodobého) správania sa siete:

Chaotické trajektórie. Sieť "blúdi" v stavovom priestore. Energia siete v tvare (8.6) nepravidelne stúpa a klesá. Takéto správanie sa je typické pre prípad synchronnej dynamiky siete s asymetrickou synaptickou maticou \mathbf{J} ($\mathbf{J} \neq \mathbf{J}^T$) a ľubovoľnými prahmi excitácie, $h_i^{ext} \in (-1,+1)$. Už veľmi malá zmena v počiatočnom stave, napr. preklopenie 1 neurónu vedie k inej trajektórii v stavovom priestore. Chaotické trajektórie sú veľmi citlivé na počiatočné stavy siete.

Limitné cykly. Trajektórie, ktoré rýchlo (okolo 5 prechodov) vedú do malých cyklov, zložených najčastejšie z 2, prípadne 4 stavov. Cykly sú typické pre synchronnú dynamiku, ale zriedka sa vyskytnú aj pri asynchronnej dynamike. Sú teda citlivé na typ dynamiky siete.

Bodové atraktory. Trajektórie rýchlo vedú do jedného, potom už sa stále opakujúceho vzorca aktivity celej siete, ξ . Z ktoréhokoľvek počiatočného stavu sa sieť po niekoľkých prechodoch dostane do jediného stavu, v ktorom už zostane aj napriek tomu, že jej stav sa aktualizuje podľa rovníc (8.4) alebo (8.5). Takýto stav sa nazýva bodový atraktor. Pre jednu a tú istú maticu synaptických váh \mathbf{J} existuje v stavovom priestore niekoľko bodových atraktorov. Bodové atraktory zodpovedajú lokálnym minimám energie siete vyjadrenej rovnicou (8.6) — pozri obr. 8.3. Evolúcia do bodového atraktora je pomerne nezávislá na počiatočnom stave siete v tom zmysle, že mnohé počiatočné stavy siete končia v tom istom bodovom atraktore. Okolo každého atraktora je oblasť takých stavov, ktoré všetky konvergujú do toho istého bodového atraktora a nie do iného. Táto oblasť sa nazýva **spádová oblasť** atraktora. Každá matica váh \mathbf{J} má iné bodové atraktory.



Evolúcia siete do bodových atraktorov je typická pre asynchrónnu dynamiku. V prípade asynchrónnej dynamiky, keď je matica váh symetrická, t.j. $(J_{ij}=J_{ji})$, a $h_i^{ext} = 0$ pre $\forall i$, **energia** v tvare (8.6) **monotónne klesá**, pokiaľ nedosiahne minimum. Odvodíme si to nasledovným postupom. Energiu takejto Hopfieldovej siete v ľubovoľnom časovom okamihu môžeme rozpísať takto:

$$\begin{aligned}
 E &= -\frac{1}{2} \sum_{i \neq m}^N \sum_{j \neq m}^N J_{ij} S_i S_j - \frac{1}{2} \sum_{i=1}^N J_{im} S_i S_m - \frac{1}{2} \sum_{j=1}^N J_{mj} S_m S_j \\
 &= -\frac{1}{2} \sum_{i \neq m}^N \sum_{j \neq m}^N J_{ij} S_i S_j - \sum_{j=1}^N J_{mj} S_m S_j
 \end{aligned} \tag{8.7}$$

Najskôr sme z celkovej sumy vyňali dva členy, ktoré prislúchajú m -tému neurónu, ktorý bol náhodne vybraný a bude aktualizovaný podľa vzťahu (8.5). V ďalšej úprave sme využili to, že synaptická matica je symetrická, a teda platí $J_{mj}=J_{jm}$. Preto môžeme tieto dva vyňaté členy pre m -tý neurón zlúčiť dohromady (je jedno, či v prvej sume pre m -tý neurón sumujeme cez i alebo cez j). Nový stav vybraného neurónu označme ako S'_m . Po zmene stavu jedného neurónu bude nová energia E' rovná

$$\begin{aligned}
 E' &= -\frac{1}{2} \sum_{i \neq m}^N \sum_{j \neq m}^N J_{ij} S_i S_j - \frac{1}{2} \sum_{i=1}^N J_{im} S_i S'_m - \frac{1}{2} \sum_{j=1}^N J_{mj} S'_m S_j \\
 &= -\frac{1}{2} \sum_{i \neq m}^N \sum_{j \neq m}^N J_{ij} S_i S_j - \sum_{j=1}^N J_{mj} S'_m S_j
 \end{aligned} \tag{8.8}$$

Teraz odčítame (8.7) od (8.8) a dostaneme rozdiel energií v dvoch po sebe nasledujúcich časových krokoch:

$$\Delta E = E' - E = (S_m - S'_m) \sum_{j \neq m}^N J_{mj} S_j = (S_m - S'_m) h_m = -\Delta S_m h_m. \quad (8.9)$$

V týchto úpravách sme využili vzťah (8.1) a fakt, že prahy excitácie sú nulové, takže platí $h_m = h_m^{int}$. Spomeňme si na prechodové pravidlo (8.2), definíciu funkcie $\text{sign}(x)$, a môžeme analyzovať všetky štyri možnosti zmeny stavu jedného neurónu, a čo z toho vyplýva pre zmenu energie:

$$\begin{aligned} (1) \quad S_m = +1 \quad a \quad S'_m = +1 \quad (\text{vtedy } h_m > 0) &\Rightarrow \Delta E = 0 \\ (2) \quad S_m = -1 \quad a \quad S'_m = -1 \quad (\text{vtedy } h_m < 0) &\Rightarrow \Delta E = 0 \\ (3) \quad S_m = +1 \quad a \quad S'_m = -1 \quad (\text{vtedy } h_m < 0) &\Rightarrow \Delta E < 0 \\ (4) \quad S_m = -1 \quad a \quad S'_m = +1 \quad (\text{vtedy } h_m > 0) &\Rightarrow \Delta E < 0 \end{aligned} \quad (8.10)$$

Dokázali sme, že pre asynchrónnu dynamiku, keď je matica váh symetrická, t.j. ($J_{ij}=J_{ji}$), a $h_i^{ext}=0$ pre $\forall i$, platí vždy $\Delta E \leq 0$, a teda energia Hopfieldovej siete v tvare (8.6) v priebehu relaxácie monotónne klesá, pokiaľ nedosiahne minimum.

Existencia viacerých rôznych bodových atraktorov je dôsledkom frustrácie väzieb medzi neurónmi. Frustrácia väzieb znamená, že neexistuje taká konfigurácia siete, v ktorej by stavy jednotlivých neurónov zodpovedali polarite všetkých svojich väzieb. To bráni tomu, aby sieť dospela do jediného stavu, v ktorom by mala energia (8.6) svoje globálne minimum, ako by sa stalo v nefrustrovanom systéme. Frustrácia väzieb teda spôsobuje rôznosť atraktorov, v ktorých má energia siete (8.6) svoje lokálne minimum. V stavovom priestore existujú aj tzv. **falošné atraktory** (viď obr. 8.3). Konfigurácie aktivity siete prislúchajúce falošným atraktorom sú rôznymi lineárnymi kombináciami konfigurácií prislúchajúcich pravým atraktorom. Falošné atraktory ležia energeticky vyššie ako pravé atraktory a sú obklopené bariérami stavov s vyššou energiou. Ak je dynamika siete taká, že monotónne minimalizuje energiu (8.6), môže sa stať, že sa systém ocitne vo falošnom atraktore, a nemôže sa dostať von.

Iste je namieste otázka, ktorý typ dynamiky, synchronná (8.4) alebo asynchrónna (8.5), vernejšie vystihuje modelovanú skutočnosť — prácu neurónov v mozgu. Stochastická povaha činnosti neurónov je lepšie vystihnúť v asynchrónnej dynamike, s náhodným výberom jednotlivých neurónov, ktoré aktualizujú svoj stav [63]. Jeden cyklus relaxácie Hopfieldovej siete možno interpretovať nasledovným spôsobom. Absolútna refraktérna doba v neurobiológii je minimálna doba kľudu medzi generovaním dvoch akčných potenciálov za sebou. Je daná charakteristikami neurónovej membrány a jej trvanie je $\Delta t = 1-2$ ms [46]. Pri modelovaní na počítači možno tento interval rozdeliť na N subintervalov, t.j. $\Delta t = N\delta t$. Po uplynutí každého δt sa s pravdepodobnosťou $1/N$ vyberie jeden neurón, ktorý bude aktualizovať svoj stav podľa vzťahu (8.5). Takto zabezpečíme, že medzi dvoma po sebe nasledujúcimi okamihmi generovania akčného potenciálu má každý neurón prestávku $\Delta t = 1-2$ ms.

8.4 Autoasociatívna pamäť

Všetky modelované kognitívne udalosti (pamäť, učenie, atď.) sa odohrávajú na úrovni celej siete, ktorá reprezentuje nejakú **populáciu neurónov** určenú pre danú úlohu. Opakujúci sa vzorec aktivity (**bodový atraktor**) je stacionárnym stavom siete. **Stacionárne stavy** siete predstavujú **pamäťové stavy** siete. To, ktorý stav siete je bodovým atraktorom, čiže jej pamäťovým stavom, je determinované maticou váh synaptických spojení \mathbf{J} . Učenie, t.j. reprezentácia nových vzorov v sieti, je spojené so zmenami hodnôt synaptických váh. Táto vlastnosť súhlasí so súčasnou predstavou o tom, že mechanizmus učenia sa a uchovávanía informácií v mozgu spočíva v zmenách účinnosti synaptických spojení medzi neurónmi [14, 26, 46].

Hopfieldova sieť dospeje do tej ktorej konfigurácie (zodpovedajúcej príslušnému bodovému atraktoru) na základe podobnosti tejto výslednej konfigurácie s počiatočnou konfiguráciou siete vyvolanou externou (vstupnou) stimuláciou. Ako sme už spomenuli, každému atraktoru (pravému i falošnému) prislúcha tzv. spádová oblasť, čo je priestor stavov, v ktorom keď sa sieť ocitne, vždy sa deterministicky vyvinie do príslušného atraktora (obr. 8.3). To znamená, že neurónová sieť je schopná nájsť príslušný atraktor (pamäťový stav) aj vtedy, keď sa jej prezentuje neúplný, resp. deformovaný vstupný vzor. Táto vlastnosť sa nazýva **autoasociatívna (obsahom adresovaná) pamäť** (angl. *content addressable memory*), a je dominantnou vlastnosťou idealizovaných neurónových sietí Hopfieldovho typu.

Z hľadiska použitia Hopfieldovej siete ako modelu kognitívnych funkcií je veľmi dôležité, že synaptické váhy sa dajú navrhnuť (skonštruovať) tak, aby sa bodovými atraktormi stali vopred vybrané konfigurácie siete. Špecifický predpis pre konštrukciu váh J_{ij} je tento [9, 40]:

$$J_{ij} = \begin{cases} \frac{1}{N} \sum_{\mu=1}^p \xi_i^{\mu} \xi_j^{\mu} & \text{pre } i \neq j, \\ 0 & \text{pre } i = j. \end{cases} \quad (8.11)$$

N -rozmerné binárne vektory $\xi^{\mu} = (\xi_1^{\mu}, \xi_2^{\mu}, \dots, \xi_N^{\mu})$, $\mu = 1, \dots, p$, predstavujú zvolené pamäťové konfigurácie a ich celkový počet je p . J_{ij} môže nadobudnúť $2p+1$ rôznych hodnôt z intervalu $\langle -p/N, +p/N \rangle$. Tento predpis je jedným z viacerých možných formálnych vyjadrení tzv. Hebbovho pravidla [36] pre zmenu synaptických váh. **Hebbovo pravidlo** hovorí, že váha synapsy rastie, ak oba neuróny spojené touto synapsou sú zároveň aktívne, a naopak váha synapsy klesá, ak je aktivita týchto dvoch neurónov nekorelovaná (nie je synchronná). V procese učenia sa (resp. zapamätávania si viacerých vzorov) sa jednotlivé synapsy modifikujú podľa časového spriemernenia minulej aktivity neurónov. Formálne to možno vyjadriť takto: $\Delta J_{ij} \propto \langle \mathcal{S}_i \mathcal{S}_j \rangle_t$. Symbol $\langle \dots \rangle_t$ značí časové spriemernenie súčinu aktivity pre- a postsynaptického neurónu počas intervalu nejakého minulého časového obdobia dĺžky t . V prípade nášho predpisu (8.10) je zmena ΔJ_{ij} lineárna, t.j. $\Delta J_{ij} = (1/N) \xi_i^{\mu} \xi_j^{\mu}$. Proces učenia v atraktorových sieťach predstavuje samostatnú bohatú problematiku teoretického štúdia, ktorá nie je ešte ani zďaleka vyčerpaná [3, 4, 5, 23, 67].

Nech sa sieť vo svojej evolúcii dostane do stavu zodpovedajúcemu jednej z pamäťových konfigurácií, napr. ξ^v , čo znamená, že pre $\forall i$ platí $S_i = \xi_i^v$. Interpretujeme to ako vybavenie si (angl. *recall*) daného vzoru z pamäti (obr. 8.4). Vzniká otázka, či skutočne predpis na konštrukciu synaptických váh (8.11) garantuje, že zvolený vzor ξ^v je naozaj stabilný vzor. Inými slovami, keď sa sieť do takéhoto stavu dostane, nemôže sa z neho dostať von? Podmienka, že určitý stav ξ^v je dynamicky stabilný je taká, že lokálne pole i -teho neurónu musí mať to isté znamienko ako stav i -teho neurónu. **Podmienka stability** pre ľubovoľný vzor ξ^v vyjadrená matematicky je (viď rovnicu 8.3)

$$\xi_i^v h_i^v > 0 \quad \text{pre } \forall i . \quad (8.12)$$

Prepíšeme vzťah (8.12) tak, že do ľavej strany tejto nerovnosti najskôr dosadíme vzťah (8.1), a do tohto vzťahu namiesto J_{ij} dosadíme vzťah (8.11):

$$\xi_i^v h_i^v = \xi_i^v \sum_{j=1}^N J_{ij} \xi_j^v = \xi_i^v \frac{1}{N} \sum_{j=1}^N \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu \xi_j^v = \sum_{\mu=1}^p \xi_i^v \xi_i^\mu \left(\frac{1}{N} \sum_{j=1}^N \xi_j^\mu \xi_j^v \right) > 0 . \quad (8.13)$$

Predposlednú sumu na ľavej strane rozdelíme na dva členy, prvý zodpovedajúci vzoru $\mu = v$, a druhý zodpovedajúci ostatným vzorom $\mu \neq v$, takto:

$$\begin{aligned} \xi_i^v h_i^v &= \sum_{\mu=1}^p \xi_i^v \xi_i^\mu \left(\frac{1}{N} \sum_{j=1}^N \xi_j^\mu \xi_j^v \right) = \xi_i^v \xi_i^v \left(\frac{1}{N} \sum_{j=1}^N \xi_j^v \xi_j^v \right) + \sum_{\mu \neq v}^p \xi_i^v \xi_i^\mu \left(\frac{1}{N} \sum_{j=1}^N \xi_j^\mu \xi_j^v \right) \\ &= \frac{N-1}{N} + \sum_{\mu \neq v}^p \xi_i^v \xi_i^\mu \left(\frac{1}{N} \sum_{j=1}^N \xi_j^\mu \xi_j^v \right) > 0 . \end{aligned} \quad (8.14)$$

V limite pre veľké N dostaneme

$$\xi_i^v h_i^v = 1 + \xi_i^v \sum_{\mu \neq v}^p \xi_i^\mu \left(\frac{1}{N} \sum_{j=1}^N \xi_j^\mu \xi_j^v \right) = 1 + C_i^v > 0 . \quad (8.15)$$

Druhý člen, C_i^v , predstavuje tzv. presluch (angl. *crossstalk*). Ak je $C_i^v = 0$, môžeme vysloviť záver, že vzor ξ^v spĺňa podmienku stability (8.12). C_i^v sa rovná nule vtedy, keď sú pamäťové vzory navzájom ortogonálne, t.j. keď ich skalárny súčin je rovný nule, teda

$$\xi^\mu \cdot \xi^v = \frac{1}{N} \sum_{j=1}^N \xi_j^\mu \xi_j^v = 0 \quad \text{pre } \mu \neq v . \quad (8.16)$$

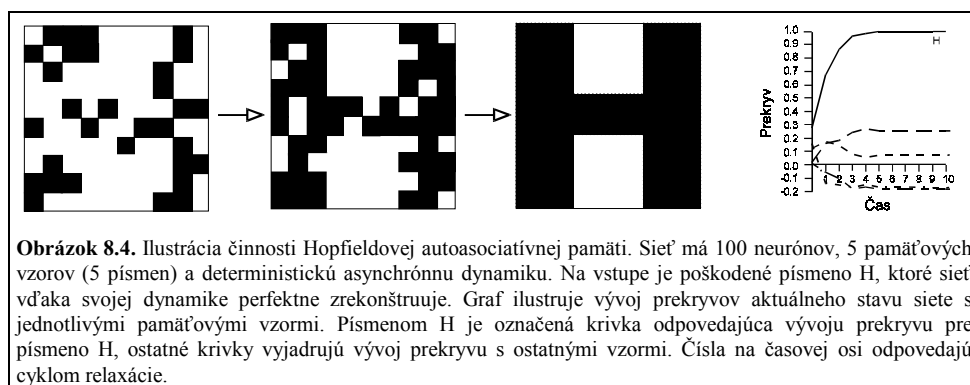
Keď sú pamäťové vzory navzájom ortogonálne, maximálny počet vzorov, ktoré možno uchovať v Hopfieldovej sieti (pamäťová kapacita) $p_{\max} \rightarrow N$. Podmienka stability (8.12) je splnená aj v prípade, keď je $|C_j^v| < 1$. Táto situácia nastáva vtedy, keď sú pamäťové vzory pseudoortogonálne (t.j. keď stredná hodnota ich skalárnych súčinov, ustrednená cez všetky dvojice vzorov $\langle \langle \xi^\mu \cdot \xi^\nu \rangle \rangle$, je približne rovná nule) a zároveň keď $p \ll N$. Z toho vyplýva, že takisto $p_{\max} \ll N$. V týchto dvoch prípadoch sú všetky pamäťové vzory stabilné, t.j. keď je sieť naštartovaná v ktoromkoľvek z nich, aj v ňom zostane. Okrem toho je sieť schopná opraviť určité percento neurónov, ktoré sa na začiatku nachádzajú v nesprávnych stavoch, takže sieť relaxuje do správneho pamäťového vzoru (vid' obr. 8.4). To znamená, že vybrané pamäťové vzory sú naozaj atraktormi systému, ktorý funguje ako autoasociatívna pamäť.

