

2. prednáška

Logické neuróny McCullocha a Pittsa

1.1 Úvodné poznámky

Ako už bolo poznamenané v predchádzajúcej prednáške, publikácia Warrena McCullocha a Waltera Pittsa [3] „*A logical calculus of the ideas immanent to nervous activity*“ z r. 1943 je medzníkom v rozvoji konekcionizmu. V tejto práci bolo s geniálnou jasnozrivosťou ukázané, že neurónové siete sú mocným modelovým prostriedkom, napr., že siete zložené z logických neurónov simulujú Boolove funkcie.

Elementárnou jednotkou McCullochovej a Pittsovej neurónovej siete je *logický neurón* (pozri ref. [4]) s binárnymi stavmi 1 a 0. Vstupný systém logického neurónu obsahuje *excitačné vstupy* (opísané binárnymi premennými x_1, x_2, \dots, x_n) a *inhibičné vstupy* (opísané binárnymi premennými $x_{n+1}, x_{n+2}, \dots, x_m$). Aktivita logického neurónu je jednotková, ak *vnútorný potenciál* neurónu definovaný ako rozdiel medzi sumou excitačných vstupných aktivít a inhibičných vstupných aktivít je väčší alebo rovný prahu ϑ , v opačnom prípade je nulová

$$y = \begin{cases} 1 & (x_1 + \dots + x_n - x_{1+n} - \dots - x_m \geq -\vartheta) \\ 0 & (x_1 + \dots + x_n - x_{1+n} - \dots - x_m < -\vartheta) \end{cases} \quad (2.1)$$

Pomocou jednoduchšej krokovej funkcie (1.1b) môžeme aktivitu y vyjadriť takto:

$$y = s \left(\underbrace{x_1 + \dots + x_n - x_{1+n} - \dots - x_m}_{\xi} + \vartheta \right) \quad (2.2)$$

Tento vzťah pre aktivitu logického neurónu môžeme alternatívne interpretovať tak, že excitačné aktivity vstupujú do neurónu cez spoje, ktoré sú ohodnotené jednotkovým váhovým koeficientom ($w = 1$), zatiaľ čo inhibičné aktivity vstupujú do neurónu cez spoje so záporným jednotkovým váhovým koeficientom ($w = -1$). Potom aktivitu logického neurónu môžeme vyjadriť takto

$$y = s \left(\underbrace{w_1 x_1 + \dots + w_m x_m}_{\xi} + \vartheta \right) = s \left(\sum_{i=1}^m w_i x_i + \vartheta \right) \quad (2.3)$$

Jednoduchá implementácia elementárnych Boolových funkcií disjunkcie, konjunkcie, implikácie a negácie je znázornená na obr. 2.1. Ako ilustratívny príklad študujeme funkciu disjunkcie pre $n = 2$, použitím formúl z definície logického neurónu dostaneme

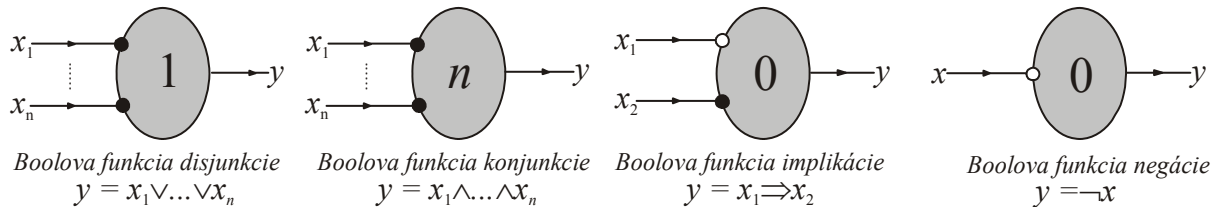
$$y_{OR}(x_1, x_2) = s(x_1 + x_2 - 1) \quad (2.4)$$

Funkčné hodnoty tejto Boolovej funkcie sú ukázané v tabuľke 2.1.

Tabuľka 2.1. Binárna Boolova funkcia disjuncie

#	x_1	x_2	$y_{OR}(x_1, x_2)$	$x_1 \vee x_2$
1	0	0	$s(-1)$	0
2	0	1	$s(0)$	1
3	1	0	$s(0)$	1
4	1	1	$s(1)$	1

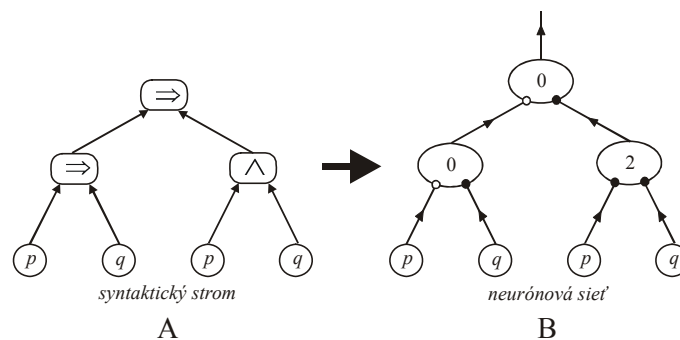
Z tabuľky vyplýva, že Boolova funkcia y_{OR} simuluje Boolovu funkciu disjuncie.



Obrázok 2.1. Tri realizácie logických neurónov na implementáciu Boolových funkcií disjuncií, konjunkcií a negácie. Excitačné spoje sú znázornené plným krúžkom, inhibičné prázdny krúžkom. Pre Boolovej funkcie implikácie je použité vetvenie excitačnej aktivity na dva vstupy, to znamená, že táto excitačná aktivita je započítaná dvakrát v (4.4) pre výpočte výstupnej aktivity neurónu.

2.1 Neurónové siete

Každá Boolova funkcia [1,2] je reprezentovaná pomocou syntaktického stromu, ktorým reprezentuje jej rekurentnú výstavbu inicializovanú Boolovými premennými a končiacu danou Boolovou funkciou (funkciou výrokovej logiky), pozri obr. 2.2, diagram A. Syntaktický strom je dôležitý keď hľadáme podformuly danej formuly, každý uzol stromu špecifikuje podformulu: najnižšie položené uzly reprezentujú triviálne podformuly p a q , nasledujúce dva vrcholy reprezentujú podformuly $p \Rightarrow q$ a $p \wedge q$, najvyššie položený vrchol – koreň stromu – reprezentuje samotnú formulu $(p \Rightarrow q) \Rightarrow (p \wedge q)$.