Ako **pozorovateľná premenná** sa pri Hopfieldových neurónových sieťach najčastejšie volí prekryv $m^\mu(t)$ okamžitej konfigurácie siete $S(t)$ s pamäťovými stavmi ξ^μ , definovaný ako

$$m^\mu(t) = \frac{1}{N} \sum_{j=1}^N \xi_j^\mu S_j(t) \quad , \quad \text{pre } \mu = 1, \dots, p. \quad (8.17)$$

Prekryv $m^\mu(t)$ je vlastne **miera podobnosti** dvoch konfigurácií. Pomocou časového vývoja okamžitých prekryvov s každým z pamäťových stavov (8.17) môžeme sledovať časovú evolúciu siete. Pamäťové stavy ξ^μ sú atraktormi deterministickej siete a sú v sieti uložené pomocou predpisu (8.11). Po krátkom čase sa sieť dostane do jedného z atraktorov a príslušný prekryv bude $m^\nu = 1$.

Pamäťová kapacita siete p_{\max} sa dá odvodiť nasledovným spôsobom. C_j^v z rovnice (8.15) závisí len na pamäťových vzoroch ξ^μ , ktoré chceme uchovať v sieti. Uvažujme, že



Obrázok 8.4. Ilustrácia činnosti Hopfieldovej autoasociatívnej pamäti. Sieť má 100 neurónov, 5 pamäťových vzorov (5 písmen) a deterministickú asynchrónnu dynamiku. Na vstupe je poškodené písmeno H, ktoré sieť vďaka svojej dynamike perfektne zrekonštruje. Graf ilustruje vývoj prekryvov aktuálneho stavu siete s jednotlivými pamäťovými vzormi. Písmenom H je označená krivka odpovedajúca vývoju prekryvu pre písmeno H, ostatné krivky vyjadrujú vývoj prekryvu s ostatnými vzormi. Čísla na časovej osi odpovedajú cyklom relaxácie.

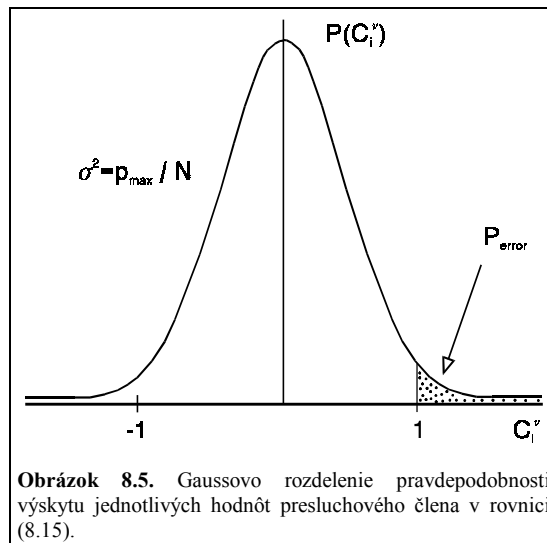
tieto vzory sú čisto náhodné konfigurácie, teda že stavy $\xi_j^\mu = +1$ a $\xi_j^\mu = -1$ sú generované s rovnakou pravdepodobnosťou. Potom môžeme určiť pravdepodobnosť P_{error} , že ľubovoľný bit (neurón) je nestabilný, ako

$$P_{error} = P(C_i^v > 1) . \quad (8.18)$$

P_{error} závisí na počte neurónov N a počte vzorov p . Predpokladajme, že $N \gg 1$ aj $p \gg 1$, čo je typický prípad a umožňuje nám to použiť vzťahy z matematickej štatistiky. Potom C_i^v je $1/N$ krát suma približne Np náhodných čísiel, z ktorých každé má hodnotu $+1$ alebo -1 (pozri vzťah (8.15)). V matematickej štatistike [25] bolo odvodené, že takáto náhodná premenná má binomické rozloženie pravdepodobnosti so strednou hodnotou nula a s varianciou $\sigma^2 = p/N$. Ale keďže Np je dostatočne veľké, môžeme binomické rozdelenie aproximovať Gaussovým rozdelením s nulovou strednou hodnotou a tou istou varianciou (obr. 8.5).

P_{error} sa rovná veľkosti vybodkovanej oblasti pod grafom Gaussovej krivky na obr. 8.5, a tak môžeme písať

$$P_{error} = \frac{1}{\sigma\sqrt{2\pi}} \int_1^{\infty} e^{-x^2/2\sigma^2} dx = \frac{1}{2} \left[1 - \operatorname{erf}(1/\sqrt{2\sigma^2}) \right] = \frac{1}{2} \left[1 - \operatorname{erf}(\sqrt{N/2p}) \right] . \quad (8.19)$$



Chybová funkcia (angl. *error function*) $\operatorname{erf}(x)$ je definovaná ako

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-u^2) du . \quad (8.20)$$

V tabuľke 8.1 sú uvedené hodnoty p_{max}/N prislúchajúce jednotlivým hodnotám P_{error} . Napríklad ak si zvolíme, že ľubovoľný bit sa môže preklopiť s pravdepodobnosťou $P_{error} < 0,01$, dostávame $p_{max} \leq 0,185N$. Toto číslo tiež hovorí, že asi 1,85% neurónov bude spočiatku nestabilných, keď štartujeme sieť z niektorého pamäťového vzoru.

Tabuľka 8.1. Pravdepodobnosť P_{error} , že ľubovoľný neurón je nestabilný, vo vzťahu k pamäťovej kapacite Hopfieldovej siete p_{max}/N .

P_{error}	p_{max}/N
0,001	0,105
0,0036	0,138
0,01	0,185
0,05	0,37
0,1	0,61

Sofistikovanejšia analýza situácie by ukázala [38], že Hopfieldova sieť funguje skutočne spoľahlivo ako autoasociatívna pamäť pre $p_{\text{max}} = 0,138N$. Ak sa budeme pokúšať zapamätať v sieti viac vzorov, výsledkom bude "výpadok pamäti" (angl. *blackout*) a môže sa nám stať, že sieť si nebude pamätať žiaden z predpísaných vzorov. Analýza kapacity autoasociatívnych neurónových sietí, čiže maximálneho počtu pamäťových konfigurácií, ktoré možno v sieti uchovať tak, aby boli vybaviteľné z pamäti, tiež tvorí bohatú časť ich štúdiá [19, 20, 28-31, 47]. Kapacita siete sa vyšetruje napr. vzhľadom na to, do akej miery sú pamäťové vzory navzájom korelované, ďalej vzhľadom na predpis ich uchovania, ktorý nemusí mať vždy tvar (8.11), atď. Napríklad do tohto predpisu možno zabudovať kontinuálne učenie a zabúdanie [59].

V tomto kontexte teda vybavenie si nejakého vzoru z pamäti predstavuje proces, v ktorom sieť znovu vytvára (obnovuje) tú istú aktivitu neurónov, ktorá sa v minulosti opakovane šírila v sieti pri zapamätávaní si daného vzoru. Vzorec aktivity siete (pamäťová konfigurácia siete) je vnútornou reprezentáciou nejakého vonkajšieho podnetu. Tento vonkajší podnet môže byť "jednoduchý" vstup z nejakého senzorického orgánu alebo komplexný vstup zložený z aktivity prichádzajúcej z jednej alebo viacerých iných neurónových sietí.

8.5 Stochastický Hopfieldov model

Ukázali sme, že Hebbovo pravidlo učenia (8.11) nám dáva (pre dostatočne malé p) dynamický systém s bodovými atraktormi, ktoré sú totožné s vopred zvolenými stavmi siete ξ^{μ} a ktoré zodpovedajú lokálnym minimám energie systému (8.6). Avšak tieto stavy nie sú jedinými atraktormi systému. Po prvé, všetky **reverzné konfigurácie** $-\xi^{\mu}$ sa automaticky tiež stávajú pamäťovými stavmi, ktoré zodpovedajú tým istým minimám energie ako ξ^{μ} . Je to preto, lebo predpis pre konštrukciu synaptických váh (8.11) a výraz pre energiu systému (8.6) sú oba perfektne symetrické čo sa týka zámenny $S_i \leftrightarrow -S_i$. Existencia týchto atraktorov odpovedajúcich reverzným stavom nás až tak netrápi a považujeme ich za pravé atraktory.

Po druhé, sú tu však aj tzv. **zmiešané stavy** ξ^{mix} , ktoré sa nerovnajú ani jednému zapamätanému vzoru, ale lineárnej kombinácii nepárneho počtu pamäťových vzorov. Najjednoduchší prípad je takáto symetrická kombinácia troch pamäťových vzorov:

$$\xi_i^{\text{mix}} = \text{sign}(\pm \xi_i^{\mu_1} \pm \xi_i^{\mu_2} \pm \xi_i^{\mu_3}), \quad \text{pre } \forall i. \quad (8.21)$$

Vďaka rôznym kombináciám znamienok dostaneme osem rôznych kombinácií troch pamäťových vzorov. Na to, aby sme si overili, či zmiešaný stav (8.21) je skutočne stabilný, treba overiť či platí (8.12). Napríklad pre takú kombináciu (8.21), kde sa vyskytujú len znamienka +, môžeme dospieť k takémuto výsledku:

$$\begin{aligned} \xi_i^{mix} h_i^{mix} &= \xi_i^{mix} \frac{1}{N} \sum_{\mu=1}^3 \sum_{j=1}^N \xi_i^{\mu} \xi_j^{\mu} \xi_j^{mix} \\ &\approx (\xi_i^1 + \xi_i^2 + \xi_i^3) \text{sign}(\xi_i^1 + \xi_i^2 + \xi_i^3) + \text{presluch} > 0. \end{aligned} \quad (8.22)$$

Môžeme vidieť, že podmienka stability (8.12) je pre tento zmiešaný stav splnená (ak je p dostatočne malé). Podobne môžeme kombinovať 5, 7, a viac vzorov. Systém si nevyberá ako falošné atraktory kombinácie z párneho počtu pamäťových vzorov, lebo ich príspevky sa môžu pre niektoré i vynulovať (pozri vzťah 8.21), čo nie je prípustné, keďže stavy neurónov sú ± 1 .

Po tretie, v stavovom priestore existuje ešte jeden druh falošných atraktorov, ktoré nie sú kombináciou žiadneho konečného počtu pamäťových vzorov [9-11]. Tieto stavy sa nazývajú **stavy spinového skla**, kvôli ich príbuznosti so stabilnými stavmi spinových skiel.

Ako vidíme, takýto model pamäti nie je dokonalý. Okrem nami vytvorených minim energie sa tam objavujú aj minimá zodpovedajúce rozličným ďalším stavom. Hoci teória a simulácie ukázali, že spádové oblasti týchto falošných atraktorov sú oveľa užšie ako spádové oblasti pravých atraktorov, je veľmi žiaduce sa ich zbaviť. Veľmi účinným spôsobom ako sa zbaviť falošných atraktorov je zaviesť do systému šum. Pritom však pravé atraktory zostávajú naďalej dobrými atraktormi. Okrem toho, **zavedenie šumu** predstavuje aj ďalšie priblíženie sa biologickej realite, pretože reálne neuróny pracujú v "zašumenom" prostredí. Zdroje šumu sú rôzne: stochastická povaha uvoľňovania neuromediátorov, fluktuácie postsynaptických potenciálov, spontánne generovanie akčných potenciálov, atď. Je zaujímavé, že šum, ktorý v iných informačných systémoch máva negatívne dôsledky, a preto sa ho snažíme eliminovať, v Hopfieldových neurónových sieťach plní vyslovene pozitívnu úlohu a pritom ich viac približuje k biologickej realite.

Uvažujme teda Hopfieldovu neurónovú sieť, v ktorej sa stavy neurónov nemenia podľa deterministického prechodového pravidla (8.3), ale podľa nejakého pravdepodobnostného (stochastického) pravidla. Toto stochastické pravidlo si teraz odvodíme. Majme v systéme nejaký bližšie nešpecifikovaný zdroj šumu, ktorého sila je parametrizovaná parametrom T , čo je analógia teploty termostatu, s ktorým je daný fyzikálny systém v kontakte. Na popis Hopfieldovho modelu pri nenulovej teplote T použijeme štatisticko-fyzikálny prístup [9, 38]. Predpokladajme, že systém po určitom čase dospeje do rovnovážneho stavu. V rovnovážnom stave je systém popísaný distribučnou funkciou Gibbsovoho kanonického systému. Teda pravdepodobnosť výskytu konfigurácie \mathbf{S} je rovná

$$P(\mathbf{S}) = \frac{1}{Z} \exp(-\beta(\mathbf{S})), \quad \text{kde} \quad Z = \sum_{\mathbf{S}} \exp(-\beta E(\mathbf{S})) = \exp(-\beta E) + \exp(-\beta E') \quad (8.23)$$

je stavová suma (partičná funkcia), a $\beta = 1/T$. Keďže ľubovoľný stav \mathbf{S}' , ktorý vznikol so stavu \mathbf{S} , sa líši len v stave jedného neurónu m , tak pravdepodobnosť prechodu do

ľubovoľného stavu \mathbf{S}' , ktorý vznikol zo stavu \mathbf{S} preklopením m -tého neurónu, je daná vzťahom

$$P(\mathbf{S} \rightarrow \mathbf{S}') = \frac{\exp(-\beta E')}{\exp(-\beta E') + \exp(-\beta E)} = \frac{1}{1 + \frac{\exp(-\beta E)}{\exp(-\beta E')}} = \frac{1}{1 + \exp(\beta \Delta E)}, \quad (8.24)$$

kde zjavne $E = E(\mathbf{S})$, $E' = E'(\mathbf{S}')$, a $\Delta E = E' - E$. Teraz si spomeňme na vzťahy (8.9) a (8.10) a môžeme písať

$$\Delta E = E' - E = (\mathbf{S}_m - \mathbf{S}'_m) h_m = \begin{cases} -2h_m \mathbf{S}'_m & \text{ak } \mathbf{S}_m = -\mathbf{S}'_m, \\ 0 & \text{ak } \mathbf{S}_m = \mathbf{S}'_m. \end{cases} \quad (8.25)$$

Pomocou vzťahu (8.25) teraz vyjadríme $P(\mathbf{S} \rightarrow \mathbf{S}')$ takto

$$P(\mathbf{S} \rightarrow \mathbf{S}') = \frac{1}{1 + \exp(-2\beta h_m \mathbf{S}'_m)}. \quad (8.26)$$

Tento vzťah zovšeobecňme pre ľubovoľný stav \mathbf{S} , ktorý vznikol aktualizáciou m -tého neurónu. Zároveň si uvedomme, že táto pravdepodobnosť je vlastne totožná s pravdepodobnosťou, že m -tý neurón zmení svoj stav, a teda

$$P(\mathbf{S}_m = \pm 1) = \frac{1}{1 + \exp(-2\beta h_m \mathbf{S}_m)}. \quad (8.27)$$

Rovnica (8.27) je **stochastické pravidlo** na zmenu stavu jedného neurónu. Z tejto rovnice vyplýva

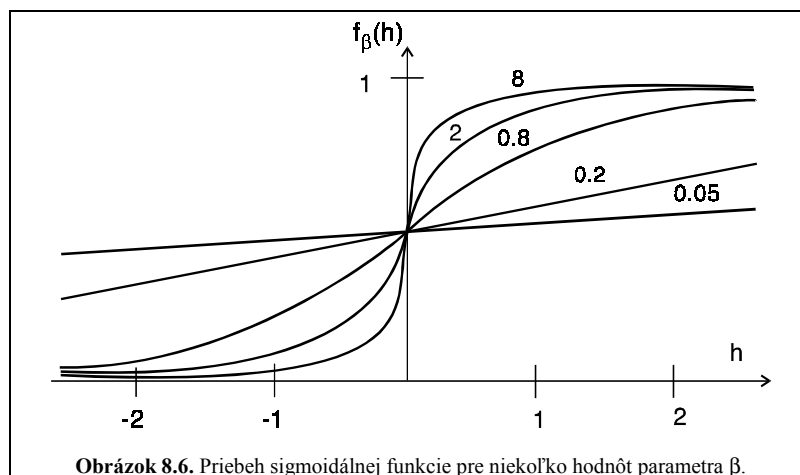
$$P(\mathbf{S}_m = +1) = \frac{1}{1 + \exp(-2\beta h_m)}, \quad (8.28)$$

a tiež

$$P(\mathbf{S}_m = -1) = \frac{1}{1 + \exp(+2\beta h_m)} = 1 - P(\mathbf{S}_m = +1). \quad (8.29)$$

Vzťah (8.28) je analógiou tzv. Glauberovej dynamiky v Isingovom modeli [32]. Vyjadruje priebeh sigmoidálnej (logistickej) funkcie (obr. 8.6).

V limite pre $T \rightarrow 0$ a $\beta \rightarrow \infty$, sa (8.28) mení na deterministické pravidlo (8.3). V limite $T \rightarrow \infty$ a $\beta \rightarrow 0$ je $P(\mathbf{S}_m = +1) = 0,5$. Všetky stavy sú rovnako pravdepodobné, systém sa stáva ergodickým a nemôže ďalej pracovať ako autoasociatívna pamäť. Z toho vyplýva, že musí existovať nejaký interval teplôt (hodnôt šumu), pre ktorý je stochastická Hopfieldova sieť stále spoľahlivou autoasociatívnou pamäťou.



Obrázok 8.6. Priebeh sigmoidálnej funkcie pre niekoľko hodnôt parametra β .

V rovnovážnom stave pri nenulovej teplote systém nedospeje do jedného stabilného stavu, do jednej stabilnej konfigurácie. Očakáva sa, že pri dostatočne malých teplotách bude jeho rovnovážny stav charakterizovaný hodnotami parametrov usporiadania, ktoré sú definované ako stredné hodnoty termodynamických pozorovateľných premenných — prekryvov (8.17). Po istom počte prechodov, keď sieť dospeje do rovnovážneho stavu, bude konfigurácia siete oscilovať okolo niektorého z atraktorov, ktorý je v stavovom priestore blízko počiatočnej konfigurácie. Parametre usporiadania sú časové stredné hodnoty prekryvov (8.17) stavov siete $\mathcal{S}(t)$ s danými atraktormi ξ^μ definované ako

$$\langle m^\mu(t) \rangle = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu \langle \mathcal{S}_i(t) \rangle, \quad \text{pre } \mu = 1, \dots, p. \quad (8.30)$$

Pre atraktor, ktorý si stochastická sieť "vyberie" býva stredná hodnota prekryvu (8.30) väčšia ako 0,9, ale nedosahuje hodnotu 1. Hodnota prekryvu (8.30) s ostatnými pamäťovými stavmi osciluje blízko nuly. V Hopfieldovej neurónovej sieti so šumom nemožno hovoriť o atraktórnych stavoch, ale len o atraktórnej pravdepodobnostnej distribúcii stavov.

Podmienky, ktoré umožňujú pretrvávanie rovnovážnej distribúcie (8.23), môžeme lepšie uvidieť, keď si odvodíme **selfkonzistentné rovnice**. Majme stochastickú sieť pozostávajúcu z N neurónov, ktorá má pri nulovej teplote $p \ll N$ pamäťových stavov ξ^μ , $m = 1, \dots, p$. Dynamika siete je asynchrónna s pravdepodobnosťou zmeny stavu i -teho neurónu (8.27), matica váh je symetrická a pre synaptické váhy platí predpis (8.11). V časovej strednej hodnote prekryvu (8.30) vystupuje časová stredná hodnota stavu i -teho neurónu $\langle \mathcal{S}_i(t) \rangle$, pre ktorú platí

$$\langle \mathcal{S}_i(t) \rangle = \sum_{\mathcal{S}_i} \mathcal{S}_i P(\mathcal{S}_i) = (+1) \times P(\mathcal{S}_i = +1) + (-1) \times P(\mathcal{S}_i = -1) = \tanh[\beta h_i(t)]. \quad (8.31)$$

Za pravdepodobnosti $P(\mathcal{S}_i)$ sme dosadili vzťahy (8.28) a (8.29). V ďalšom môžeme uvažovať všeobecne, že celkové pole $h_i(t) = h_i^{int}(t) + h_i^{ext}(t)$, kde $h_i^{ext}(t) = h_i^{ext}$. Ale na to, aby v rovnovážnom stave toto celkové pole pretrvávalo dlhšiu dobu, musia byť príspevky k jeho internej časti h_i^{int} (t.j. časti, ktorá je tvorená príspevkami od všetkých

ostatných neurónov v sieti) reprodukované strednými hodnotami všetkých neurónov spojených s miestom i . To znamená, že v (8.31) prejdeme k zámene skutočného fluktuujúceho poľa za jeho priemernú hodnotu $\langle h_i^{int}(t) \rangle$. Táto aproximácia je známa ako **teória stredného poľa** (angl. *mean-field approximation*). Sústreďujeme sa na jediný neurón a zanedbávame zmeny stavov individuálnych okolitých neurónov, ktoré berieme do úvahy len ako priemerné pozadie.

Pre $\langle h_i^{int}(t) \rangle$ platí:

$$\langle h_i^{int}(t) \rangle = \frac{1}{N} \sum_{j=1}^N J_{ij} \langle \mathcal{S}_j(t) \rangle = \frac{1}{N} \sum_{j=1}^N \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu \langle \mathcal{S}_j(t) \rangle = \sum_{\mu=1}^p \xi_i^\mu \langle m^\mu(t) \rangle. \quad (8.32)$$

Pri týchto úpravách sme využili vzťahy (8.11) a (8.30). Keďže $J_{ii}=0$, ale vzťah pre prekryv (8.30) obsahuje súčiny všetkých prvkov, pri prechode k strednej hodnote prekryvu chýba jeden člen prislúchajúci i -temu neurónu, rovný $(p/N) \langle \mathcal{S}_i(t) \rangle$. Môžeme ho však zanedbať, lebo $p \ll N$. Teraz dosadíme (8.31) do (8.32), a tento výraz dosadíme do vzťahu pre strednú hodnotu prekryvu (8.30). Dostávame sústavu N viazaných nelineárnych rovníc (8.33), známych pod názvom **selfkonzistentné rovnice**, alebo rovnice stredného poľa [9, 62]:

$$\begin{aligned} \langle h_i^{int}(t) \rangle &= \frac{1}{N} \sum_{j=1}^N J_{ij} \langle \mathcal{S}_j \rangle = \sum_{\mu=1}^p \xi_i^\mu \langle m^\mu(t) \rangle \\ &= \sum_{\mu=1}^p \xi_i^\mu \left\{ \sum_{j=1}^N \xi_j^\mu \tanh \left[\beta \left(\sum_{\mu=1}^p \xi_j^\mu \langle m^\mu(t) \rangle + h_j^{ext} \right) \right] \right\}. \end{aligned} \quad (8.33)$$

Zo selfkonzistentnosti vyplývajú pre rovnovážny stav tieto navzájom ekvivalentné podmienky:

- (1) Priemerné vnútorné pole všetkých neurónov $\langle h_i^{int}(t) \rangle$ (8.33) má práve takú hodnotu ako vnútorné pole, ktoré dalo vznik danej distribúcii stredných hodnôt stavov neurónov $\langle \mathcal{S}_i(t) \rangle$.
- (2) Stredné hodnoty stavov $\langle \mathcal{S}_i(t) \rangle$ vstupujú do vnútorného poľa práve s takými hodnotami, aké im dáva vnútorné pole, ktoré ich vyvolalo.
- (3) Keďže $\langle h_i^{int}(t) \rangle$ (8.33) závisí len na hodnotách $\langle m^\mu \rangle$, $\mu = 1, \dots, p$, tak tieto prekryvy sú reprodukované len tým vnútorným poľom $\langle h_i^{int}(t) \rangle$, ktoré ich vyvolalo.

V skutočnosti je simulovanie vybavovania vzorov z pamäti efektívne aj v tomto prípade, keď dynamika dovedie sieť iba do okolia uchovávaného vzoru a istá časť neurónov neustále mení svoj stav. Ak je časová stredná hodnota prekryvu meniacich sa konfigurácií s niektorým zo vzorov dostatočne vysoká, môžeme predpokladať, že výstupný vzorec aktivity siete reprezentuje vyvolanie daného vzoru z pamäti.

Riešeniami selfkonzistentných rovníc (8.33) sú parametre usporiadania (stredné hodnoty prekryvov). Parametre usporiadania charakterizujú rôzne fázy, v ktorých sa systém môže v rovnovážnom stave nachádzať v závislosti od teploty a histórie systému, t.j. stavu, z ktorého sieť vyštartovala. V ideálnom prípade, po dosiahnutí termodynamической rovnováhy, rozpoznanie fázy podľa parametrov usporiadania zodpovedá obnoveniu informácie

adresovanej obsahom resp. degradovanej chybami v počiatočnom stave. Systém môže slúžiť ako pamäť, ak vďaka dynamike na ľubovoľný podnet (počiatočný stav) spontánne prechádza do stavov asociovaných s atraktorom, a tieto stavy majú významný prekryv s týmto naučeným vzorom. Zjednodušene možno interpretovať systém s ideálnymi pamäťovými vlastnosťami ako systém, ktorý z nerovnovážneho počiatočného stavu prejde do termodynamickej rovnováhy v jednej z možných fáz systému. Každá fáza zodpovedá jednému naučenému vzoru a spontánna voľba fázy je determinovaná (pri konštantnej teplote) iba počiatočnou konfiguráciou.

Teraz vysvetlíme, ako je to možné, že šum T destabilizuje falošné atraktory. Uvažujme stochastickú Hopfieldovu neurónovú sieť s $p \ll N$, $N \rightarrow \infty$ a $h_i^{ext} = 0$ pre $\forall i$. Pre strednú hodnotu stavu i -teho neurónu môžeme selfkonzistentné rovnice (8.31) prepísať takto:

$$\langle S_i(t) \rangle = \tanh \left(\beta \sum_{j=1}^N J_{ij} \langle S_j \rangle \right) = \tanh \left(\frac{\beta}{N} \sum_{j=1}^N \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu \langle S_j(t) \rangle \right). \quad (8.34)$$

Táto sústava nie je vo všeobecnosti riešiteľná, lebo pozostáva z N nelineárnych viazaných rovníc s N neznámymi. Pri jej riešení však môžeme postupovať heuristicky a navrhnúť hypotézu, že riešenie $\langle S_i(t) \rangle$ je úmerné jednému zo zapamätaných vzorov, tak že

$$\langle S_i(t) \rangle = m \xi_i^v \quad \text{pre } \forall i. \quad (8.35)$$

Parameter $m \in \mathbf{R}$ je konštantou úmernosti. Už sme si ukázali, že takéto stavy sú stabilné v deterministickej limite $T = 0$ (pre $m = 1$), takže je prirodzené hľadať podobné priemerné stavy v stochastickom prípade. Po dosadení (8.35) do (8.34) dostávame

$$m \xi_i^v = \tanh \left(\frac{\beta}{N} \sum_{j=1}^N \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu m \xi_j^v \right). \quad (8.36)$$

Takisto ako v prípade deterministickej siete môžeme argument funkcie tanh rozložiť na dva členy; jeden prislúchajúci ξ_i^v a druhý člen vyjadrujúci presluch, obsahujúci prekryvy medzi ξ_i^v a ostatnými zapamätanými vzormi. Keďže $p \ll N$ (a pamäťové vzory sú pseudoortogonálne), môžeme presluchový člen zanedbať a dostaneme vzťah

$$m \xi_i^v = \tanh(\beta m \xi_i^v). \quad (8.37)$$

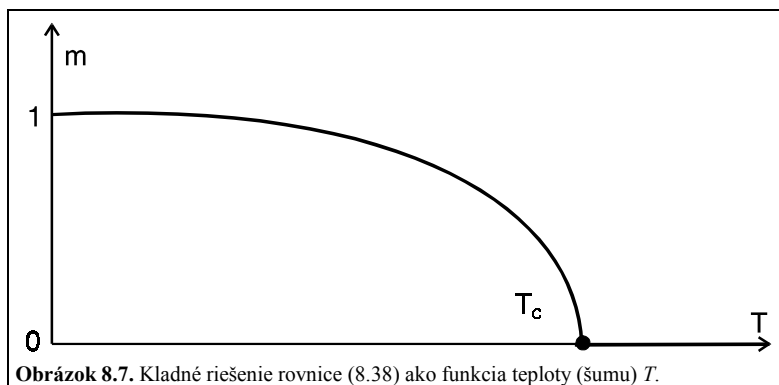
Keďže $\tanh(\tilde{x}) = \tilde{\tanh}(x)$, dostávame rovnicu na výpočet parametra m pre rovnicu (8.35):

$$m = \tanh(\beta m). \quad (8.38)$$

Grafické znázornenie riešenia rovnice (8.38) pre kladné m v závislosti od teploty T je znázornené na obr. 8.7. Z riešenia vidíme, že pre teploty väčšie ako **kritická teplota**, teda pre $T \geq T_c$ a $T_c = 1$, je $m = 0$ a žiadny pamäťový stav nie je stabilný¹. Inými slovami, teraz vieme, že pamäťové stavy budú stabilné pre teploty $T < 1$. Chceme poukázať na to, že v správaní sa stochastickej Hopfieldovej siete je prudká skokovitá zmena pri kritickej teplote T_c . Je to ďalší príklad fázového prechodu. Mohli by sme očakávať, že správanie sa siete sa

¹ Pre $T < T_c$ je $m = 0$ tiež riešením rovnice (8.38) avšak nie je to stabilné riešenie.

spojito mení ako teplota rastie, ale v systémoch s veľkým N sa takáto závislosť často nepozoruje. Keď sa prekročí istá úroveň šumu, veľká sieť prestane fungovať ako autoasociatívna pamäť.



Podarilo sa nám ukázať, že stavy so strednou hodnotou $\langle S_i(t) \rangle$ úmernou jedinému pamäťovému vzoru (viď (8.35)) sú pri nízkych teplotách $0 < T < 1$ stabilné. Avšak, hoci pre $p \ll N$ nie sú stavy spinového skla aktuálne, ešte stále sú tu reverzné stavy a hlavne zmiešané stavy. V prácach Amita so spolupracovníkmi [10, 11] bolo ukázané, že **každý zmiešaný stav má svoju vlastnú kritickú teplotu** v intervale $0 < T_c \leq 0,46$. Keď šum prekročí túto hodnotu, zmiešané stavy už nie sú stabilné. Inými slovami, **pre šum z intervalu $0,46 < T < 1$ sú falošné atraktory (zmiešané stavy) destabilizované**, a iba pamäťové stavy (a ich reverzné konfigurácie) sú stále dobrými atraktormi.

Cenou za to je nižšia vyvolávajúca kvalita pamäťových vzorov a o niečo dlhšie relaxačné časy siete (doba príchodu siete do okolia atraktora). Napriek všetkému, určitá úroveň šumu zlepšuje fungovanie Hopfieldovej neurónovej siete ako autoasociatívnej pamäti. Pre vyššie hodnoty šumu, a to $T \geq 1$, sa destabilizujú aj pravé atraktory (pamäťové stavy) a systém sa stáva ergodický.

8.6 Poškodzovanie a vymazávanie synaptických spojení

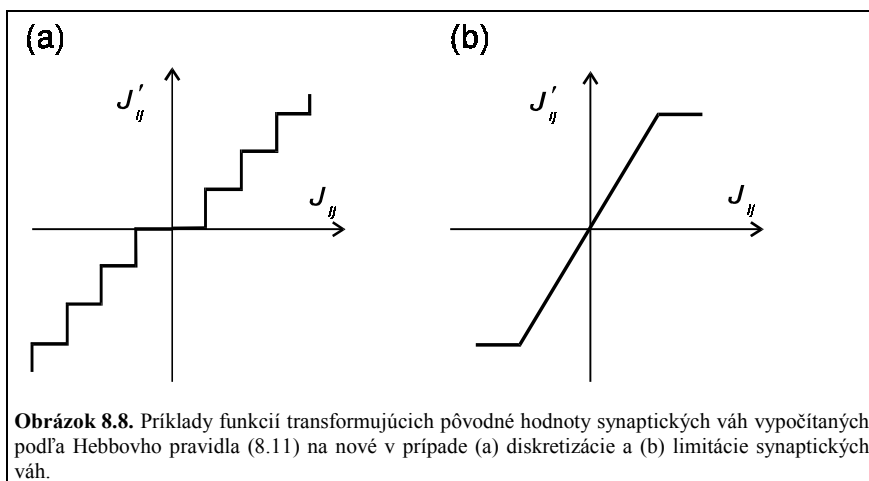
Skutočné neurónové siete v mozgu dokážu vykonávať svoju funkciu aj po poškodení ich prvkov a spojení, ak tieto poškodenia nie sú príliš súvislé a priestorovo rozsiahle. V tejto časti spomenieme výsledky štúdií, ktoré sa zamerali práve na tzv. **robustnosť** resp. odolnosť **Hopfieldových sietí voči nepresnostiam zaťažujúcim ich synaptické spojenia** vrátane ich vynulovania. Vo väčšine prípadov **kvalitatívne vlastnosti autoasociatívnej pamäti ostávajú nezmenené**, a poškodenie má iba kvantitatívne dôsledky, konkrétne ovplyvňuje najmä kapacitu siete $\alpha_c = p_{\max}/N$, predlžuje dobu príchodu do atraktora, konečný prekryv s pamäťovým vzorom je o niečo menší, a pod.