Obrázok 2.2. (A) Syntaktický strom Boolovej funkcie (výrokovej formuly) $(p \Rightarrow q) \Rightarrow (p \wedge q)$. Koncové vrcholy stromu reprezentujú Boolove premenné (výrokové premenné) p a q , vrcholy z nasledujúcich vrstiev sú priradené spojкам implikácie a konjunkcie. Vyhodnocovanie tohto stromu prebieha postupne zdola nahor. (B) Neurónová sieť obsahujúca logické neuróny spojok, ktoré sa vyskytujú v príslušnom syntaktickom strome diagramu A. Vidíme, že medzi syntaktickým stromom a príslušnou neurónovou sieťou existuje veľmi tesná previazanosť, ich topológia je identická, odlišujú sa len vo vrcholoch. Obrazne môžeme povedať, že neurónovú sieť pre Boolovu funkciu ϕ zostrojíme pomocou jej syntaktického stromu tak, že vrcholy zo syntaktického stromu, ktoré reprezentujú logické spojky, nahradíme príslušnými logickými neurónmi, koreň syntaktického stromu je zamenený.

Pomocou syntaktického stromu ľahko zostrojíme neurónovú sieť tak, že jednotlivé vrcholy reprezentujúce Boolove elementárne funkcie (logické spojky) nahradíme príslušnými logickými neurónmi podľa obr. 2.1. Na obr. 2.2 je znázornená táto konštrukcia neurónovej siete pre formulu $(p \Rightarrow q) \Rightarrow (p \wedge q)$. Tento postup konštrukcie neurónovej siete zosumarizujeme pomocou nasledujúcej vety.

Veta 2.1. Každá Boolova funkcia môže byť vyjadrená pomocou „neurónovej siete“ zloženej z logických neurónov vyjadrujúcich logické spojky.

Táto veta patrí medzi základné výsledky McCullocha a Pittsa [3,4], ukázali, že pomocou niekoľkých jednoduchých logických neurónov je možné zostrojiť neurónovú sieť, ktorá simuluje danú Boolovu funkciu. Môžeme teda hovoriť o tom, že neurónové siete s logickými neurónmi majú univerzálny charakter v doméne Boolových funkcií, vyššie naznačený konštruktívny postup reprezentuje všeobecne platnú metódu pre zostrojenie neurónovej siete, ktorá simuluje Boolove funkcie.

Aktivita i -teho neurónu v neurónovej sieti je vyjadrená zovšeobecnenou formulou (2.3)

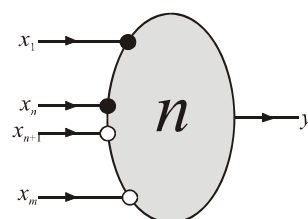
$$x_i = t \left(\sum_j w_{ij} x_j + \vartheta_i \right) \quad (2.5)$$

kde sumácia prebieha nad všetkými neurónmi, ktoré v neurónovej sieti predchádzajú i -ty neurón. Váhové koeficienty w_{ij} sú definované takto

$$w_{ij} = \begin{cases} 1 & (\text{spoj } j \rightarrow i \text{ má exitačný charakter}) \\ -1 & (\text{spoj } j \rightarrow i \text{ má inhibičný charakter}) \\ 0 & (\text{spoj } j \rightarrow i \text{ neexistuje}) \end{cases} \quad (2.6)$$

To znamená, že v neurónovej sieti váhové koeficienty sú fixne a dané a sú určené topológiou syntaktického stromu špecifikujúceho Boolovu funkciu.

Architektúra neurónovej siete, ktorá je zostrojená pre danú Boolovu funkciu môže byť podstatne zjednodušená na tzv. 3-vrstvovú neurónovú sieť, t. j. obsahuje vrstvu vstupných neurónov (ktoré len kopírujú vstupné aktivity, nie sú výpočtovými jednotkami), vrstva skrytých neurónov a posledná vrstva obsahujúca výstupný neurón. Táto architektúra je minimalistická a už zrejme nemôže byť zjednodušená. Ukážeme ako zostrojiť takúto neurónovú sieť pre danú Boolovu funkciu.



Obrázok 2.3. Logický neurón, ktorý simuluje klauzulu, ktorá obsahuje konjunkciu konečného prvku výrokových premenných alebo ich negácií, $y = x_1 \wedge \dots \wedge x_n \wedge \neg x_{n+1} \wedge \dots \wedge \neg x_m$.

Jednoduchým zovšeobecnením logických spojok možno ukázať, že logický neurón je schopný simulovať aj Boolovu funkciu, ktorá obsahuje konjunkcie premenných alebo ich negácií, $x_1 \wedge \dots \wedge x_n \wedge \neg x_{n+1} \wedge \dots \wedge \neg x_m$, pozri obr. 2.3. Táto Boolova funkcia sa rovná 1 len pre $x_1 = \dots = x_n = 1$ a $x_{n+1} = \dots = x_m = 0$, vo všetkých ostatných pravdivostných kombináciách argumentov jej hodnota je 0 (nepravdivý výrok)

$$val_{\tau}(x_1 \wedge \dots \wedge x_n \wedge \neg x_{n+1} \wedge \dots \wedge \neg x_m) = \begin{cases} 1 & (\text{pre } \tau = \tau_0) \\ 0 & (\text{pre } \tau \neq \tau_0) \end{cases} \quad (2.7)$$

kde $\tau_0 = (x_1/1, \dots, x_n/1, x_{n+1}/0, \dots, x_m/0)$ je špecifikácia pravdivostných hodnôt premenných. Lahko sa presvedčíme o tom, že táto klauzula je simulovaná logickým neurónom znázorneným na obr. 2.3. jeho výstupná aktivita je určená formulou

$$y = s(x_1 + \dots + x_n - x_{n+1} - \dots - x_m - n) \quad (2.8a)$$

Funkčná hodnota tejto funkcie sa rovná 1 vtedy, ak

$$x_1 + \dots + x_n - x_{n+1} - \dots - x_m \geq n \quad (2.8b)$$

Táto podmienka je dosiahnutá len vtedy, keď prvých n vstupných (excitačných) aktivít sa rovná 1 a ďalších $(m-n)$ vstupných (inhibičných) aktivít sa rovná 0.