Ako prvé poškodenie budeme uvažovať deviaciu hodnôt synaptických váh od hodnôt získaných pomocou Hebbovho pravidla (8.11). Táto deviacia môže byť realizovaná ako **pridanie malého náhodného čísla** k hodnote J_{ij} vypočítanej podľa vzťahu (8.11). Dôsledkom je zníženie α_c , spomalenie nájdenia atraktora, ale i odstránenie falošných atraktorov typu stavov spinového skla [20, 61, 67]. **Diskretizácia** spojení znamená, že synapsy môžu nadobúdať len diskkrétne hodnoty (viď obr. 8.8a). Extrémnym prípadom diskretizácie spojení je ich **binarizácia**, t.j. synapsy môžu nadobúdať iba dve hodnoty:

$$J_{ij} = \begin{cases} \text{sign}\left(\sum_{\mu=1}^p \xi_i^{\mu} \xi_j^{\mu}\right) & \text{pre } i \neq j, \\ 0 & \text{pre } i = j. \end{cases} \quad (8.39)$$

Teoretickou analýzou tohoto modelu bolo ukázané, že sieť i naďalej spoľahlivo funguje ako autoasociatívna pamäť, len kapacita α_c sa zníži z hodnoty 0,138 na hodnotu 0,1 [37].

Limitácia znamená obmedzenie pre synaptické váhy v tvare $|J_{ij}| \leq A$. To znamená, že hodnotu synaptickej váhy počítame podľa vzťahu (8.11), a keď vypočítaná hodnota prekročí hranicu A , nahradíme ju touto hodnotou (viď obr. 8.8b).



Limitácia synáps poskytuje zaujímavé možnosti, čo sa týka modelovania učenia a zabúdania. Predstavme si, že záznam nových pamäťových vzorov do Hopfieldovej siete inkrementujeme tak, že pridanie každého nového vzoru μ prebieha podľa vzťahu

$$J_{ij} \leftarrow J_{ij} + \eta \xi_i^\mu \xi_j^\mu . \quad (8.40)$$

Koeficient η označuje rýchlosť synaptickej modifikácie (rýchlosť učenia). Pritom môžu synaptické váhy nadobúdať iba hodnoty z intervalu $(-A, +A)$. Takéto učenie sa nazýva **učenie v intervale** (angl. *learning within bounds*) [60, 61]. Posledne pridané pamäťové vzory sa budú dobre vyvolávať z pamäti, avšak staršie budú pomaly miznúť. Počet zapamätaných vzorov závisí na hodnotách η a A . Ak je hodnota η príliš veľká, iba posledne zapamätané vzory sú vyvolateľné. Ak je hodnota η príliš malá, sieť sa môže predimenzovať ($\alpha > \alpha_c$) a naplniť sa príliš veľa vzormi ešte predtým, ako váhy dosiahnu hranice intervalu A , a sieť prestane fungovať ako autoasociatívna pamäť. Pri optimálnych hodnotách η a A sa bude množina zapamätaných vzorov meniť tak, že pridávaním nových vzorov budeme vymazávať najstaršie zapamätané vzory. Takto sa dá modelovať **učenie a zabúdanie** [22, 60, 61].

Ďalší typ poškodenia je vymazanie alebo **zriedenie spojení** (angl. *dilution*). Náhodne vynulujeme konečný počet spojení medzi neurónmi tak, že

$$J_{ij} = \begin{cases} J_{ij}^{Hebb} & \text{s pravdepodobnosťou } c, \\ 0 & \text{s pravdepodobnosťou } (1-c). \end{cases} \quad (8.41)$$

J_{ij}^{Hebb} je hodnota synaptickej váhy vypočítaná podľa Hebbovho pravidla (8.11). Po zriedení bude percento zachovaných spojení rovné $100c$ a každý neurón bude mať približne cN spojení. Pri symetrickom riedení s hodnotou J_{ij}

vymazávame zároveň aj J_{ji} , pri asymetrickom riedení pristupujeme k J_{ij} a J_{ji} nezávisle. Nech $c_{ij}=1$, keď spojenie nie je vymazané a $c_{ij}=0$, keď sme spojenie zrušili. Potom pre každú váhu môžeme písať

$$J_{ij} = c_{ij} J_{ij}^{Hebb}. \quad (8.42)$$

Lokálne pole na i -tom neuróne je rovné

$$h_i = \sum_{j=1}^N c_{ij} J_{ij}^{Hebb} S_j. \quad (8.43)$$

Rozoznávame tzv. **slabé zriedenie**, keď $cN \geq \ln N$. **Silné zriedenie** nastáva, keď platí $cN \ll \ln N$. Pri slabom zriedení a veľkom N je ešte stále v sume (8.43) veľa členov, a preto môžeme aplikovať teóriu stredného poľa. Pre strednú hodnotu lokálneho potenciálu na i -tom neuróne dostaneme vzťah

$$\langle h_i \rangle = c \sum_{j=1}^N J_{ij}^{Hebb} \langle S_j \rangle. \quad (8.44)$$

S príslušným preškálovaním pomocou parametra c dostaneme analogické výsledky ako pre nezriedený model, v ktorom je každý neurón spojený s každým.

Silné zriedenie v deterministickej a stochastickej Hopfieldovej sieti ako prví študovali Derrida, Gardner a Zippelius [20]. Autoasociatívne vlastnosti sú zachované v jednom aj druhom prípade. Na príklade si to ukážeme pre deterministickú sieť. Nech priemerný počet spojení na jednom neuróne je $K=cN$ a $K \ll \ln N$. Zriedenie je asymetrické, takže c_{ij} a c_{ji} sú nezávislé náhodné premenné. Matica \mathbf{J} sa teda stala asymetrickou. Pre výpočet váh sa používa vzťah

$$J_{ij} = \frac{1}{K} c_{ij} \sum_{\mu=1}^p \xi_i^{\mu} \xi_j^{\mu}, \quad (8.45)$$

s $1/K$ namiesto $1/N$, aby sme dostali hodnoty blízke 1. Pre ľubovoľný stav $\mathbf{S}(t)$ uvažujme lokálne pole príspevkov od všetkých neurónov spojených s i -tým neurónom rozdelené na člen pochádzajúci od konkrétneho pamäťového vzoru v a presluchový člen:

$$h_i = \sum_{j=1}^N J_{ij} S_j = \frac{1}{K} \sum_{j=1}^N c_{ij} \sum_{\mu=1}^p \xi_i^{\mu} \xi_j^{\mu} S_j = \frac{1}{K} \xi_i^v \sum_{j=1}^N c_{ij} \xi_j^v S_j + C_i^v, \quad (8.46)$$

kde presluch

$$C_i^v = \frac{1}{K} \sum_{\mu \neq v} \xi_i^{\mu} \sum_{j=1}^N c_{ij} \xi_j^{\mu} S_j. \quad (8.47)$$

Ak v (8.46) položíme $S_j = \xi_j^v$ pre $\forall i$, vidíme, že pre dostatočne malý počet zapamätaných vzorov p nám prvý člen na pravej strane rovnice (8.46) dáva ξ_i^v , a to vďaka tomu, že sme ako normalizačnú konštantu zvolili $1/K$, čo nám umožňuje vykrátiť $1/K$ a sumu s c_{ij} , lebo $\sum_j c_{ij} = K$. Keď položíme $S_j = \xi_j^v$ pre $\forall i$ v rovnici pre presluchový člen (8.47), na pravej

strane dostaneme $1/K$ krát suma Kp nezávislých náhodných ± 1 . V limite pre $Kp \rightarrow \infty$ má presluchový člen Gaussovo rozdelenie (viď obr. 8.5) a ďalej platí podobný záver ako pre úplne pospájanú sieť, teda že pamäťové vzory sú stabilné pre $p \ll N$.

Zachovanie autoasociatívnych vlastností Hopfieldových sietí aj pri silnom rozriedení synaptickej matice umožňuje priblížiť sa k biologickej realite. Po prvé napríklad v tom, že v reálnych neurónových sieťach nie sú neuróny pospájané každý s každým monosynaptickými spojmi a po druhé v tom, že váhy spojení nie sú symetrické. Pomocou selektívneho riedenia možno skonštruovať Hopfieldovu neurónovú sieť, v ktorej je zachovaný tzv. Daleov princíp, ktorý hovorí, že jeden neurón môže vytvárať na druhých neurónoch len inhibičné alebo len excitačné synapsy, ale nie oba druhy naraz [66]. Zachovanie autoasociatívnych vlastností pri silnom rozriedení sa hodí aj pri konštrukcii viacvrstvových sietí s Hopfieldovou dynamikou, ktoré sa dajú použiť na hierarchické uchovávanie dát [34].

8.7 Neortogonálne vzory

Doteraz sme sa zaoberali Hopfieldovou neurónovou sieťou, ktorá je schopná zaznamenať a vyvolať ortogonálne resp. pseudoortogonálne pamäťové vzory, t.j. také pre ktoré platí, že stredná hodnota ich skalárneho súčinu cez všetky vzory sa rovná nule:

$$\langle\langle \xi^\mu \cdot \xi^\nu \rangle\rangle = \frac{1}{N} \sum_{j=1}^N \xi_j^\mu \xi_j^\nu = 0 \quad \text{pre } \mu \neq \nu. \quad (8.48)$$

Keď je v sieti veľký počet neurónov N , je možné vyhovieť tejto podmienke len ak vyberáme resp. konštruujeme pamäťové vzory tak, že v každom z nich sa počet aktívnych neurónov rovná počtu neaktívnych neurónov, t.j. že v sieti je 50%-ná priemerná aktivita, to znamená

$$\langle\langle \xi_j^\mu \rangle\rangle = \frac{1}{N} \sum_{j=1}^N \xi_j^\mu = 0 \quad \text{pre } \forall \mu. \quad (8.49)$$

O takýchto vzoroch hovoríme tiež, že sú nekorelované. V praxi však často pracujeme so vzormi, pre ktoré tieto podmienky nie sú splnené. Aj v neurónových populáciách, napríklad na kôre mozgu, je pomer počtu aktívnych neurónov k počtu neaktívnych neurónov podstatne nižší ako 50%. Pri modelovaní tejto skutočnosti sa stav každého neurónu ξ_j^μ v pamäťovom vzore vyberá nezávisle s pravdepodobnosťou $P(\xi) = \frac{1}{2}(1+b)\delta(\xi-1) + \frac{1}{2}(1-b)\delta(\xi+1)$. Symbol δ označuje Diracovu δ funkciu, kde $\delta(x) = 0$, ak $x \neq 0$, $\delta(x) = 1$, ak $x = 0$. Parameter b sa nazýva **bias** a platí preň vzťah

$$b = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu \quad \text{pre } \forall \mu. \quad (8.50)$$

Keďže b nadobúda hodnoty v intervale $-1 < b < +1$ môžeme pracovať s ľubovoľnou úrovňou aktivity. Stredná aktivita siete je rovná $\frac{1}{2}(1+b)$ a namiesto (8.48) platí

$\langle\langle \xi^\mu \cdot \xi^\nu \rangle\rangle = \delta^{\mu\nu} + b^2(1 - \delta^{\mu\nu})$ pre $\mu \neq \nu$. Pre takéto neortogonálne (korelované) vzory bola navrhnutá modifikácia dynamiky stochastickej Hopfieldovej neurónovej siete [7]. Predpis na konštrukciu synaptických váh (8.11) sa mení takto:

$$J_{ij} = \begin{cases} \frac{1}{N} \sum_{\mu=1}^p (\xi_i^\mu - b)(\xi_j^\mu - b) & \text{pre } i \neq j, \\ 0 & \text{pre } i = j. \end{cases} \quad (8.51)$$

Globálna kontrola dynamiky neurónovej siete sa dosiahne tak, že sa k energii (8.6) pridá člen, ktorý bude zabraňovať, aby sa sieť dostala do oblastí s aktivitou príliš vzdialenou pôvodnému biasu vzorov. Modifikovaná energia siete je rovná (prahy excitácie sú nulové)

$$E(\mathbf{S}) = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N J_{ij} S_i S_j + \frac{r}{2N} \left(\sum_{i=1}^N S_i - Nb \right)^2. \quad (8.52)$$

Koeficient $r > 0$ je mierou sily, ktorou je aktivita siete "tlačená" do oblasti určenej biasom vzorov. Lokálne pole pôsobiace na jeden neurón (8.1) sa modifikuje takto:

$$h_i = \sum_{j=1}^N J_{ij} S_j - r \left(\frac{1}{N} \sum_{j=1}^N S_j - b \right). \quad (8.53)$$

Vzťahy (8.50) až (8.53) zodpovedajú tzv. jemnému obmedzeniu dynamiky (angl. *softly constrained dynamics*). Táto modifikácia zachováva všetky hlavné vlastnosti pôvodného stochastického Hopfieldovho modelu. Okrem toho platí:

- (1) Ak je koeficient $r \ll 10$, dostávame spoľahlivé fungovanie siete pre akúkoľvek hodnotu biasu b .
- (2) Pamäťová kapacita siete s modifikovanou dynamikou sa pohybuje v intervale: $0,12N < p_{\max} < 0,18N$ pre $r \ll 10$. Konkrétna hodnota kapacity závisí od hodnoty biasu b .
- (3) Kritická teplota závisí od biasu b . Pri teplotách nižších ako T_c sú všetky pamäťové vzory stabilné.

$$T_c = (\tilde{\Gamma} b^2)^2. \quad (8.54)$$

Dokonca aj pri veľmi nízkych hodnotách T sú falošné atraktory (zmiešané stavy) eliminované.

- (4) Symetria $S_i \leftrightarrow -S_i$ je zrušená.

8.8 Časové postupnosti vzorov

Podnety, ktoré spracováva náš mozog, prichádzajú v časových postupnostiach. Príkladom je ľudská reč, ale i vizuálne prehľadávanie okolia, sled myšlienok, predstáv, atď. Na druhej

strane, mozog sám generuje časové postupnosti, napríklad pri generovaní povelov pre svaly, čo sa vlastne deje neustále. Tzv. centrálné generátory rytmu sú neurónové obvody, ktoré generujú biologické rytmy, od riadenia cyklu spánku a bdenia až po rytmické pohyby napríklad pri dýchaní. Štúdium umelých neurónových sietí sa snaží svojimi výsledkami tiež prispieť k pochopeniu podstaty týchto procesov. Pozrime sa teraz ako sa dajú rozpoznávať a generovať časové postupnosti konfigurácií aktivít v Hopfieldovej neurónovej sieti. Vieme, že v Hopfieldovej sieti reprezentuje konfigurácia aktivity celej populácie modelových neurónov nejaký ľubovoľný podnet, resp. pamäťový stav. Pamäťové stavy siete sú atraktormi v stavovom priestore a priťahujú im podobné stavy. Formulujme si úlohu spracovania časových postupností takto:

(1) Externý podnet spôsobí vyvolanie pamäťového vzoru, ktorý je s ním asociovaný. Nech je tento pamäťový stav členom nejakej postupnosti stavov.

(2) Neurónová sieť zotrúva v tomto atraktore, ale po uplynutí nejakej doby sa presunie do iného atraktora, ktorý je asociovaný s nasledujúcim členom v časovej postupnosti vzorov.

(3) Sieť sa postupne presúva z jedného atraktora do ďalšieho sledujúc tak preddefinovanú postupnosť stavov. V každom atraktore zotrúva istú dobu. Správnejšie by sa teda malo hovoriť o pseudo-atraktorech alebo kvázi-atraktorech, no pre jednoduchosť budeme i naďalej používať termín atraktor. Presun z atraktora do atraktora môže byť buď spontánny — vtedy sa jedná o **generovanie časovej postupnosti** stavov vyvolanej nejakým podnetom, alebo sa presúva z jedného atraktora do druhého na základe príchodu ďalšieho podnetu — vtedy sa jedná o **rozpoznanie časovej postupnosti**.

Podľa princípu, ktorý navrhli viacerí autori, menovite Kanter a Sompolinsky [47], Kleinfeld a Kanter [50], Gutfreund a Mézard [35], a Amit [8], dosiahneme takéto správanie sa autoasociatívnej neurónovej siete zavedením masívnej a koherentnej asymetrie do matice synaptických spojení medzi neurónmi. Majme sieť pozostávajúcu z N neurónov, ktoré môžu nadobúdať stav $+1$ alebo -1 . Sieť uchováva p pamäťových vzorov, ktorých konfigurácia aktivity je vygenerovaná náhodne, čiže v sieti je 50%-ná priemerná aktivita. Dynamika je asynchrónna, a aktualizovanie neurónov môže byť buď deterministické alebo stochastické. V sieti sú **dve množiny synáps** J_{ij} a J_{ij}^{Δ} . J_{ij} sa vypočíta podľa vzťahu (8.11). V tejto synaptickej matici je zapamätaných p dopredu zvolených vzorov a jej úlohou je oprava chýb a stabilizácia individuálnych vzorov. Druhá množina synáps J_{ij}^{Δ} má na starosti prechod medzi atraktormi v preddefinovanej postupnosti. Vypočíta sa podľa vzťahu

$$J_{ij}^{\Delta} = \frac{\lambda}{N} \sum_{\mu=1}^q \xi_i^{\mu+1} \xi_j^{\mu}, \quad (8.55)$$

kde $q \leq p$ je počet vzorov spojených v časovej postupnosti $\mu=1 \square 2 \square 3 \square \dots \square (q+1)$. J_{ij}^{Δ} definuje poradie q vzorov v postupnosti. Uzavretý cyklus sa dá zostrojiť pomocou (8.55) tak, že položíme $\xi_j^{q+1} = \xi_j^1$. Parameter λ je kladná konštanta. Nech sú prahy excitácie rovné nule. Celkové pole pôsobiace na i -ty neurón má tieto dve zložky

$$h_i(t) = \frac{1}{N} \sum_{j=1}^N \sum_{\mu=1}^p \xi_i^{\mu} \xi_j^{\mu} S_j(t) + \frac{\lambda}{N} \sum_{j=1}^N \sum_{\mu=1}^q \xi_i^{\mu+1} \xi_j^{\mu} S_j(t - \tau). \quad (8.56)$$

Prepíšme si (8.56) pomocou prekryvov

$$m^\mu(t) = \frac{1}{N} \sum_{j=1}^N \xi_j^\mu \mathcal{S}_j(t) \quad \text{a} \quad m^\mu(t-\tau) = \frac{1}{N} \sum_{j=1}^N \xi_j^\mu \mathcal{S}_j(t-\tau), \quad (8.57)$$

tak, že dostaneme

$$h_i(t) = \sum_{\mu=1}^p \xi_i^\mu m^\mu(t) + \lambda \sum_{\mu=1}^q \xi_i^{\mu+1} m^\mu(t-\tau). \quad (8.58)$$

Prvý člen závisí na okamžitých aktivitách neurónov v čase t , a teda na momentálnom prekryve stavu siete s jednotlivými pamäťovými vzormi. Druhý, **prechodový člen** závisí na tom, aká bola konfigurácia aktivity neurónov pred τ relaxačnými cyklami. J_{ij}^Δ sú tzv.

pomalé synapsy, alebo **prechodové synapsy**, lebo sprostredkujú neurónom informáciu o stave ich susedov s **časovým oneskorením** τ . Predstavme si, že sa sieť ocitne v stave μ , napríklad že doňho zrelaxuje po príchode nejakého podnetu. Keďže je to atraktor, sieť v ňom zotrúva. Avšak po uplynutí τ cyklov sa hodnota prekryvu $m^\mu(\tau)$ bude rovnáť 1 (alebo sa bude blížiť hodnote 1, podľa toho, či je dynamika deterministická alebo stochastická). Pozrime sa na vzťah pre lokálne pole i -teho neurónu (8.58) a vidíme, že ak je $\square \square 1$, sieť prejde do ďalšieho atraktora s poradovým číslom $\mu+1$. V tomto atraktore zotrúva opäť po dobu τ , a potom prejde do atraktora $\mu+2$, atď. V tomto prípade sa jedná o **spontánne generovanie časovej postupnosti** stavov siete, ktoré bolo iniciované nejakým vonkajším podnetom. **Rozpoznávanie časových postupností** bez spontánneho pokračovania v postupnosti dosiahneme tak, keď v (8.58) položíme $\square \square 1$ a ku každému lokálnemu poľu (8.58) pripočítame v časovom momente t príspevok z vonkajšieho signálu $h_i^{out} = \rho \xi_i^{\mu+1}$, kde $\rho \square 1$. Prechod do ďalšieho atraktora v časovej postupnosti nastane teda iba vtedy, keď sa tento stav (alebo jemu podobný stav) ocitne na vstupe do siete.

Vo vzťahu (8.56) môžeme namiesto $\mathcal{S}_j(t-\tau)$ použiť aj iné závislosti. Položme $\mathcal{S}_j(t-\tau) = \bar{\mathcal{S}}_j(t)$. Tzv. **pamäťová stopa** $\bar{\mathcal{S}}_j(t)$ môže mať vo všeobecnosti tvar $\bar{\mathcal{S}}_j(t) = \int_{-\infty}^t G(t-t') \mathcal{S}_j(t') dt'$, kde sme doteraz uvažovali funkciu $G(t)$ v tvare Diracovej delta funkcie $G(t) = \delta(t-\tau)$, ale môže to byť aj lineárna závislosť $G(t) = (\tau-t)/\tau$ alebo exponenciálna závislosť $G(t) = \tau^{-1} \exp(-t/\tau)$.

Vzťah (8.58) môže slúžiť na **rozpoznávanie viac ako jednej postupnosti** pomocou jednej a tej istej neurónovej siete len vtedy, ak každá postupnosť obsahuje iné členy μ . Gutfreund a Mézard [35] navrhli postup, ako môže Hopfieldova sieť rozpoznávať viacero rôznych časových postupností aj vtedy, keď každý stav μ má niekoľko možných pokračovaní. Tento postup si najprv ukážeme pre prípad, keď chceme, aby sieť uchovávala dve rôzne postupnosti, $\{\xi_i^{1,\mu}\}$ a $\{\xi_i^{2,\mu}\}$. Druhý, prechodový člen v (8.58) teraz nahradíme týmto výrazom

$$h_i^\Delta(t) = \lambda \sum_{\mu=1}^q (\xi_i^{1,\mu+1} + \xi_i^{2,\mu+1}) (m^{1,\mu}(t-\tau) + m^{2,\mu}(t-\tau)). \quad (8.59)$$

Ak je systém v stave $(1, \mu)$, môže pokračovať buď do stavu $(1, \mu+1)$ alebo $(2, \mu+1)$. Skutočný prechod bude determinovaný vonkajším signálom $h_i^{out} = \rho \xi_i^{s,\mu+1}$, kde $\rho \leq 1$. Ak bude tento signál z postupnosti $s=1$, tak sa bude pokračovať v postupnosti 1, ak bude z postupnosti $s=2$, tak sa bude pokračovať v postupnosti 2. Zovšeobecnenie (8.59) pre viac ako dve postupnosti je triviálne. Pri vhodnom výbere hodnôt parametrov môže sieť rozlíšiť až 2^p rôznych časových postupností.

8.9 Invariantné rozpoznávanie vzorov

Invariantné rozpoznávanie obrázcov (vzorov) robí mozog neustále. Vo vizuálnej sfére je to napríklad rozpoznávanie objektov bez ohľadu na ich veľkosť, posunutie a otočenie. V sluchovej sfére je to napríklad rozpoznávanie melódií bez ohľadu na rozdielnu výšku tónovej oktávy, v ktorej sú zahraté, alebo invariantné rozpoznávanie hlások a slov v reči rôznych ľudí. Neurobiologické mechanizmy tohto procesu (či procesov) sú stále predmetom intenzívneho experimentálneho a teoretického výskumu.

Zaujímajú nás možnosti ako prispieť k hľadaniu riešenia pomocou autoasociatívnych neurónových sietí. Vnútorňa reprezentácia nejakého vzoru je v autoasociatívnej sieti zakódovaná v aktivite celej siete. Táto aktivita zodpovedá len istej konkrétnej vstupnej stimulácii, ktorú opakovane vyvoláva daný vzor, najskôr v procese učenia a potom v procese rozpoznávania. Táto predstava zodpovedá neurobiologickej realite. Avšak ani mohutné autoasociatívne vlastnosti Hopfieldových sietí, takých ako sme ich doteraz predstavili, nestačia na simulovanie invariantného rozpoznávania vzorov. To, ktorý atraktor si sieť vo svojej evolúcii vyberie, silno závisí na tom, ako ďaleko je v stavovom priestore počiatočná konfigurácia od jednotlivých atraktorov. Pri posunutí, zmenšení, prevrátení, atď. vzoru treba počítať s tým, že iníciaľny stav siete je v priestore stavov veľmi vzdialený od príslušnej pamäťovej konfigurácie. Inými slovami, posunutému, či inak zmenenému vzoru, zodpovedá úplne iný vzorec aktivity siete, ktorý je len málo, resp. vôbec nie podobný na aktivitu siete v čase zapamätávania tohto vzoru. Keby sme sieť nechali spontánne sa vyvíjať, skončila by v atraktore, ktorý sa najviac podobá vstupnej stimulácii.

Dotsenkov algoritmus pre invariantné rozpoznávanie vzorov v autoasociatívnej sieti vzhľadom k ich posunutiu, otočeniu a zmene rozmerov [24] je založený na využití neurónových prahov excitácie pri riadení evolúcie siete v stavovom priestore. Tento prístup si najskôr vysvetlíme pre prípad invariantného rozpoznania vzoru vzhľadom k jeho posunutiu. Neurónová sieť, pre ktorú je algoritmus navrhnutý, je pôvodným modelom Hopfieldovej stochastickej siete so symetrickou váhovou maticou, s neurónmi pospájanými každý s každým, a s asynchrónnou dynamikou. V ďalšom budeme pracovať s geometrickou predstavou neurónovej siete ako 2-rozmernej mriežky, a preto namiesto indexu i budeme používať vektor polohy \mathbf{r} , resp. pre j -ty neurón vektor \mathbf{r}^j . Vzdialenosť okamžitej konfigurácie siete $\mathcal{S}(t)$ od jednotlivých pamäťových vzorov ξ^μ vyjadruje okamžitý prekryv $m^\mu(t)$

$$m^\mu(t) = \frac{1}{N} \sum_{\mathbf{r}} \xi^\mu(\mathbf{r}) S(\mathbf{r}, t) \quad , \quad \text{pre } \mu = 1, \dots, p. \quad (8.60)$$

Táto rovnica je tá istá rovnica ako (8.17), len prepísaná pomocou vektorov polohy \mathbf{r} namiesto indexu i . Relaxačná dynamika neurónov je asynchrónna s pravdepodobnosťou zmeny stavu pre neurón so súradnicou \mathbf{r} vyjadrenou vzťahom (vid' rovnice 8.28 a 8.29):

$$S(\mathbf{r}, t) = \begin{cases} +1 & \text{s pravdepodobnosťou } P_+ = \frac{1}{1 + \exp(-2\beta h(\mathbf{r}, t))} \\ -1 & \text{s pravdepodobnosťou } P_- = 1 - P_+ \end{cases} \quad (8.61)$$

Celkové pole $h(\mathbf{r}, t)$ pôsobiace na neurón na mieste \mathbf{r} v čase t je rovné súčtu príspevkov od ostatných neurónov s indexami \mathbf{r}' váhovanými cez synaptické váhy plus príspevok od lokálneho nenulového prahu excitácie $\theta(\mathbf{r}, t) \equiv h^{\text{ext}}(\mathbf{r}, t) \neq 0$, t.j.

$$h(\mathbf{r}, t) = \sum_{\mathbf{r}'} J_{\mathbf{r}\mathbf{r}'} S(\mathbf{r}', t) + \theta(\mathbf{r}, t) \quad . \quad (8.62)$$

Prahy excitácie sú nielenže nenulové, ale majú aj časovú závislosť, ktorú špecifikujeme neskôr. Pri splnení určitých podmienok môžu prostredníctvom (8.61) veľmi účinne riadiť evolúciu siete v stavovom priestore, lebo pravdepodobnosť zmeny stavu neurónov bude závisieť hlavne od ich relatívnej hodnoty vzhľadom k príspevku od ostatných neurónov.

Váhová matica \mathbf{J} je symetrická a neuróny na sebe netvorí synapsy. Hodnotu synaptickej váhy medzi dvoma neurónmi vypočítame podľa vzťahu (8.11), ktorý pre naše účely prepíšeme pre vektorové súradnice neurónov \mathbf{r} a \mathbf{r}' :

$$J_{\mathbf{r}\mathbf{r}'} = \begin{cases} \frac{J_0}{N} \sum_{\mu=1}^p \xi^\mu(\mathbf{r}) \xi^\mu(\mathbf{r}') & \text{pre } \mathbf{r} \neq \mathbf{r}' \\ 0 & \text{pre } \mathbf{r} = \mathbf{r}' \end{cases} \quad (8.63)$$

Tento vzťah je všeobecnejší ako (8.11), pretože sme doňho zaviedli konštantu J_0 , ktorá vo všeobecnosti nemusí byť rovná 1.

Zatiaľ budeme pracovať s ortogonálnymi (resp. pseudoortogonálnymi) vzormi, v ktorých je 50%-ná priemerná aktivita. Nech $\mathbf{S}^{(0)}$ je počiatková konfigurácia siete, a nech je to jeden z pamäťových vzorov ξ^μ posunutý o vektor $(-\mathbf{a}^*)$. Keďže v sieti musí byť zachovaná 50%-ná priemerná aktivita, uvažujeme periodické okrajové podmienky. To znamená, že pri posunutí vzoru cez okraj mriežky sa na protíľahlom okraji objaví tá časť vzoru, ktorá sa pri posune dostala mimo mriežky (vid' obr. 8.9). Pre pamäťovú konfiguráciu ξ^μ platí vzťah:

$$\xi^\mu(\mathbf{r}) = \mathbf{S}^{(0)}(\mathbf{r} + \mathbf{a}^*), \quad \text{pre } \forall \mathbf{r}. \quad (8.64)$$

Samozrejme, sieť "nevie" aká je hodnota \mathbf{a}^* .

V Dotsenkovom algoritme sa predpokladá, že počiatkový vzor $\mathbf{S}^{(0)}$ sa v čase $t = 0$ premietne na neurónové prahey excitácie tak, že tieto nadobudnú hodnoty

$$\theta(\mathbf{r}, t = 0) = \theta_0 \mathbf{S}^{(0)}(\mathbf{r}), \quad \text{pre } \forall \mathbf{r}. \quad (8.65)$$

θ_0 je kladná konštanta. V každom ďalšom časovom momente sa hodnoty prahov excitácie menia podľa tejto rovnice:

$$\theta(\mathbf{r}, t) = \theta_0 \mathbf{S}^{(0)}(\mathbf{r} + \mathbf{a}(t)), \quad \text{pre } \forall \mathbf{r}. \quad (8.66)$$

Dynamická premenná $\mathbf{a}(t)$ vyjadruje fakt, že konfigurácia prahov excitácie je v každom časovom okamihu iná, že je daná momentálnou hodnotou vektora $\mathbf{a}(t)$. Je to teda konfigurácia prahov excitácie, ktorá hľadá príslušný pamäťový atraktor, a to prostredníctvom dynamiky premennej $\mathbf{a}(t)$. Prahy budú dominovať pri zmene stavu neurónov (8.61) a príspevok od ostatných neurónov bude slúžiť hlavne na korekciu chýb vo vzore. Dotsenko navrhol relaxačnú dynamiku pre dynamickú premennú posunu prahov $\mathbf{a}(t)$ nasledovne:

$$\frac{d\mathbf{a}}{dt} \equiv \mathbf{a}(t+1) - \mathbf{a}(t) = -\frac{\delta E}{\delta \mathbf{a}} + \boldsymbol{\eta}(t). \quad (8.67)$$

V tomto vzorci je E energia siete (8.6) v tvare

$$E(\mathbf{S}) = -\frac{1}{2} \sum_{\mathbf{r}} \sum_{\mathbf{r}'} J_{\mathbf{r}\mathbf{r}'} \mathbf{S}(\mathbf{r}) \mathbf{S}(\mathbf{r}') - \sum_{\mathbf{r}} \mathbf{S}(\mathbf{r}) \theta(\mathbf{r}). \quad (8.68)$$

Pomocou relaxačnej prahovej dynamiky, ktorej podstata je vyjadrená rovnicou (8.67), hľadáme lokálne minimum energie systému (8.68). Rovnica (8.67) odpovedá optimalizácii metódou záporného gradientu. Pri tejto optimalizácii sa minimum danej optimalizačnej funkcie hľadá proti smeru gradientu tejto funkcie. Štartuje sa z posunutého počiatočného bodu (8.65), posunieme sa s prahmi niektorým náhodným smerom, necháme sieť zrelaxovať podľa (8.61) a vypočítame energiu (8.68). Ak je energia nového stavu menšia ako pôvodná, pokračujeme v posúvaní sa týmto smerom. Ak je energia väčšia, vygenerujeme $\mathbf{a}(t)$ v smere opačnom, tak ako predpisuje (8.67). Veľkosť posunu závisí od veľkosti $\square E / \square \mathbf{a}$. Takto by sa iteratívnym spôsobom mala sieť usadiť so svojou aktivitou v konfigurácii zodpovedajúcej lokálnemu energetickému minimu príslušného pamäťového vzoru. Gradientová metóda optimalizácie, ktorá teraz riadi dynamiku systému, má tú nevýhodu, že keď sa systém dostane do falošného minima, nemôže sa dostať von. Preto je nutné

zaviesť do systému náhodnosť, napríklad v podobe obyčajného teplotného šumu $\eta(t)$. V práci [13] sme dospeli k záveru, že namiesto gradientovej minimalizácie je v tomto prípade efektívnejšie použitie optimalizácie pomocou simulovaného žihania [16, 49]. Výsledky simulácie Dotsenkovho algoritmu pre invariantné rozpoznávanie vzorov vzhľadom k posunutiu sú ilustrované na obr. 8.9.