V teórii Boolových funkcií je dokázana dôležitá vlastnosť Boolových funkcií, podľa ktorej každá Boolova funkcia môže byť prepísaná do ekvivalentného disjunktívneho tvaru [1,2]

$$\varphi = \bigvee_{\tau \text{ (val}_{\tau}(\varphi)=1)} x_1^{(\tau)} \wedge x_2^{(\tau)} \wedge \dots \wedge x_n^{(\tau)} \quad (2.9a)$$

kde

$$x_i^{(\tau)} = \begin{cases} x_i & (\text{ak } val_{\tau}(x_i) = 1) \\ \neg x_i & (\text{ak } val_{\tau}(x_i) = 0) \end{cases} \quad (2.9b)$$

Ako ilustračný príklad použitia tejto vlastnosti, študujme Boolovu funkciu, ktorej funkčné hodnoty sú zadané tabuľkou 2.2. V tejto tabuľke v riadkoch 3, 4 a 6 sú jednotkové funkčné hodnoty (Boolova funkcia je pre tieto tri hodnoty premenných pravdivá), v poslednom stĺpci sú uvedené v týchto riadkoch aj príslušné klauzuly zostrojene pomocou (2.9b). Použitím (2.9a) dostaneme „analytický“ tvar Boolovej funkcie určenej len tabuľkou jej funkčných hodnôt

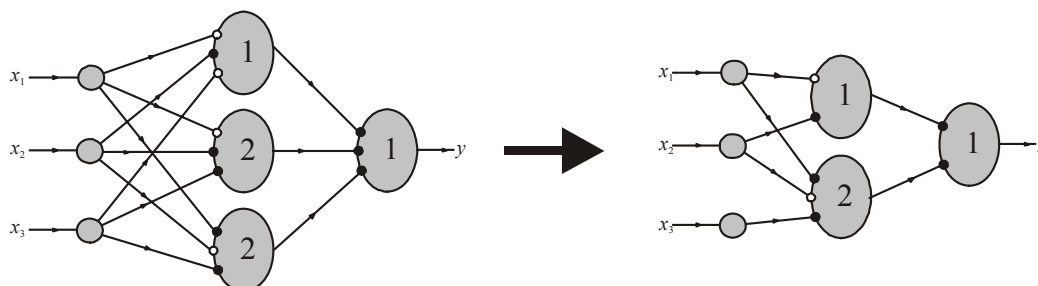
$$y = f(x_1, x_2, x_3) = (\bar{x}_1 \wedge x_2 \wedge \bar{x}_3) \vee (\bar{x}_1 \wedge x_2 \wedge x_3) \vee (x_1 \wedge \bar{x}_2 \wedge x_3) \quad (2.10)$$

Táto Boolova funkcia môže byť zjednodušená tak, že prvá a druhá klauzula sa zjednodušia

$$(\bar{x}_1 \wedge x_2 \wedge \bar{x}_3) \vee (\bar{x}_1 \wedge x_2 \wedge x_3) = \bar{x}_1 \wedge x_2 \wedge \underbrace{(\bar{x}_3 \vee x_3)}_1 = \bar{x}_1 \wedge x_2 \quad (2.11)$$

Potom

$$y = f(x_1, x_2, x_3) = (\bar{x}_1 \wedge x_2) \vee (x_1 \wedge \bar{x}_2 \wedge x_3) \quad (2.12)$$



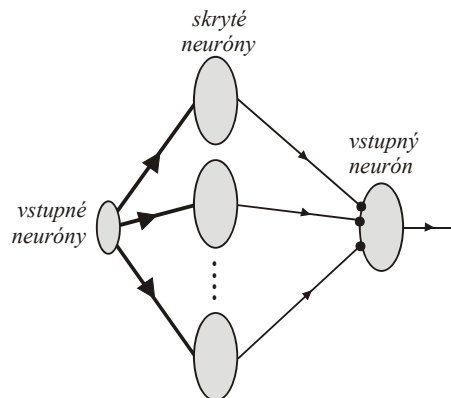
Obrázok 2.4. Trojvrstvová neurónová sieť, ktorá simuluje Boolovu funkciu zadanou primárne Tab. 2.2, z ktorej je zostrojený pomocou formule (2.9) ja „analytický“ tvar (2.10). Skryté neuróny reprezentujú jednotlivé klauzuly z Tab. 2.2, ich disjunkcia je realizovaná pomocou výstupného neurónu.. Táto neurónová sieť môže byť zjednodušená tak, že prvé dve klauzuly sa spoja do jednej jednoduchšej klauzule, pozri (2.11-12).

Klauzule $x_1^{(\tau)} \wedge x_2^{(\tau)} \wedge \dots \wedge x_n^{(\tau)}$ môžeme vyjadriť jedným logickým neurónom, pozri obr. 2.3. Výstupy z týchto neurónov spojíme do disjunkcie pomocou neurónu reprezentujúceho disjunkciu (pozri obr. 2.1). Potom neurónová sieť, reprezentujúca Boolovu funkciu (2.10) má tvar znázornený na obr. 2.4.

Výsledky z tohto ilustračného príkladu sú zosumarizované pomocou tejto vety.

Veta. 2.2. *Lubovoľná Boolova funkcia f je simulovaná pomocou 3-vrstvovej neurónovej siete.*

Všeobecný tvar 3-vrstvovej neurónovej siete je znázornený na obr. 2.5.



Obrázok 2.5. Schématické znázornenie 3-vrstvovej neurónovej siete. Idúc zľava doprava, prvá je vstupná vrstva obsahujúca vstupné neuróny, ktoré nie sú výpočtové elementy, ale len formálne reprezentujú vstupné aktivity. Druhá vrstva obsahuje skryté neuróny, ktoré reprezentujú jednotlivé klauzule danej Boolovej funkcie. Tretia (posledná) vrstva obsahuje výstupný neurón, ktorý vykonáva disjunkciu aktivít zo skrytých neurónov.

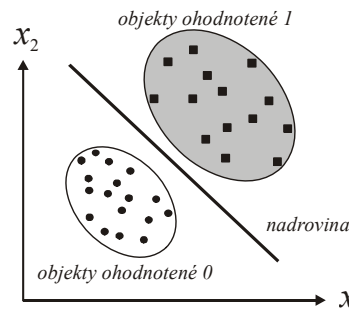
Tabuľka 2.2. Zadanie funkčných hodnôt Boolovej funkcie.

#	x_1	x_2	x_3	$y = f(x_1, x_2, x_3)$	klauzula
1	0	0	0	0	-
2	0	0	1	0	-
3	0	1	0	1	$\bar{x}_1 \wedge x_2 \wedge \bar{x}_3$
4	0	1	1	1	$\bar{x}_1 \wedge x_2 \wedge x_3$
5	1	0	0	0	-
6	1	0	1	1	$x_1 \wedge \bar{x}_2 \wedge x_3$
7	1	1	0	0	-
8	1	1	1	0	-

Poznamenajme, že podľa vety 2.2 3-vrstvové neurónové siete obsahujúce logické neuróny sú *univerzálne výpočtové zariadenia* pre doménu Boolových funkcií. Tento výsledok predznačil moderný výsledok neurónových sietí z počiatku 90. rokov minulého storočia, podľa ktorého trojvrstvové dopredné neurónové siete so spojenou aktivačnou funkciou (1.1) majú vlastnosť univerzálneho aproximátora (pozri kapitolu X.X). Navyše, pretože dôkaz vety bol konštruktívny, poznáme jednoduchý postup, ako túto sieť systematickým spôsobom zostrojiť pre ľubovoľnú Boolovu funkciu. Žiaľ, problém minimalnosti takto zostrojenej siete nie je vetou riešený. Vo všeobecnosti môže existovať sieť, ktorá pre danú Boolovu sieť obsahuje menej logických (skrytých) neurónov ako tá, ktorá bola zostrojená systematickým postupom z dôkazu vety. V teórii Boolových funkcií boli vypracované metódy (napr. Quinova a McCluskeyho metóda [xx]), ktoré minimalizujú Boolovu funkciu zostrojené pomocou (2.9a)

vzhľadom k veľkosti konjunktívnych členov (klauzúl). Potom môžeme zostrojiť „optimalizovanú“ neurónovú sieť pomocou takto minimalizovanej Boolovej funkcie..