Áké sú podmienky toho, aby prahová dynamika dominovala nad relaxáciou neurónov? Rovnicu pre celkové pole pôsobiace na jeden neurón môžeme pomocou vzťahu pre synaptické váhy (8.63) a vzťahu pre prahy excitácie (8.66) prepísať na tvar

$$h(\mathbf{r}, t) = \frac{J_0}{N} \sum_{\mu=1}^p \sum_{\mathbf{r}'} \xi^{\mu}(\mathbf{r}) \xi^{\mu}(\mathbf{r}') S(\mathbf{r}', t) + \theta_0 S^{(0)}(\mathbf{r} + \mathbf{a}(t)). \quad (8.69)$$

Pomocou prekryvu

$$m^{\mu}(t) = \frac{1}{N} \sum_{\mathbf{r}'} \xi^{\mu}(\mathbf{r}') S(\mathbf{r}', t), \quad (8.70)$$

možno celkové pole vyjadriť takto:

$$h(\mathbf{r}, t) = J_0 \sum_{\mu=1}^p \xi^{\mu}(\mathbf{r}) m^{\mu}(t) + \theta_0 S^{(0)}(\mathbf{r} + \mathbf{a}(t)). \quad (8.71)$$

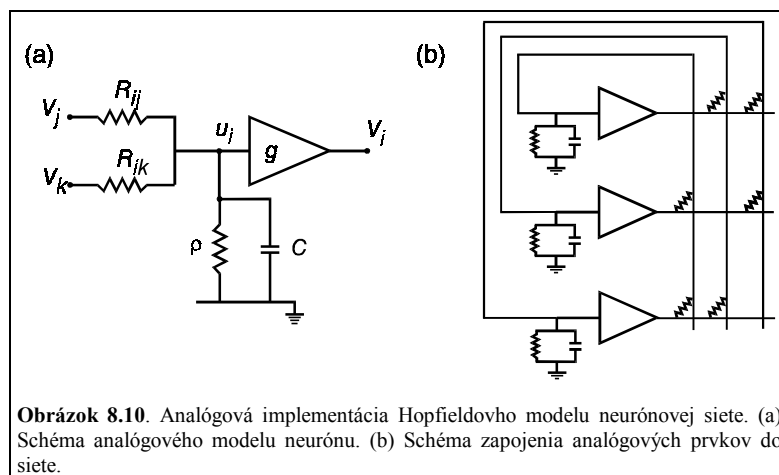
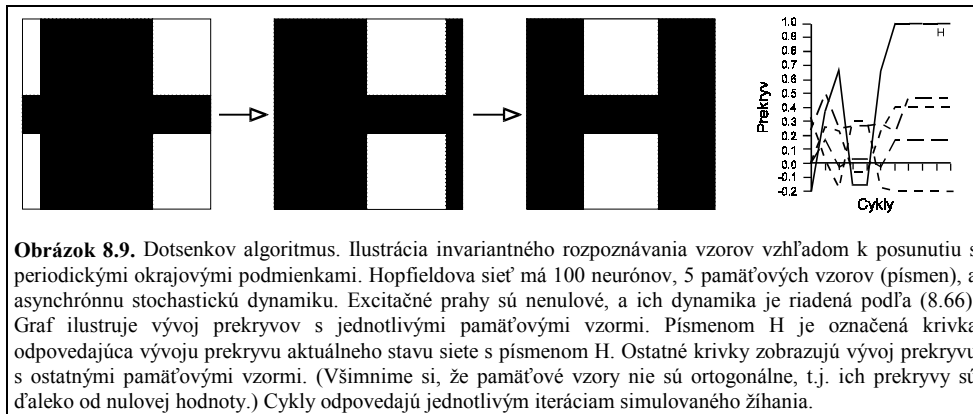
Vidíme, že druhý, prahový člen v rovnici (8.71) bude dominovať, ak $\theta_0 \gg J_0$, a prvý člen môžeme zanedbať. Podľa rovnice (8.60) pre strednú hodnotu stavu neurónu \mathbf{a} (8.71) zistíme, že stavy neurónov budú kopírovať prahové pole

$$\langle S(\mathbf{r}, t) \rangle = \tanh[\beta h(\mathbf{r}, t)] \cong \beta \theta_0 S^{(0)}(\mathbf{r} + \mathbf{a}(t)). \quad (8.72)$$

V praxi stačí, keď $\theta_0/J_0 = 2$.

Dotsenkov algoritmus, ktorý sme práve popísali pre invariantné rozpoznávanie vzorov vzhľadom k posunutiu, sa môže rovnakým spôsobom použiť aj na invariantné rozpoznávanie vzorov vzhľadom na ich veľkosť alebo otočenie. Rovnica (8.64) vyjadrujúca vzťah medzi pamäťovou konfiguráciou a počiatočným stavom siete bude pre zmenu veľkosti vzoru vyzerat' takto:

$$\xi^{\mu}(\mathbf{r}) = S^{(0)}(\lambda^* \mathbf{r}), \quad \text{pre } \forall \mathbf{r}. \quad (8.73)$$



Skalár λ^* hovorí, koľkokrát je počiatočná konfigurácia zväčšená (zmenšená) oproti pamäťovej konfigurácii. Pre otočenie zavedieme operátor otočenia Ω , takže rovnica analogická rovniciam (8.64) a (8.73) bude

$$\xi^\mu(\mathbf{r}) = S^{(0)}(\Omega^* \mathbf{r}), \quad \text{pre } \forall \mathbf{r}. \quad (8.74)$$

Namiesto dynamickej premennej $\mathbf{a}(t)$ bude v prípade transformácie veľkosti vzoru riadiť dynamiku dynamická premenná $\lambda(t)$, a v prípade otáčania dynamická premenná $\Omega(t)$.

Na záver spomenieme, že existujú aj iné prístupy k riešeniu invariantného rozpoznávania vzorov pomocou Hopfieldových autoasociatívnych sietí, napríklad prístupy, ktoré navrhli Bienenstock a von der Malsburg [15] alebo Kree a Zippelius [56].

8.10 Analógový Hopfieldov model

Skutočné neuróny v mozgu nie sú dvojestavové prvky, ale majú spojitú, nelineárnu vstupno-výstupnú charakteristiku. Táto skutočnosť ako aj požiadavky hardwarovej implementácie Hopfieldovej neurónovej siete motivovali Hopfielda k tomu, aby svoj model rozšíril aj na siete so spojitými analógovými prvkami [41]. Skutočné neuróny aj elektronické zariadenia (ako napríklad operačný zosilňovač) majú integračné časové oneskorenia kvôli ich membránovej resp. elektrickej kapacitancii, a evolúciu ich stavu (t.j. elektrického napätia) v čase popisujú diferenciálne rovnice príslušného elektrického obvodu.

Na obr. 8.10a je nakreslený elektrický obvod, ktorý reprezentuje jeden prvok (**analógový neurón**) v analógovej implementácii Hopfieldovho modelu. Obr. 8.10b zobrazuje sieť tvorenú takýmito prvkami.

Premenná u_i označuje vstupné napätie i -teho prvku, V_i je výstupné napätie. Ako u_i tak aj V_i sú funkciou času, vyvíjajú sa v čase. Operačný zosilňovač má vstupno-výstupnú funkciu g takú, že

$$V_i = g(u_i). \quad (8.75)$$

Najčastejšie sa používa **nelineárna funkcia** $g(u)=\tanh(\beta u)$, ktorá má hodnoty v intervale $[-1,+1]$ alebo sigmoidálna funkcia $g(u)=f_{\beta}(u)$, s hodnotami v intervale $[0,+1]$ (viď obr. 8.6). Vstup každého prvku je uzemnený cez odpor ρ a kondenzátor s kapacitou C . Tieto premenné reprezentujú transmembránový odpor a kapacitu skutočného neurónu (viď obr. 1.9). Výstup j -teho prvku je spojený so vstupom i -teho prvku cez odpor R_{ij} . Z Kirchhoffovho zákona o rovnosti prúdov do uzla vtekajúcich a z uzla vytekajúcich vyplýva, že

$$C \frac{du_i}{dt} + \frac{u_i}{\rho} = \sum_{j \neq i}^N \frac{1}{R_{ij}} (V_j - u_i). \quad (8.76)$$

alebo tiež

$$\tau_i \frac{du_i}{dt} = -u_i + \sum_{j \neq i}^N w_{ij} g(u_j), \quad (8.77)$$

kde

$$\tau_i = R_i C; \quad \frac{1}{R_i} = \frac{1}{\rho} + \sum_{j \neq i}^N \frac{1}{R_{ij}}; \quad w_{ij} = \frac{R_i}{R_{ij}}. \quad (8.78)$$

Podiel $w_{ij}=R_i/R_{ij}$ reprezentuje hodnotu synaptickej váhy synapsy tvorenej j -tým prvkom na i -tom prvku. Ak zvolíme hodnotu ρ dostatočne malú a R_{ij} veľké, môžeme členy $1/R_{ij}$ zanedbať, a potom $R_i \approx \rho$ a $w_{ij} \approx \rho/R_{ij}$ pre $\forall i$. Ak má mať niektorá váha negatívnu hodnotu, môžeme to v hardwarovej implementácii vyriešiť napríklad tak, že pred ňu zapojíme invertor, ktorý zmení hodnotu vstupného napätia na $-V_j$.

Keď systém skonverguje do atraktora, pre $\forall i$ platí $du_i/dt=0$ a $dV_i/dt=0$. Vtedy $u_i = \sum_j w_{ij}V_j$. Rovnaké riešenie ako systém dynamických rovníc (8.77) má aj systém diferenciálnych rovníc pre V_i :

$$\tau_i \frac{dV_i}{dt} = -V_i + g(u_i) = -V_i + g\left(\sum_{j \neq i}^N w_{ij}V_j\right). \quad (8.79)$$

Dynamika tohto systému je spojitá, t.j. **všetky prvky aktualizujú svoj stav spojite a súčasne** podľa rovníc (8.75) a (8.77) resp. (8.79). Sledovať evolúciu vyššie popísanej analógovej neurónovej siete nám umožní **numerická integrácia systému rovníc** (8.77) alebo (8.79) na počítači. V literatúre sa odporúča napríklad integrovanie s adaptívnou veľkosťou kroku ako je napríklad Bulirschova-Stoerova metóda [65]. Ak je strmota \square funkcie $g(u)$ vysoká, výstupy prvkov sa blížia k hraniciam ± 1 , resp. 0 a 1. Takto môžeme získať binárne odpovede aj od prvkov spojitej analógovej siete. Pri určitých úlohách je však potrebné pracovať so spojitémi hodnotami (viď aplikácie Hopfieldovho modelu).

Analógia so stochastickou sieťou: všimnime si, že rovnica (8.75) s $g(u) = \tanh(\beta u)$ je vlastne taká istá ako rovnice stredného poľa (8.34) pre stochastickú Hopfieldovu sieť. Vstupné napätie analógového prvku u_i hrá úlohu priemerného interného poľa $\langle h_i \rangle$ a výstupné napätie V_i je ekvivalentné priemernej hodnote stavu i -teho neurónu $\langle S_i \rangle$. To znamená, že analógovú sieť môžeme použiť na riešenie rovníc stredného poľa pre stochastickú Hopfieldovu sieť pri nenulovej "teplote" T .

Takisto ako pre binárnu Hopfieldovu sieť, aj pre analógovú sieť môžeme definovať energetickú funkciu tak, aby sa energia systému v priebehu evolúcie minimalizovala. Rovnovážne riešenia rovníc (8.77) a (8.79) zodpovedajú minimám tejto energetickej funkcie a im prislúchajúce konfigurácie aktivity odpovedajú atraktorom v stavovom priestore. **Energia** analógovej Hopfieldovej siete je [41]

$$E = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N w_{ij}V_iV_j + \sum_{i=1}^N \int_0^{V_i} g^{-1}(V)dV, \quad (8.80)$$

kde g^{-1} je inverzná funkcia k funkcii g . Aby sme ukázali, že **E sa v priebehu evolúcie systému minimalizuje**, derivujeme (8.80) podľa času t :

$$\begin{aligned} \frac{dE}{dt} &= -\frac{1}{2} \sum_{ij} w_{ij} \frac{dV_i}{dt} V_j - \frac{1}{2} \sum_{ij} w_{ij} V_i \frac{dV_j}{dt} + \sum_i g^{-1}(V_i) \frac{dV_i}{dt} \\ &= -\sum_i \frac{dV_i}{dt} \left(\sum_j w_{ij} V_j - u_i \right) = -\sum_i \tau_i \frac{dV_i}{dt} \frac{du_i}{dt} = -\sum_i \tau_i g'(u_i) \left(\frac{du_i}{dt} \right)^2 \leq 0. \end{aligned} \quad (8.81)$$

Pri týchto úpravách sme využili predpoklad, že synaptické váhy sú symetrické ($w_{ij} = w_{ji}$) a že $w_{ii} = 0$ a tiež vzťahy (8.75) a (8.77). Časová derivácia energie je záporná preto, lebo $\tau_i > 0$, $g(u)$ je monotónne rastúca funkcia, a teda jej prvá derivácia podľa času g' je nezáporná, a

derivácia u_i tu vystupuje v druhej mocnine. Výraz (8.81) sa rovná nule iba v rovnovážnom bode, keď $du_i/dt=0$ a $dV_i/dt=0$. Vzťah (8.81) ukazuje, že dynamická rovnica (8.77) spojitou minimalizuje energiu systému, pokiaľ systém nedospeje do stavu zodpovedajúceho lokálnemu energetickému minimu, kde $E=0$. **Rovnovážne stavy** systému určené synaptickou maticou W sú jeho atraktormi, a sú to **jediné atraktory** systému. Napríklad limitné cykly nemôžu byť atraktormi, lebo energia nemôže na uzavretej krivke spojitou klesať. Symetria matice W garantuje existenciu bodových atraktorov. Ak by sme do synaptickej matice zaviedli asymetriu, môžu vzniknúť oscilácie alebo chaotický vývoj $V_i(t)$. V extrémnom prípade čisto náhodných váh w_{ij} s hodnotami, ktoré majú priemernú hodnotu 0 a strednú kvadratickú odchýlku σ^2 , môžeme zaznamenať prechod od stabilného ku chaotickému vývoju, keď zvyšujeme hodnotu σ^2 [68].

Ďalšou možnosťou ako sledovať evolúciu tohto systému je jeho **hardwarová implementácia** pomocou elektronických prvkov. Najťažší problém pri vytváraní obvodov tejto neurónovej siete vo VLSI je výroba rezistorov R_{ij} . Pre úplne pospájanú sieť s N prvkami potrebujeme $2N^2$ rezistorov (keď používame invertory), a ich odpor musí byť dostatočne veľký, aby sa obmedzili energetické straty. Graf so spolupracovníkmi [33] vyrobil čip s $N = 256$ analógovými neurónmi, ktoré boli pospájané každý s každým cez približne $2N^2 \gg 130\,000$ rezistorov, ktoré mali približne rovnaký odpor. Každý rezistor sa pridával individuálne na hotový CMOS čip pomocou elektrónovo-lúčovej litografie. Pri tejto metóde je hodnota odporu rezistorov raz navždy daná, ale nestráca sa po vypnutí zdroja napätia. Iní použili namiesto pasívnych odporov aktívne elektronické elementy (tranzistory) [2]. Takto získali flexibilnejšie, programovateľné čipy v zmysle modifikácie hodnôt ich synaptických spojení. Existujú aj optické a optoelektronické implementácie Hopfieldovej analógovej neurónovej siete [1].

8.11 Využitie Hopfieldovho modelu

Veľkou nevýhodou Hopfieldovho modelu neurónovej siete je, že sa nedokáže učiť z príkladov tak ako napríklad viacvrstvá sieť na základe metódy spätného šírenia sa chýb alebo Kohonenova samoorganizujúca sa mapa. Geometria pamäťových vzorov musí byť dopredu určená a slúži na naprogramovanie synaptickej matice.

Prvou oblasťou aplikácie Hopfieldovho modelu je **oblasť modelovania neurobiologických a psychických javov**. Pomocou Hopfieldovho modelu neurónovej siete a jeho rôznych variácií môžeme simulovať nasledovné kognitívne procesy:

- **Rozpoznávanie** — vnímaný objekt už bol raz kedysi vnímaný, je v pamäti.
- **Asociatívne vybavenie si z pamäti** — rekonštrukcia kompletnej položky (reprezentácie objektu) v pamäti, iniciovaná prítomnosťou fragmentu známeho vzoru na vstupe siete.
- **Klasifikácia** — konkrétny atraktor (resp. okolie atraktora) reprezentuje spracovaný a zapamätaný vzor. Pri modelovaní vnímania sú vstupom siete signály zo senzorických orgánov. Keď by sme modelovali iné mozgové funkcie ako napríklad myslenie, vstupmi budú konfigurácie aktivity z iných kortikálnych (resp. asociačných) častí mozgu. Opakujúci sa vzorec neurónovej aktivity celej siete (atraktor) je signálom toho, že nastala kognitívne významná udalosť. Konkrétny vstup (zodpovedajúci počiatočnej konfigurácii aktivity siete) je kognitívne zmysluplný vtedy, keď vedie sieť rýchlo do okolia atraktora. Zmysel tomuto vstupu dáva práve daný atraktor, ktorý v minulosti vznikol na danej konkrétnej úrovni kognitívneho spracovania.

Ak chceme pomocou autoasociatívnych sietí simulovať kognitívne procesy, ako sú vnímanie, pamäť a iné, musia byť splnené o.i. tieto podmienky:

□ Relaxačný čas (doba príchodu do blízkosti atraktora) musí byť dostatočne krátky, aby stav siete dosiahol atraktor ešte pred tým, ako je sieť postavená pred novú úlohu. Nová úloha môže byť reprezentovaná príchodom novej externej stimulácie, prípadne nejakým vnútorným mechanizmom, ktorý resetuje sieť. Pri väčšine kognitívnych úloh ide o časy rádovo desiatok až stoviek milisekúnd.

□ Doba zotrávania siete v blízkosti atraktora musí byť dostatočne dlhá. Ak pripisujeme atraktoru kognitívnu úlohu, výstupný mechanizmus musí byť schopný detekovať fakt, že systém skutočne dospel do blízkosti atraktora, čiže musí byť schopný odlišiť tento stav od prechodných vzorcov aktivity siete. Odhaduje sa, že by sa mohlo uvažovať o časovom priemernení aktivity počas obdobia 3040 ms [9].

Ďalej na niektorých príkladoch naznačíme, ako možno konkrétne aplikovať autoasociatívne neurónové siete na modelovanie kognitívnych procesov prebiehajúcich v mozgu resp. v nervovom systéme ako takom. Podrobnú informáciu možno nájsť v citovanej literatúre. Modely týkajúce sa nervového systému je v súčasnosti ťažko možno navrhovať a overovať na základe priamych meraní na neurónoch tak, ako to bolo možné v prípade modelu Kleinfelda a Sompolinského [51, 52] alebo Amita so spolupracovníkmi [6]. Kleinfeld a Sompolinsky navrhli analógovú autoasociatívnu neurónovú sieť, ktorá má také dynamické správanie ako majú tzv. **centrálne generátory rytmu** (neurónové okruhy generujúce a riadiace cyklické vzorce motorickej aktivity). Svoj model aplikovali na prípad únikového plávania mäkkýša *Tritonia diomedea*. Vzorce aktivity v sieti sa zhodovali s hodnotami nameranými pomocou elektród na skutočných neurónoch. S týmito hodnotami a Hopfieldovou dynamikou, umelá neurónová sieť produkovala cyklický výstup slúžiaci na riadenie svalov. Amit so spolupracovníkmi [6] skonštruovali analógovú autoasociatívnu sieť, ktorá reprodukuje **reverberácie aktivity** namerané v **mozgovej kôre** opíc učiacich sa zapamätať a rozpoznať časové sekvencie zrakových stimulov. Zrakové stimuly boli náhodné bitové obrázky a prezentovali sa vo fixnom poradí. Opica bola potom testovaná tak, že pri obrázkoch prezentovaných tentoraz v náhodnom poradí mala určiť, či sa dva po sebe idúce obrázky pri učení vyskytli za sebou alebo nie. Napriek tomu, že obrázky boli nekorelované (náhodné) vzorce aktivity, vzorce aktivity namerané v mozgovej kôre mali korelovanú distribúciu aktivity v závislosti od toho, či sa obrázky vyskytovali v postupnosti za sebou alebo nie. Veľkosť tejto korelácie bola funkciou vzájomnej "vzdialenosti" obrázkov v postupnosti (ich vzájomného poradia). Všetky empirické výsledky tejto štúdie boli reprodukovateľné pomocou atraktorovej neurónovej siete. Treves s Rollsom [69] sa pomocou modelu atraktorovej siete pokúsili vysvetliť funkciu a **činnosť hippocampu** (mozgovej štruktúry, ktorá hrá dôležitú úlohu pri pamäťových procesoch u zvierat aj u človeka).

My sme skonštruovali stochastickú Hopfieldovu sieť s pamäťovými vzormi reprezentujúcimi hudobné tóny, pomocou ktorej sme simulovali **spektrálne invariantné rozpoznávanie hudobných tónov** [12]. Konkrétne išlo o teoretické vysvetlenie javu chýbajúcej základnej frekvencie. Treba vedieť, že zložené hudobné tóny používané v psychofyzikálnych experimentoch sa skladajú zo základnej frekvencie a z určitého konečného počtu vyšších frekvencií, ktoré sú vždy celistvým násobkom základnej frekvencie (tzv. vyššie harmonické). Tieto frekvencie definujú farbu tónu. Nota alebo výška tónu je určená hodnotou základnej frekvencie. Zaujímavosť fenoménu chýbajúcej

základnej frekvencie spočíva v tom, že človek počuje tón prislúchajúci určitej základnej frekvencii aj v prípade, že sluchový stimul obsahuje iba vyššie harmonické frekvencie a základná frekvencia v spektre zvuku fyzicky chýba. Pamäťové vzory sme zostrojili v súhlase so známou geometriou reprezentácie zvukových frekvencií v mozgovej kôre. Frekvencie počuteľného spektra sú reprezentované zhruba rovnobežnými pásmi neurónov, ktoré sú orientované kolmo v smere frekvenčného gradientu. Sledovali sme kvalitu vybavovania tónov (resp. ich neurónových reprezentácií) v závislosti od toho, ktoré vyššie harmonické frekvencie boli v zvuku prítomné a dosiahli sme zhodné výsledky s výsledkami psychofyzikálnych experimentov uskutočnených na ľuďoch. Neskôr sme použili Dotsenkov model invariantného rozpoznávania modifikovaný pomocou metódy simulovaného žihania na **modelovanie transpozične invariantného rozpoznávania melódií**, vzhľadom ku konštantnému frekvenčnému posunu všetkých tónov [13]. Kvôli tomu bolo potrebné vytvoriť sieť s asymetrickou váhovou maticou, v ktorej boli zahrnuté synaptické časové oneskorenia, ktoré umožňujú rozpoznávať a generovať časové postupnosti konfigurácií.

Objavili sa aj **psychiatrické špekulácie** využívajúce pojmy atraktora, falošného atraktora a šumu na vysvetlenie obsahových a formálnych porúch reči a myslenia pri schizofrénii a mánii [39]. Normálna reč a myslenie by v tejto metafore súviseli s konvergenciou príslušnej mozgovej autoasociatívnej siete do správneho jedného bodového atraktora zodpovedajúceho jednému pojmu. Porucha reči a myslenia pri mánii by v tejto metafore zodpovedala spontánnym prechodom z jedného atraktora do druhého vplyvom príliš veľkého vnútorného šumu. Vysvetlenie špecifického charakteru obsahovej a formálnej poruchy reči a myslenia pri schizofrénii by zodpovedalo metafore, podľa ktorej mozgové autoasociatívne siete majú takú nízku úroveň šumu, že falošné atraktory, zmiešaniny pravých atraktorov (pojmov), sú veľmi dobrými atraktormi.

Druhou oblasťou problémov, na riešenie ktorých môžeme úspešne použiť algoritmy inšpirované teóriou neurónových sietí, sú **optimalizačné problémy**. Na neurónovú sieť Hopfieldovho typu sa môžeme pozeráť ako na masívne paralelný algoritmus, ktorého cieľom je minimalizácia energie systému. V teórii optimalizácie sa vo všeobecnosti hovorí, že v priebehu optimalizácie sa minimalizuje **cenová resp. objektívna funkcia** (angl. *cost resp. objective function*), ktorá v teórii neurónových sietí odpovedá účelovej (kriteriálnej) funkcii. Predtým ako si konkrétne načrtneme niektoré optimalizačné problémy, chceme ešte raz zdôrazniť, že keď sa hovorí, že tieto problémy sa riešia pomocou neurónových sietí, v skutočnosti sa pre riešenie každého z týchto problémov konštruuje špeciálne prispôsobený masívne paralelný algoritmus.

Ako prvé spomenieme **kombinatorické optimalizačné problémy** [38]. Pri týchto problémoch hľadáme riešenie v množine veľkého množstva možných kombinácií základných prvkov systému. Celkový počet riešení pre veľkosť problému s počtom prvkov N je obvykle exponenciálnou funkciou N , a tomu je úmerný aj čas potrebný na vyriešenie problému. Na ilustráciu si spomenieme ako prvý **problém váhovaného priradenia** (angl. *weighted matching problem*). V priestore majme množinu N bodov, pričom poznáme vzdialenosti medzi jednotlivými dvojicami týchto bodov d_{ij} . Tieto body sa môžu nachádzať v euklidovskom priestore a d_{ij} môže reprezentovať Euklidovu vzdialenosť, alebo tieto body môžu byť abstraktné entity a hodnoty d_{ij} môžu reprezentovať ich vzťahy. Vo všeobecnosti sa môžu d_{ij} považovať za nezávislé náhodné premenné s pravdepodobnostnou distribúciou $P(d_{ij})$. Našou úlohou je pospájať dvojice bodov tak, aby každý bod bol spojený len s jedným iným bodom tak, aby celková dĺžka spojení bola minimálna. Praktickými príkladmi

takéhoto problému môže byť spájanie prvkov v nejakom elektronickom zariadení, optimálne mapovanie procesov na dva ekvivalentné procesory, priradovanie žiakov do škôl, a pod. Každému páru bodov i a j , priradíme prvok n_{ij} , $i < j$. Nech $n_{ij} = 1$, keď medzi i -tým a j -tým bodom existuje spojenie a $n_{ij} = 0$, keď medzi nimi spojenie nie je. Sám so sebou sa bod nemôže spojiť, takže $n_{ii} = 0$ a pre $j < i$ platí $n_{ij} = n_{ji}$. Funkcia, ktorej minimum hľadáme je celková dĺžka spojení

$$L = \sum_{i < j} d_{ij} n_{ij} . \quad (8.82)$$

Túto funkciu budeme musieť trochu modifikovať, lebo každé riešenie musí spĺňať jedno obmedzenie, a síce, že každý jeden bod môže byť spojený iba s jedným ďalším bodom, a teda $\sum_j n_{ij} = 1$ pre $\forall i$. Dodržiavanie tohto obmedzenia sa rieši penalizáciou v objektívnej funkcii. Ako základ pre objektívnu funkciu zoberieme (8.82) a pridáme k nej člen, ktorý bude rásť priamo úmerne porušeniu danej podmienky. Objektívna funkcia E bude rovná:

$$E = \sum_{i < j} d_{ij} n_{ij} + \frac{\gamma}{2} \sum_i \left(1 - \sum_j n_{ij} \right)^2 . \quad (8.83)$$

Veľkosť konštanty γ by mala byť asi taká ako priemerná hodnota d_{ij} . Pravdepodobnosť, že hodnota ľubovoľného prvku n_{ij} sa zmení, je rovná

$$P(n_{ij} \rightarrow n'_{ij}) = \frac{1}{1 + \exp(\beta \Delta E)} \quad \text{kde} \quad \Delta E = E(n'_{ij}) - E(n_{ij}) . \quad (8.84)$$

Toto pravidlo hovorí, že zmeny, ktoré minimalizujú objektívnu funkciu E sú pravdepodobnejšie ako tie, ktoré ju zvyšujú. Hľadáme teda takú distribúciu n_{ij} , ktorá zodpovedá jednému z miním (8.83). Systém môžeme naštartovať z náhodnej konfigurácie n_{ij} a každú ďalšiu konfiguráciu generovať náhodne ako v **Monte Carlo simulácii** s tým, že každú novú konfiguráciu akceptujeme s pravdepodobnosťou (8.84). Je vhodné použiť i **metódu simulovaného žihania** a v priebehu evolúcie znižovať teplotu $T = \beta^{-1}$ [16, 49]. Problém môžeme riešiť aj pomocou analógového systému buď počítačovou simuláciou alebo hardwarovou implementáciou, vtedy n_{ij} blízke 1 (0) budeme považovať za rovné 1 (0).

Ďalší kombinatorický optimalizačný problém, ktorý sa stal štandardom na testovanie efektívnosti optimalizačných metód je **problém obchodného cestujúceho** (angl. *traveling salesman problem*) [42, 43]. Máme v priestore N bodov alebo miest, medzi dvojicami ktorých sú vzdialenosti d_{ij} . Našou úlohou je nájsť najkratšiu uzavretú dráhu, ktorá prechádza cez každé mesto len raz a vracia sa do pôvodného mesta, z ktorého sme vyštartovali. Praktické príklady zahŕňajú efektívne plánovanie ciest kamiónov, vlakov, lietadiel, ale aj úlohy v robotike, napr. pri plánovaní efektívneho pohybu ramena robota. Na reprezentáciu riešenia si zvolíme stochastický binárny prvok n_{ia} , pričom $n_{ia} = 1$, keď mesto i je a -tou zastávkou na okružnej ceste. Celková dĺžka okružnej cesty je

$$L = \frac{1}{2} \sum_{ij,a} d_{ij} n_{ia} (n_{j,a+1} + n_{j,a-1}) . \quad (8.85)$$

Zároveň musia byť splnené dve obmedzenia, a to

$$\sum_a n_{ia} = 1 \quad \text{pre } \forall i \quad \text{a} \quad \sum_i n_{ia} = 1 \quad \text{pre } \forall a . \quad (8.86)$$

Prvé obmedzenie hovorí, že každé mesto i sa objaví počas okružnej cesty len raz, a druhé obmedzenie znamená, že každá zastávka na ceste sa týka len jedného mesta. Teraz môžeme vytvoriť objektívnu funkciu E s dvoma penalizáciami, ktoré sa minimalizujú vtedy, keď sú podmienky (8.86) splnené:

$$E = \frac{1}{2} \sum_{ij,a} d_{ij} n_{ia} (n_{j,a+1} + n_{j,a-1}) + \frac{1}{2} \left[\sum_a \left(1 - \sum_i n_{ia} \right)^2 + \sum_i \left(1 - \sum_a n_{ia} \right)^2 \right] . \quad (8.87)$$

Riešenie problému obchodného cestujúceho, teda konkrétne hodnoty n_{ia} zodpovedajúce optimálnej okružnej ceste, hľadáme tak ako v prípade problému váhovaného priradenia, teda pomocou počítačovej simulácie alebo hardwarovej implementácie príslušného systému. Používame buď binárne prvky so stochastickou dynamikou (8.84) a simulované žihanie, alebo analógové prvky.

Posledným príkladom kombinatorického optimalizačného problému je **delenie grafov** (angl. *graph bipartitioning*) [27]. Predstavme si, že navrhujeme čip s N prvkami, ale všetky sa nám naň nezmestia. Potom by sme chceli urobiť dva čipy, tak aby polovica prvkov bola na jednom a polovica prvkov na druhom a aby počet spojení medzi týmito dvoma čipmi bol minimálny. Uvažujme všeobecný graf, t.j. množinu N bodov, vrcholov grafu pospájaných hranami, ktoré spájajú dvojice vrcholov. Nech je N párne. Nech p je fixná pravdepodobnosť, že každý vrchol je spojený s iným. Priemerný počet vrcholov pN , ktoré sú navzájom spojené, sa nazýva valencia grafu. Našou úlohou je rozdeliť vrcholy do dvoch rovnako veľkých množín, medzi ktorými je minimálny počet hrán. Definujme si $c_{ij}=1$, keď sú vrcholy i a j spojené hranou, a $c_{ij}=0$, keď nie sú spojené. Ďalej si pre každý vrchol definujme premennú $S_i=+1$, keď sa vrchol nachádza v prvej množine a $S_i=1$, keď sa vrchol nachádza v druhej množine. Chceme minimalizovať funkciu

$$L = - \sum_{\langle ij \rangle} c_{ij} S_i S_j , \quad (8.88)$$

kde $\langle ij \rangle$ znamená indexovanie cez každý pár vrcholov zvlášť. Premenné S_i podliehajú obmedzeniu $\sum_i S_i = 0$. V termínoch teórie magnetických látok tieto rovnice zodpovedajú feromagnetu s nulovou celkovou magnetizáciou. Každú celkovú magnetizáciu, ktorá sa vzdáľuje od 0 budeme v objektívnej funkcii penalizovať, takže objektívna funkcia systému je

$$E = - \sum_{\langle ij \rangle} c_{ij} S_i S_j + \mu \left(\sum_i S_i \right)^2 . \quad (8.89)$$

Hodnotu μ vyberáme v intervale 0 až 1/2. Rovnica (8.89) pripomína energiu systému analogického spinového sklu. Systém naštartujeme z náhodnej nekorelovanej konfigurácie S a necháme relaxovať, pričom S_i sa menia podľa stochastického pravidla (8.84) s S_i namiesto n_{ij} , kde energia E je vyjadrená ako (8.89). Opäť bude užitočné použiť simulované žihanie. Feromagnet s nulovou celkovou magnetizáciou bol teoreticky extenzívne študovaný a boli nájdené mnohé zaujímavé vlastnosti (napríklad zánik symetrie replík) [59].

Ako **optimalizačný problém** môžeme formulovať aj **spracovanie obrazu** (angl. *image processing*), čiže rekonštrukciu objektu zo zašumeného alebo rozmazaného obrazu [54]. Vstupom do paralelného systému je množina hodnôt d_i zodpovedajúcich dvojdimenzionálnej množine pixlov (*pixel* = angl. *picture cell*). Konkrétne hodnoty d_i môžu reprezentovať napríklad jas, alebo prvú či druhú priestorovú deriváciu jasu, alebo časovú deriváciu jasu, alebo binokulárnu disparitu (t.j. malú vzdialenosť medzi obrazmi toho istého bodu vnímaného pravým a ľavým okom, ktorú používa mozog na konštrukciu tretieho rozmeru objektu, resp. vzdialenosti jednotlivých objektov od pozorovateľa), alebo nejakú inú premennú, ktorá sa mení v priestore. Napríklad v prípade jasu môžu byť dáta binárne, t.j. svetlé alebo tmavé pixle, alebo spojité, reprezentujúce rozličné úrovne šedi v jednotlivých pixloch (angl. *gray-level data*). Vstupom do nášho systému budú zašumené dáta d_i a výstupom budú hodnoty V_i , ktoré reprezentujú rekonštruovaný objekt. Takto formulovaná úloha je veľmi ťažká, a preto sa vo väčšine prípadov používajú *a priori* znalosti o objektoch, ako napríklad že majú (nemajú) hladký povrch, že majú rovné (zakrivené) okraje, že sa prekrývajú (neprekrývajú) s inými objektmi, alebo akékoľvek iné pomôcky. V prípade, že samotné dáta obsahujú príliš málo informácie, hovoríme, že tieto **problémy sú zle formulované** (angl. *ill-posed problems*). Existujú postupy a techniky ako transformovať zle formulované problémy na dobre formulované problémy (angl. *well-posed problems*). Týmto sa zaoberá tzv. **regularizačná teória** [64]. Postupy používané pri počítačovej rekonštrukcii objektu si naznačíme na najjednoduchšom prípade, keď máme obraz jednoliateho hladkého povrchu bez hrán. Uvažujme jednorozmerný prípad, viď obr. 8.11a. Najlepšiu rekonštrukciu hladkej nekonečnej krivky získame, keď budeme nútiť výstupy V_i , aby spĺňali dve podmienky zároveň: hladkosť a čo najväčšiu blízkosť k dátam, teda E bude

$$E = \frac{1}{2} \kappa \sum (V_i - V_{i+1})^2 + \frac{1}{2} \lambda \sum (V_i - d_i)^2, \quad (8.90)$$

kde koeficienty κ a λ vyjadrujú relatívny dôraz na jednu alebo druhú podmienku. Minimum tejto objektívnej funkcie môžeme hľadať napríklad metódou záporného gradientu:

$$\kappa \tau \frac{dV_i}{dt} = - \frac{\partial E}{\partial V_i} = \kappa (V_{i+1} + V_{i-1} - 2V_i) + \lambda (d_i - V_i). \quad (8.91)$$

Ak by sme položili $\lambda = 0$, a teda nevyžadovali blízkosť k dátam, pre každé V_i by sme dostali difúziu rovnicu. Jednotlivé hodnoty V_i by sa postupne rovnomerne rozptýlili. Nenulová hodnota koeficientu λ však núti V_i prídŕžať sa dát d_i , takže dostaneme hladkú krivku, ktorá sa snaží čo najviac približovať k dátam. Zovšeobecnenie pre dvojrozmerný prípad je v uvažovaní diferencií v (8.90) v rovine. Mohli by sme skonštruovať aj analógový obvod simulujúci rovnicu (8.91), ktorá je analogická dynamickej rovnici pre spojenú Hopfieldovu

sieť (8.77) alebo (8.79), keď k nej pridáme vonkajší vstup λd_i a $g(u) = u$. Ďalej s pomocou (8.76) môžeme identifikovať $\kappa\tau$ ako C , λ ako ρ^{-1} , a κ ako $1/R_{i,i+1}$.