Môžeme si položiť otázku, aké Boolove funkcie je schopný logický neurón vyjadriť [5]? Táto otázka sa dá pomerne jednoducho vyriešiť pomocou geometrickej interpretácie výpočtu prebiehajúceho v logickom neuróne. Výpočtová funkcia logického neurónu rozdeľuje priestor vstupov na dva polpriestory pomocou roviny $w_1x_1 + w_2x_2 + \dots + w_nx_n = \vartheta$, pre koeficienty $w_i=0, \pm 1$. Hovoríme, že Boolova funkcia $f(x_1, x_2, \dots, x_n)$ je *lineárne separovateľná*, ak existuje taká rovina $w_1x_1 + w_2x_2 + \dots + w_nx_n = \vartheta$, ktorá separuje priestor vstupných aktivít tak, že v jednej časti priestoru sú vrcholy ohodnotené 0, zatiaľ čo v druhej časti priestoru sú vrcholy ohodnotené 1 (pozri obr. 2.6).



Obrázok 2.6. Schematické znázornenie pojmu lineárnej separovateľnosti, kde okrúhle a štvorcové objekty sú separované nadrovinou $w_1x_1 + \dots + w_nx_n - \vartheta = 0$ tak, že v jednom polpriestore sú objekty jedného druhu, zatiaľ čo v druhom polpriestore sú objekty druhého druhu.

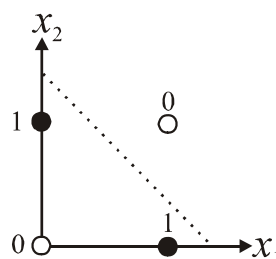
Veta 2.3. *Logický neurón je schopný simulovať len tie Boolove funkcie, ktoré sú lineárne separovateľné.*

Klasický príklad Boolovej funkcie, ktorá nie je lineárne separovateľná, je výroková spojka exkluzívna disjunkcia, ktorá je formálne definovaná ako negácia ekvivalencie $(x \oplus y) \Leftrightarrow \neg(x \equiv y)$, v informatickej literatúre sa obvykle označuje ako Boolova funkcia *XOR*, $\varphi_{XOR}(x, y) = x \oplus y$, jej funkčné hodnoty sú uvedené v tab. 2.3.

Tabuľka 2.3. Boolova funkcia XOR

#	x	y	$\varphi_{XOR}(x, y)$
1	0	0	0
2	0	1	1
3	1	0	1
4	1	1	0

Ak si jej funkčné hodnoty vynesieme do stavového priestoru $x - y$ dostaneme obr. 2.7, z ktorého jasne vyplýva, že táto funkcia nie je lineárne separovateľná.

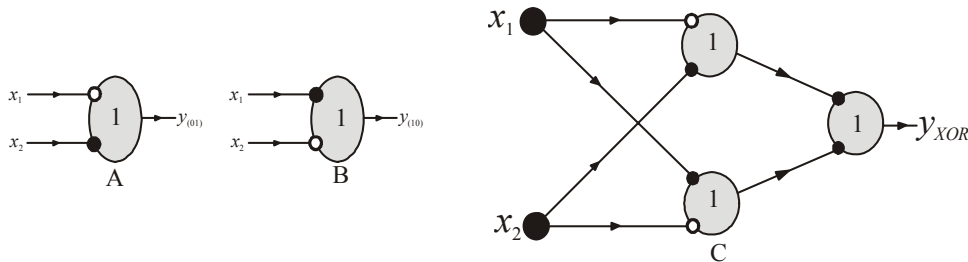


Obrázok 2.7. Znázornenie objektov Boolovej funkcie XOR v stavovom priestore jej argumentov. Z obrázku jasne plynie, že neexistuje priamka, ktorá by separovala celú rovinu na dve polroviny tak, že objekty (otvorené krúžky) s ohodnotením 0 budú v jednej polrovine, zatiaľ čo objekty s ohodnotením 1 (čierny krúžky) budú v druhej polrovine.

Použitím techniky z predchádzajúcej časti tejto kapitoly, zostrojíme neurónovú sieť, ktorá simuluje túto elementárnu lineárne neseparovateľnú Boolovu funkciu. Pomocou Tab. 2.3. zostrojíme jej ekvivaletnú funkciu pomocou klauzúl

$$\Phi_{XOR}(x_1, x_2) = (\neg x_1 \wedge x_2) \vee (x_1 \wedge \neg x_2) \quad (2.13)$$

Potom neurónová sieť zostrojená pomocou tejto funkcie má tvar znázornený na obr. 2.8.

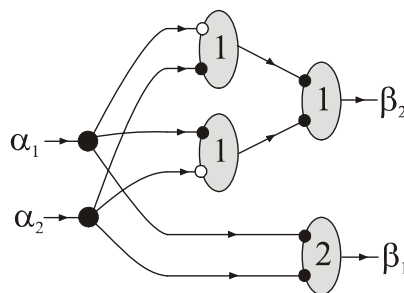


Obrázok 2.8. Diagramy A a B znázorňujú jednotlivé klauzuly z formuly (2.11). Diagram C reprezentuje 3-vrstvovú neurónovú sieť, ktorá ako skryté neuróny obsahuje neuróny z diagramu A a B, výstupný neurón reprezentuje disjunkciu výstupných aktivít zo skrytých neurónov.

Príklad 2.1. Zostrojme neurónovú sieť, ktorá simuluje sčítanie dvoch binárnych číslíc:

$$\begin{array}{r} \alpha_1 \\ \alpha_2 \\ \hline \beta_1 \beta_2 \end{array}$$

kde jednotlivé binárne premenné výsledku – sumy sú určené takto: $\beta_2 = \alpha_1 \oplus \alpha_2$ a $\beta_1 = \alpha_1 \wedge \alpha_2$. Druhú výstupnú premennú β_2 prepíšeme do disjunktívnej formy $\beta_2 = (\neg \alpha_1 \wedge \alpha_2) \vee (\alpha_1 \wedge \neg \alpha_2)$, príslušná neurónová sieť je znázornená na obrázku 2.9.



Obrázok 2.9. Neurónová sieť vykonávajúca súčet dvoch binárnych číslíc.

V predchádzajúcej časti tejto kapitoly bolo ukázané, že samotné logické neuróny sú schopné korektné klasifikovať len lineárne separovateľné Boolove funkcie. Toto podstatné obmedzenie logických neurónov môže byť odstránené pomocou logických neurónov vyšších rádov [5], ktorých aktivita je určená zovšeobecnením formuly (2.3) o členy vyšších rádov

$$y = s \left(\underbrace{\sum_{i=1}^n w_i x_i + \sum_{\substack{i,j=1 \\ (i < j)}}^n w_{ij} x_i x_j + \dots + \vartheta}_{\xi} \right) \quad (2.14)$$

Ak vnútorný potenciál neurónu ξ je určený len ako lineárna kombinácia vstupných aktivít (t.j. len prvou sumou), potom logický neurón je štandardný a nazýva sa "logický neurón prvého rádu". Ak tento potenciál neurónu ξ obsahuje aj kvadratické a prípadne aj ďalšie členy, potom sa nazýva "logický neurón vyššieho rádu". Podľa Minského a Paperta [5] platí nasledujúca veta, ktorá hovorí o tom, že perceptróny vyššieho rádu sú schopné simulovať aj množiny objektov, ktoré nie sú lineárne separovateľné.

Veta 2.4. Ľubovoľná Boolova funkcia f je simulovaná logickým neurónom vyššieho rádu.

Táto veta hovorí o tom, že každá Boolova funkcia môže byť simulovaná logickým neurónom vyššieho rádu, existujú také váhové koeficienty a prah, že pre každú špecifikáciu premenných x_1, x_2, \dots, x_n vypočítaná binárna aktivita sa rovná požadovanej hodnote.

Príklad 2.2. Ako ilustračný príklad vety študujme Boolovu funkciu XOR , ktorá nie je lineárne separovateľná a ktorej funkčné hodnoty sú uvedené v tabuľke 4.1. Nech aktivita logického neurónu je určená pomocou kvadratického potenciálu (čiže sa jedná o logický neurón druhého rádu)

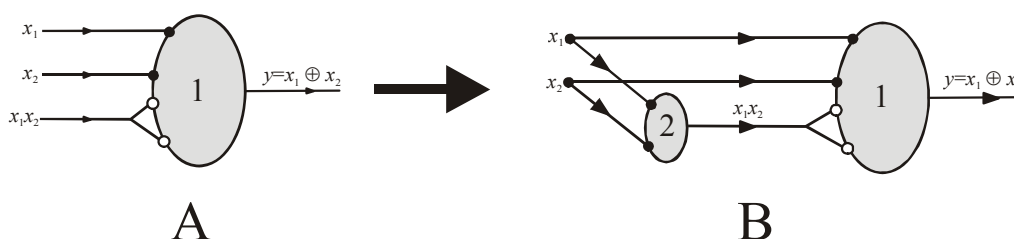
$$y = s \left(\underbrace{w_1 x_1 + w_2 x_2 + w_{12} x_1 x_2}_{\xi} - \vartheta \right) \quad (2.15)$$

Pre jednotlivé riadky tabuľky 2.3 funkcie XOR platí

$$\begin{aligned} -\vartheta &< 0 \\ w_2 &-\vartheta \geq 0 \\ w_1 &-\vartheta \geq 0 \\ w_1 + w_2 + w_{12} - \vartheta &< 0 \end{aligned} \quad (2.16)$$

Postupným riešením týchto nerovniíc dostaneme, že váhové koeficienty a prah majú napr. takéto riešenie

$$\vartheta = 1, w_1 = w_2 = 1, w_{12} = -2 \quad (2.17)$$



Obrázok 2.10. (A) Znáročenie logického neurónu druhého rádu simulujúceho funkciu XOR , kde excitačné vstupné aktivity sú binárne premenné x_1 a x_2 , inhibičná aktivita je priradená premennej druhého rádu $x_1 x_2$. Výstupná aktivita z je určená pomocou krokovej funkcie, $z = s(x_1 + x_2 - 2x_1 x_2 - 1)$, priamou verifikáciou sa presvedčíme, že simuluje funkciu XOR . Na záver poznamenajme, že "vidlička" inhibičného vstupu znamená, že vstupná aktivita je počítaná dvakrát. (C) Transformácia logického neurónu druhého rádu, ktorý implementuje spojku XOR (diagram A), na neurónovú sieť, ktorá obsahuje len logické neuróny prvého rádu (diagram B). Táto transformácia je založená na výpočte súčinu $x_1 x_2$ pomocou logického neurónu pre spojku konjunkcie, výstup

z tohto neurónu je potom dvojnásobný inhibičný vstup do pravého neurónu. Táto architektúra je asi najjednoduchšou neurónovou sieťou, ktorá implementuje spojku XOR pomocou logických neurónov prvého rádu (porovnaj s obr. 4.7).

V tomto ilustračnom príklade bolo ukázané, že lineárne neseparovateľná binárna spojka disjunkcie s vylúčením (XOR) je implementovateľná pomocou logického neurónu, ktorý má tri vstupy pre x_1 , x_2 a x_1x_2 . V tejto súvislosti sa vynára otázka výpočtu súčinu x_1x_2 , ktorý však môže byť jednoducho vykonaný pomocou spojky konjunkcie, $x_1x_2 = x_1 \wedge x_2$. Ak táto operácia bude uskutočnená pomocou logického neurónu pre konjunkciu (pozri obr. 2.10, diagram B), potom môžeme vytvoriť jednoduchú sieť obsahujúcu len dva neuróny, pričom na vstupe máme len vstupné aktivity x_1 a x_2 .