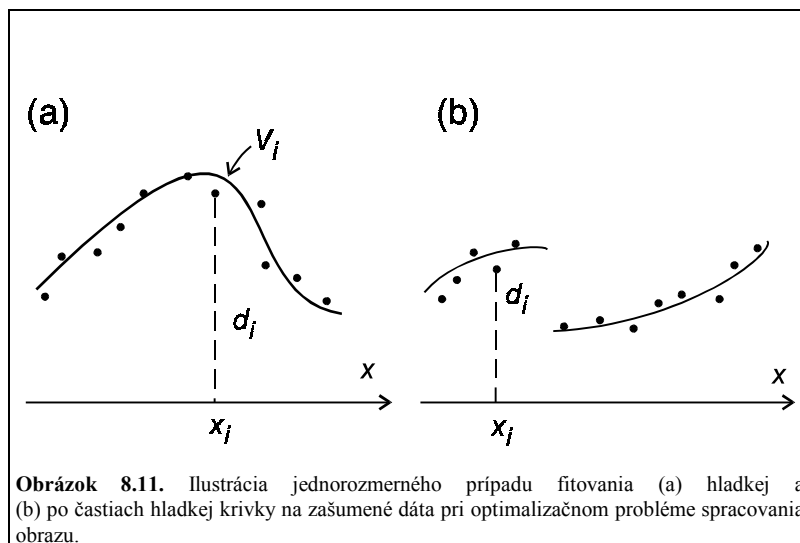
Aby sme mohli riešiť problém spracovania obrazu, ktorý zobrazuje povedzme dva objekty, musíme systém rovníc upraviť tak, aby bol schopný detekovať nespojitosti, a teda nachádzať po častiach hladké (spojité) riešenia (ilustráciu pre jednorozmerný prípad pozri na obr. 8.11b). Rieši sa to tak, že sa do systému zavedú diskrétné prvky detekujúce nespojitosti.

Ilustrujme si to na jednorozmernom prípade. Keďže nevieme dopredu, medzi ktorými bodmi sa bude nachádzať medzera (nespojitosť), medzi každý pár analógových prvkov s výstupom V_i a V_{i+1} dáme jeden binárny prvok s výstupom S_i . Hodnota $S_i = +1$ bude znamenať hypotézu, že medzi i -tým a $(i+1)$ -vým bodom je medzera a hodnota $S_i = -1$, bude znamenať, že je tam krivka spojitá.

Objektívnu funkciu (8.90) treba modifikovať tak, že (a) zrušíme penalizáciu za rozdiel medzi hodnotou V_i a V_{i+1} , keď je medzi nimi predpokladaná nespojitosť a (b) budeme penalizovať príliš veľa hypotéz o nespojitostiach. Nová objektívna funkcia bude teraz vyzeráť takto:

$$E = \frac{1}{2} \kappa \sum_i \frac{1}{2} (1 - S_i) (V_i - V_{i+1})^2 + \frac{1}{2} \lambda \sum_i (V_i - d_i)^2 + \mu \sum_i S_i \quad (8.92)$$

Pri relaxácii systému budú diskrétné prvky skúšať hypotézy o nespojitostiach stochastickým spôsobom. Výsledok môžeme vylepšiť aj tým, že v priebehu relaxácie budeme meniť hodnoty koeficientov κ , λ a μ . Objektívna funkcia (8.92) pre spracovanie dvojrozmerného obrazu zobrazujúceho viacero objektov bude



$$E = \frac{1}{2} \kappa \sum_{\langle ij \rangle} \frac{1}{2} (1 - S_{ij}) (V_i - V_j)^2 + \frac{1}{2} \lambda \sum_i (V_i - d_i)^2 + \mu \sum_{\langle ij \rangle} S_{ij} - \gamma \sum_{\langle ijkl \rangle} S_{ij} S_{jk} S_{kl} S_{li}, \quad (8.93)$$

kde sa sumuje cez dvojice bezprostredných susedov $\langle ij \rangle$ na dvojrozmernej mriežke, alebo cez štvorice susedov $\langle ijkl \rangle$ tvoriace štvorce. Posledný člen nabáda systém, aby sa nespojitosti tvoriace kontúry objektu spájali do uzavretých kriviek.

Rovnice spracovania obrazu prezentované v tomto oddieli nemusia slúžiť len ako základ na interpoláciu povrchov objektov. Môžu slúžiť na detekciu hrán, na rekonštrukciu trojrozmerného tvaru objektov z ich tieňovania alebo binokulárnej disparity, rekonštrukciu tvaru objektov na základe farby alebo pohybu, a pod. Záleží to na príslušnej modifikácii týchto rovníc, na použití analógových alebo stochastických binárnych prvkov, na hodnotách koeficientov, a samozrejme na interpretácii premenných d_i a V_j .

Literatúra

- [1] Y.S. Abu-Mostafa and D. Psaltis. Optical neural computers. *Scientific American*, 256: 88-95, 1987.
- [2] J. Alspector and R.B. Allen. A neuromorphic VLSI learning system. In: P. Losleben, editor, *Advanced Research in VLSI: Proceedings of the 1987 Stanford Conference*, pages 313-349, MIT Press, Cambridge, 1987.
- [3] D. J. Amit and S. Fusi. Constraints on learning dynamic synapses. *Network: Computation in Neural Systems*, 3: 443-449, 1992.
- [4] D. J. Amit and S. Fusi. Dynamic learning in neural networks with material synapses. *Neural Computation*, 4: 957-982, 1994.
- [5] D. J. Amit and N. Brunel. Adequate input for learning in attractor neural network. *Network: Computation in Neural Systems*, 4: 177-186, 1993.
- [6] J.A. Amit, N. Brunel, and M.V. Tsodyks. Correlations of cortical Hebbian reverberations: Theory versus experiment. *The Journal of Neuroscience*, 14: 6435-6445, 1994.
- [7] D.J. Amit, H. Gutfreund, and H. Sompolinsky. Information storage in neural networks with low levels of activity. *Physical Review A*, 35: 2293-2303, 1987.
- [8] D.J. Amit. Neural networks counting chimes. *Proceedings of the National Academy of Sciences of the USA*, 85: 2141-2145, 1988.
- [9] D.J. Amit. *Modeling Brain Function. The World of Attractor Neural Networks*. Cambridge University Press: Cambridge, New York, Sydney, 1989.
- [10] D.J. Amit, H. Gutfreund, and H. Sompolinsky. Spin-glass models of neural networks. *Physical Review A*, 32: 1007-1018, 1985a.
- [11] D.J. Amit, H. Gutfreund, and H. Sompolinsky. Storing infinite numbers of patterns in a spin-glass model of neural network. *Physical Review Letters*, 55: 1530-1533, 1985b.
- [12] L. Beňušková. Modelling the effect of the missing fundamental with an attractor neural network. *Network: Computation in Neural Systems*, 5: 333-349, 1994.
- [13] L. Beňušková. Modelling transpositional invariancy of melody recognition with an attractor neural network. *Network: Computation in Neural Systems*, 6: 313-331, 1995.
- [14] L. Beňušková, M.E. Diamond, and F.F. Ebner. Dynamic synaptic modification threshold: computational model of experience-dependent plasticity in adult rat barrel cortex. *Proceedings of the National Academy of Sciences of the USA*, 91: 4791-4795, 1994.
- [15] E. Bienenstock and C. von der Malsburg. A neural network for invariant pattern recognition. *Europhysics Letters*, 4: 121-126, 1987.
- [16] V. Černý. Thermodynamical approach to the traveling salesman problem: an efficient simulation algorithm. *J. Optim. Appl.*, 45: 41-51, 1985.
- [17] B.G. Cragg and H.N.V. Temperley. The organization of neurones: a cooperative analogy. *EEG Clin. Neurophysiol.*, 6: 85-92, 1954.
- [18] B.G. Cragg and H.N.V. Temperley. Memory: the analogy with ferromagnetic hysteresis. *Brain*, 78: 304-316, 1955.

- [19] A. Crisanti, D.J. Amit, and H. Gutfreund. Saturation level of the Hopfield model for neural network. *Europhysics Letters*, 2: 337-341, 1986.
- [20] A. Crisanti and H. Sompolinsky. Dynamics of spin systems with randomly asymmetric bonds: Langevin dynamics and a spherical model. *Physical Review A*, 36: 4922-4939, 1986.
- [21] B. Derrida, E. Gardner, and A. Zippelius. An exactly solvable asymmetric neural network model. *Europhysics Letters*, 4: 167-173, 1987.
- [22] B. Derrida and J.P. Nadal. Learning and forgetting on asymmetric, diluted neural networks. *Journal of Statistical Physics*, 49: 993-1009, 1987.
- [23] D. W. Dong and J.J. Hopfield. Dynamic properties of neural networks with adapting synapses. *Network: Computation in Neural Systems*, 3: 267-275, 1992.
- [24] V.S. Dotsenko. Neural networks: translation-, rotation- and scale-invariant pattern recognition. *Journal of Physics A: Mathematics and General*, 21: L783-L787, 1988.
- [25] W. Feller. *An Introduction to Probability Theory and Its Applications, Vol. I*, John Wiley and Sons, New York, 1968.
- [26] P. Fedor, L. Beňušková, H. Jakeš, and V. Majerník. An electrophoretic coupling mechanism between efficiency modification of spine synapses and their stimulation. *Studia Biophysica*, 92: 141-146, 1982.
- [27] Y. Fu and P.W. Anderson. Application of statistical mechanics to NP-complete problems in combinatorial optimization. *Journal of Physics A: Mathematics and General*, 19: 1605-1620, 1986.
- [28] E. Gardner. Maximum storage capacity in neural networks. *Europhysics Letters*, 4: 481-485, 1987.
- [29] E. Gardner. Multiconnected neural network models. *Journal of Physics A: Mathematics and General*, 20: 3453-3464.
- [30] E. Gardner. The space of interactions in neural networks models. *Journal of Physics A: Mathematics and General*, 21: 257-270, 1988.
- [31] E. Gardner and B. Derrida. Optimal storage properties of neural networks models. *Journal of Physics A: Mathematics and General*, 21: 270-284, 1988.
- [32] R.J. Glauber. Time-dependent statistics of the Ising model. *Journal of Mathematical Physics*, 4: 294-307, 1963.
- [33] H.P. Graf, L.D. Jackel, R.E. Howard, B. Straughn, J.S. Denker, W. Hubbard, D.M. Tennant, and D. Schwartz. VLSI implementation of a neural network memory with several hundreds of neurons. In: J.S. Denker, editor, *Neural Networks for Computing (Snowbird 1986)*, pages 182-187, American Institute of Physics, New York, 1986.
- [34] H. Gutfreund. Neural networks with hierarchically correlated patterns. *Physical Review A*, 37: 570-586, 1988.
- [35] H. Gutfreund and M. Mézard. Processing temporal sequences in neural networks. *Physical Review Letters*, 61: 235-247, 1988.
- [36] D.O. Hebb. *Organization of Behavior*. J. Wiley and Sons, New York, 1949.
- [37] J.L. van Hemmen and R. Kuhn. Nonlinear neural networks. *Physical Review Letters*, 57: 913-916, 1986.
- [38] J. Hertz, A. Krogh, and R.G. Palmer. *Introduction to the Theory of Neural Computation*, Addison-Wesley Publ. Comp., Redwood City, 1991.
- [39] R.E. Hoffman. Computer simulations of neural information processing and the schizophrenia-mania dichotomy. *Archives of general Psychiatry*, 44: 178-190, 1987.

- [40] J.J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the USA*, 79: 2554-2558, 1982.
- [41] J.J. Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences of the USA*, 81: 3088-3092, 1984.
- [42] J.J. Hopfield and D.W. Tank. "Neural" computation of decisions in optimization problems. *Biological Cybernetics*, 52: 141-152, 1985.
- [43] J.J. Hopfield and D.W. Tank. Computing with neural circuits: a model. *Science*, 233: 625-633, 1986.
- [44] E. Ising. Beitrag zur Theorie des Ferromagnetismus, *Zeitschrift Für Physik*, 31: 253-287, 1925.
- [45] J.J.B. Jack, D. Noble, and R.W. Tsien. *Electric Current Flow in Excitable Cells*. Clarendon Press, Oxford, 1975.
- [46] E.R. Kandel, J.H. Schwartz, and T.M. Jessell. *Principles of Neural Sciences*, Elsevier Science Publishing Co., New York, 1985.
- [47] I. Kanter and H. Sompolinsky. Associative recall of memory without errors. *Physical Review A*, 35: 380--392, 1987.
- [48] S. Kirkpatrick and D. Sherrington. Infinite-ranged models of spin-glasses. *Physical Review B*, 17: 4384-4403, 1978.
- [49] S. Kirkpatrick, C.D. Jr. Gelatt, and M.P. Vecchi. Optimization by simulated annealing. *Science*, 220: 671-680, 1983.
- [50] D. Kleinfeld and I. Kanter. Temporal association in asymmetric neural networks. *Physical Review Letters*, 57: 2861-2864, 1986.
- [51] D. Kleinfeld and H. Sompolinsky. Associative neural network model for the generation of temporal patterns. *Biophysics Journal*, 54: 1039-1051, 1988.
- [52] D. Kleinfeld and H. Sompolinsky. Associative network models for central pattern generators. In: C. Koch and I. Segev, editors, *Methods in Neuronal Modeling: From Synapses to Networks*, The MIT Press, Cambridge, 1989.
- [53] C. Koch and I. Segev, editors. *Methods in Neuronal Modeling: From Synapses to Networks*, The MIT Press, Cambridge, 1989.
- [54] C. Koch, J. Marroquin, and A. Yuille. Analog "neuronal" networks in early vision. *Proceedings of the National Academy of Sciences of the USA*, 83: 4263-4267, 1986.
- [55] R. Kree and A. Zippelius. Continuous-time dynamics of asymmetrically diluted neural networks. *Physical Review A*, 36: 4421-4427, 1987.
- [56] R. Kree and A. Zippelius. Recognition of topological features of graphs and images in neural networks. *Journal of Physics A: Mathematics and General*, 21: L813-L818, 1988.
- [57] W.A. Little. The existence of persistent states in the brain. *Mathematical Biosciences*, 19: 101-120, 1974.
- [58] W.A. Little and G.L. Shaw. Analytic study of the memory storage capacity of a neural network. *Mathematical Biosciences*, 39: 281-290, 1978.
- [59] M. Mézard, J.-P. Nadal, and G. Toulouse. Solvable models of working memories. *J. Physique*, 47: 1457-1468, 1986.
- [60] J.-P. Nadal, G. Toulouse, J.-P. Changeux, and S. Dehaene. Networks of formal neurons and memory palimpsests. *Europhysics Letters*, 1: 535-542, 1986.

- [61] G. Parisi. Asymmetric neural networks and the process of learning. *Journal of Physics A: Mathematics and General*, 19: L675-L680, 1986.
- [62] P. Peretto. Collective properties of neural networks: a statistical physics approach. *Biological Cybernetics*, 50: 51-62, 1984.
- [63] P. Peretto and J.-J. Niez. Stochastic dynamics of neural networks. *IEEE Trans. Syst. Man Cybern.*, 16: 73-83, 1986.
- [64] T. Poggio, V. Torre, and C. Koch. Computational vision and regularization theory. *Nature*, 347: 314-319, 1985.
- [65] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *Numerical Recipes*. Cambridge University Press, Cambridge, 1986.
- [66] S. Shinomoto. A cognitive associative memory. *Biological Cybernetics*, 57: 197-214, 1987.
- [67] H. Sompolinsky. The theory of neural networks: The Hebb rules and beyond. In: J.L. van Hemmen and I. Morgenstern, editors, *Heidelberg Colloquium on Glassy Dynamics*, pages 485-527, Heidelberg, 1987.
- [68] H. Sompolinsky, A. Crisanti, and H.J. Sommers. Chaos in random neural networks. *Physical Review Letters*, 61: 259-262, 1988.
- [69] A. Treves and E.T. Rolls. Computational constraints suggest the need for two distinct input systems to the hippocampal CA3 network. *Hippocampus*, 2: 625-639, 1992.
- [70] M.V. Tsodyks and M.V. Feigel'man. The enhanced storage capacity in neural networks with low activity level. *Europhysics Letters*, 6: 101-105, 1988.