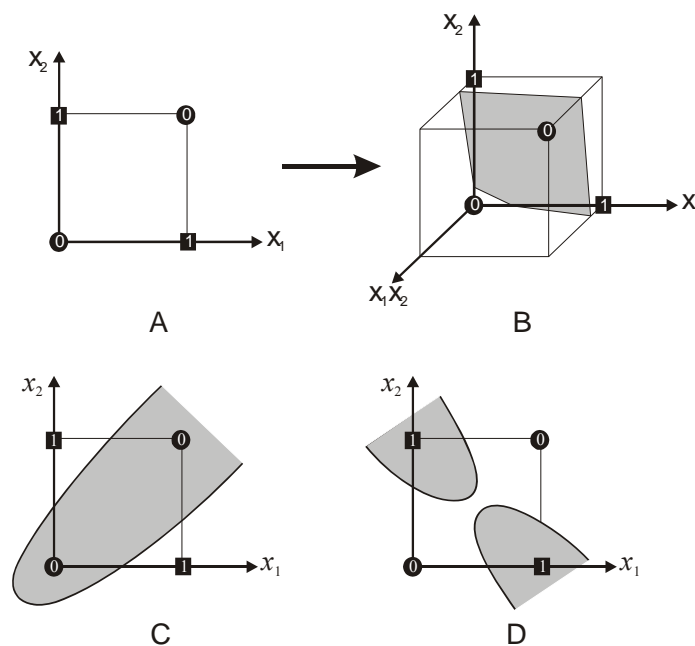
Výsledky tohto príkladu 2.2 ukážujú, že logický neurón druhého rádu je schopný korektne klasifikovať Boolovu funkcia XOR, ktorá je lineárne neseparovateľná v 2-rozmernom fázovom priestore x_1 - x_2 , ale je už lineárne separovateľná v 3-rozmernom fázovom priestore x_1 - x_2 - x_1x_2 , pozri obr. 2.11.

Pojem lineárnej separovateľnosti Boolových funkcií je ľahko zovšeobecniteľný na kvadratickú (kubickú,...) separovateľnosť pomocou kvadratickej (kubickej,...) nadplochy.

Definícia 2.1. Boolova funkcia f sa nazýva *kvadraticky separovateľná*, ak existujú také váhové koeficienty w_i , w_{ij} a prahový faktor ϑ , že pre každú špecifikáciu premenných x_1, x_2, \dots, x_n platí

$$y_{req}(x_1, x_2, \dots, x_n) = 1 \Rightarrow \sum_{i=1}^n w_i x_i + \sum_{\substack{i,j=1 \\ (i < j)}}^n w_{ij} x_i x_j \geq \vartheta$$

$$y_{req}(x_1, x_2, \dots, x_n) = 0 \Rightarrow \sum_{i=1}^n w_i x_i + \sum_{\substack{i,j=1 \\ (i < j)}}^n w_{ij} x_i x_j < \vartheta$$
(2.18)



Obrázok 2.11. Diagramatická reprezentácia Boolovej funkcie XOR. (A) Ak je funkcia XOR reprezentovaná v 2-rozmernom stavovom priestore x_1 - x_2 , objekty s jednotkovou klasifikáciou nie sú lineárne separovateľné od objektov s nulovou klasifikáciou. (B) Ak znázorníme XOR funkciu v 3-rozmernom fázovom priestore x_1 - x_2 - x_1x_2 ,

potom existuje v tomto priestore nadrovina, ktorá navzájom separuje objekty s rôznou klasifikáciou. Priemet tejto nadroviny do roviny poskytuje kvadratickú čiaru, ktorá separuje objekty, pozri diagramy C a D.

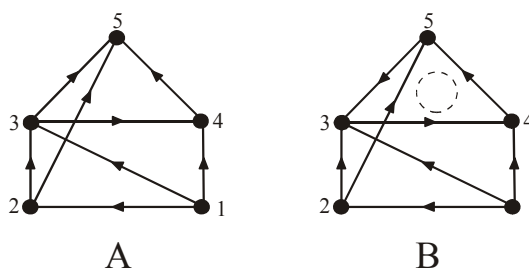
Princíp použiteľnosti logického neurónu pre klasifikáciu Boolových funkcií, ktoré nie sú lineárne separovateľné, spočíva v skutočnosti, že zavedením nových "kvadratických premenných" sa objekty stanú lineárne separovateľnými. Naznačený postup môže byť zovšeobecnený do nasledujúcej vety.

2.1.1 Formálna špecifikácia neurónovej siete

Z našich predchádzajúcich úvah vyplýva, že koncepcia neurónovej siete patrí medzi fundamentálne pojmy teórie neurónových sietí (nielen tých, ktoré obsahujú logické neuróny). Neurónová sieť je definovaná pomocou usporiadanej trojice

$$\mathcal{N} = (G, \mathbf{w}, \mathfrak{G}) \quad (2.19)$$

kde G je súvislý orientovaný graf, \mathbf{w} je matica váhových koeficientov a \mathfrak{G} je vektor prahových koeficientov.



Obrázok 2.12. Orientované súvislé grafy, ktoré reprezentujú topológiu neurónovej siete. V týchto grafoch vrchol 1 reprezentuje vstupný neurón, vrcholy 2, 3, 4 reprezentujú skryté neuróny a vrchol 5 reprezentuje výstupný neurón. Diagram A znázorňuje acyklický graf, diagram B reprezentuje cyklický graf (vznikol z ľavého grafu A zmenou orientácie hrany 3-4).

Zatiaľ sme nepoužili čas v explicitnej forme pri popise siete zloženej z logických neurónov. Budeme predpokladať, že čas v neurónovej sieti je diskretný a je špecifikovaný nezáporným celým číslom, aktivity neurónovej siete v čase t označíme $\mathbf{x}^{(t)}$, v čase $t = 0$ vektor $\mathbf{x}^{(0)}$ špecifikuje *počiatočné aktivity* neurónovej siete. v. Vzťah (2.5) pre aktivitu i -teho neurónu v čase t je má tvar

$$x_i^{(t)} = s \left(\sum_j w_{ij} x_j^{(t-1)} - \mathfrak{G}_i \right) \quad (2.20)$$

kde sumácia prebieha cez neuróny, ktoré v sieti "predchádzajú" i -ty neurón, pričom aktivity týchto neurónov sú z predchádzajúceho času $t-1$. Pre ilustráciu študujme neurónovú sieť znázornenú na obr. 2.12, kde neurónová sieť je špecifikovaná acyklickým orientovaným grafom. Aktivity jednotlivých neurónov sú určené pomocou (2.20) takto

$$\begin{aligned} x_1^{(t)} &= \text{externý vstup} \\ x_2^{(t)} &= s(-x_1^{(t-1)} - 0) \\ x_3^{(t)} &= s(x_1^{(t-1)} + x_2^{(t-1)} - 2) \quad (t = 1, 2, \dots, t_{max}) \\ x_4^{(t)} &= s(-x_1^{(t-1)} + x_3^{(t-1)} - 1) \\ x_5^{(t)} &= s(x_2^{(t-1)} + x_3^{(t-1)} + x_4^{(t-1)} - 1) \end{aligned} \quad (2.21)$$

Poznamenajme, že ako dôsledok acykličnosti neurónovej siete je skutočnosť, že pri výpočte $x_i^{(t)}$ potrebujeme poznať aktivity neurónov, ktoré daný neurón predchádzajú a sú z predošlého časového momentu. Táto skutočnosť je hlavným dôvodom prečo sa neurónové siete reprezentované acyklickým orientovaným grafom nazýva *dopredné neurónové siete*.

Binárne aktivity jednotlivých neurónov môžu byť vyjadrené pomocou vektora

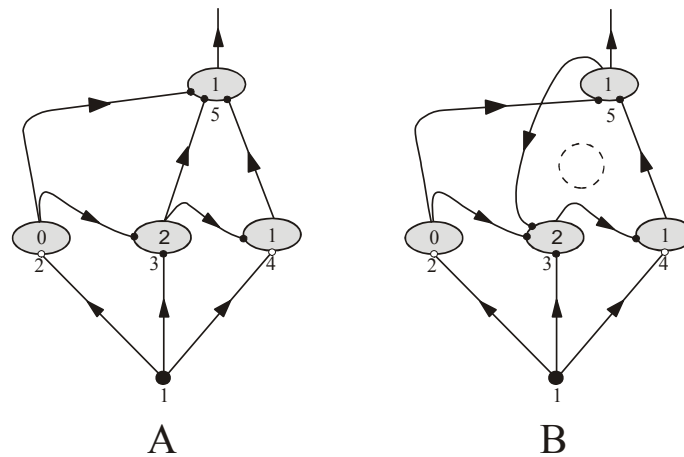
$$\mathbf{x} = \mathbf{x}_I \oplus \mathbf{x}_H \oplus \mathbf{x}_O \quad (2.22)$$

kde \mathbf{x}_I je vektor aktivít vstupných neurónov, \mathbf{x}_H je vektor aktivít skrytých neurónov a \mathbf{x}_O je vektor výstupných aktivít. Neurónovú sieť \mathcal{N} môžeme chápať ako funkciu, ktorá zobrazuje vektor aktivít $\mathbf{x}^{(t-1)}$ v čase $t-1$ na vektor aktivít $\mathbf{x}^{(t)}$ v čase t

$$\mathbf{x}^{(t)} = F(\mathbf{x}^{(t-1)}; \mathcal{N}) \quad (2.23)$$

kde funkcia F obsahuje ako „parameter“ špecifikáciu neurónovej siete \mathcal{N} .

Neurónové siete môžeme rozdeliť na dve veľké podskupiny podľa topológie grafu G z (2.19). Ak graf G je acyklický, potom neurónová sieť sa nazýva *dopredná*; v opačnom prípade, ak graf G je cyklický, potom sieť sa nazýva *rekurentná* (pozri obr. 2.13).



Obrázok 2.13. Neurónové siete, ktoré sú špecifikované orientovanými grafmi z obr. 2.12. (A) Dopredná špecifikovaná acyklickým grafom znázorneným diagramom A na obr. neurónová sieť 2.12, (B) rekurentná neurónová sieť špecifikovaná cyklickým grafom B na obr. 2.12.

Pre úplnosť budeme predpokladať, že aktivity neurónov 2-4 pre $t=1$ sú nulové. Nech vstupné aktivity sú špecifikované binárnym reťazcom dĺžky $t_{max}=10$, $\mathbf{x}_I=(1101101010)$, aktivity ostatných neurónov v doprednej neurónovej sieti z obr. 2.13, diagram A sú uvedené v nasledujúcej tabuľke pre časy $1 \leq t \leq 10$.

t	x_1	x_2	x_3	x_4	x_5
1	1	0	0	0	0
2	1	0	0	0	0
3	0	1	0	0	0
4	1	1	0	0	1
5	1	0	1	0	1
6	0	0	0	0	1
7	1	1	0	0	0
8	0	0	1	0	1
9	1	1	0	1	1
10	0	0	1	0	1

Vo všeobecnosti môžeme konštatovať, že neurónová sieť tvorí "parametrické" (excitačný alebo inhibičný charakter spojov a prahy jednotlivých neurónov) zobrazenie postupnosti vstupných aktivít na nejakú postupnosť výstupných aktivít

$$(1101101010) = \tilde{F}(\square 001110111, \text{parametre siete}) \quad (2.24)$$

Rekurentné neurónové siete sú určené cyklickým orientovaným grafom, pozri obr. 2.13, diagram B. V tomto prípade môžeme povedať, že neurónová sieť má pamäť, v dôsledku existencie orientovaných cyklov v rekurentných sieťach, niektoré spôsoby závislosti aktivít od aktivít iných neurónov sa môžu neustále opakovať. Napríklad, pri výpočte aktivity neurónu 2 v čase t potrebujeme poznať aktivity neurónov 1, 2 a 5 v čase $t-1$, avšak v dôsledku existencie cyklu, aktivita neurónu 5 v čase $t-1$ vyžaduje poznať aktivitu neurónov 2 a 4 v čase $t-2$. Z tejto jednoduchej diskusie vyplýva, že aktivita neurónu 5 v čase t je určená aktivitami neurónu v časoch $t-1$ a $t-2$. V ďalších krokoch sa môžeme toto „okno do minulosti“ zväčšovať, túto skutočnosť špecifickú pre rekurentné siete nazývajú „pamäť“ rekurentných sietí. Aktivity neurónov z rekurentnej siete sú určené pomocou (2.19) takto

$$\begin{aligned} x_1^{(t)} &= \text{externý vstup} \\ x_2^{(t)} &= s(-x_1^{(t-1)} - 0) \\ x_3^{(t)} &= s(x_1^{(t-1)} + x_2^{(t-1)} + x_5^{(t-1)} - 2) \quad (t = 1, 2, \dots) \\ x_4^{(t)} &= s(-x_1^{(t-1)} + x_3^{(t-1)} - 1) \\ x_5^{(t)} &= s(x_2^{(t-1)} + x_4^{(t-1)} - 1) \end{aligned} \quad (2.25)$$

Pre podobnú sekvenciu vstupných aktivít, aká bola použitá v predchádzajúcom príklade, $x_1=(1101101010)$, dostaneme použitím vzťahov (2.22) aktivity neurónov siete, ktoré sú uvedené v nasledujúcej tabuľke

t	x_1	x_2	x_3	x_4	x_5
1	1	0	0	0	\square
2	1	0	0	0	0
3	0	0	0	0	0
4	1	1	0	0	0
5	1	0	1	0	1
6	0	0	1	0	0
7	1	1	0	1	0
8	0	0	1	0	1
9	1	1	0	1	0
10	0	0	1	0	1

Podobne ako v predchádzajúcom príklade doprednej neurónovej siete (pozri obr. 2.13, diagram A), aj rekurentná neurónová sieť môže byť interpretovaná ako zobrazenie vstupného reťazca $x_1=(1101101010)$ na výstupný reťazec $x_5=(\square 000100101)$, kde symbol \square reprezentuje prázdny znak..

2.2 Konečnostavový stroj [6]

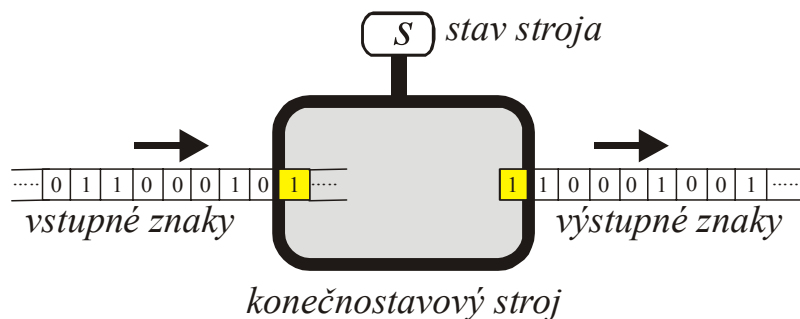
Konečnostavový stroj je znázornený na obr. 2.14., tento stroj pracuje v diskretných časových $1, 2, \dots, t, t+1, \dots$ okamžikoch. Obsahuje dve pásky, vstupných symbolov resp. výstupných symbolov, pričom výstupné symboly sú určené na základe vstupných symbolov a vnútorných stavov s stroja

$$stav_{t+1} = f(stav_t, vstupný\ symbol_t) \quad (2.26a)$$

$$výstupný\ symbol_{t+1} = g(stav_t, vstupný\ symbol_t) \quad (2.26b)$$

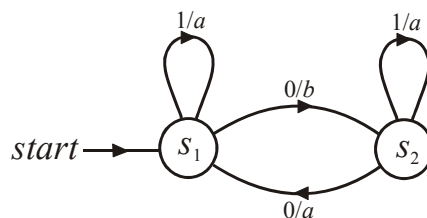
kde funkcie f a g špecifikujú daný stroj a sú chápané ako jeho základná charakteristika:

- (1) prechodová funkcia f určuje nasledujúci stav na základe aktuálneho stavu a vstupného symbolu,
- (2) výstupná funkcia g určuje výstupný symbol na základe aktuálneho stavu a vstupného symbolu.



Obrázok 2.14. Konečnostavový stroj pracuje v diskretných časových okamžikoch $1, 2, \dots, t, t+1, \dots$. Obsahuje dve hlavice, jednu na čítanie vstupnej pásky, pomocou ktorej v danom časovom okamžiku t identifikuje vstupný symbol, pomocou výstupnej hlavice vytlačí v čase $t+1$ výstupný symbol. V každom časovom okamžiku t je stroj v určitom vnútornom stave s , pričom v čase $t+1$ je vnútorný stav určený vnútorným stavom v čase t a vstupným symbolom taktiež v čase t (pozri vzťahy (2.26a-b)).

Definícia 2.2. Konečnostavový stroj (s výstupom, nazývaný taktiež *Mealyho automat*) je definovaný pomocou usporiadanej šesticice $M = (S, I, O, f, g, s_{mi})$, kde $S = \{s_1, \dots, s_m\}$ je konečná množina stavov, $I = \{i_1, i_2, \dots, i_n\}$ je konečná množina vstupných symbolov, $O = \{o_1, o_2, \dots, o_p\}$ je konečná množina výstupných symbolov, $f : S \times I \rightarrow S$ je prechodová funkcia, $g : S \times I \rightarrow O$ je výstupná funkcia a $s_{mi} \in S$ je počiatkový stav.



Obrázok 2.15. Príklad konečnostavového stroja, ktorý obsahuje dva stavy, $S = \{s_1, s_2\}$, dva vstupné znaky, $I = \{0, 1\}$, dva výstupné znaky, $O = \{a, b\}$ a počiatkový stav je s_1 . Prechodová a výstupná funkcia je špecifikovaná pomocou Tabuľky 2.4

Tabuľka 2.4. Prechodová a výstupná funkcia konečnostavového stroja z obr. 2.15.

stav	<i>f</i>		<i>g</i>	
	prechodová funkcia		výstupná funkcia	
	0	1	0	1
<i>s</i> ₁	<i>s</i> ₂	<i>s</i> ₁	<i>b</i>	<i>a</i>
<i>s</i> ₂	<i>s</i> ₁	<i>s</i> ₂	<i>a</i>	<i>a</i>

Prechodová a výstupná funkcia môžu byť použité na konštrukcia formálneho obvodu, ktorý simuluje daný konečnosťavový stroj, pozri obr. 2.16.

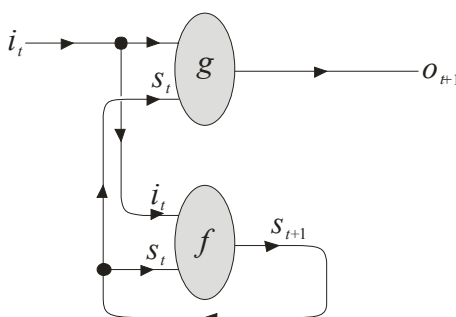
Postupnosti stavov a výstupných znakov pre konečnosťavový stroj z obr. 2.15 sú špecifikované Tabuľkou 2.5 pre vstupnú postupnosť znakov (100111010...). Zariadenie môže byť interpretované ako zobrazenie vstupného reťazca znakov na výstupný reťazec znakov

$$G \left(\underbrace{100111010\dots}_{\text{vstupný reťazec } x}; f, g \right) = \underbrace{\square abaaaabaa\dots}_{\text{výstupný reťazec } y}$$

kde symbol \square vo výstupnom reťazci znamená „prázdny znak“, znaky výstupného reťazca sú posunuté o jeden časový krok – takt, vzhľadom k vstupnému reťazcu. Zobrazenie G obsahuje dve funkcie f a g , ktoré určujú „topológiu“ konečnosťavového stroja. Pre konštrukciu vzťahu medzi neurónovou sieťou a konečnosťavového stroja budeme tento prístup špecifikovať takto: nech $\mathbf{i} = i^{(1)}i^{(2)}i^{(3)}\dots i^{(t)}\dots$, $\mathbf{o} = o^{(2)}o^{(3)}o^{(4)}\dots o^{(t+1)}\dots$ a $\mathbf{s} = s^{(1)}s^{(2)}s^{(3)}\dots s^{(t)}\dots$ sú reťazce vstupných znakov, výstupných znakov, resp. stavov stroja (pozri Tabuľku 2.5). Jednotlivé znaky z týchto reťazcov sú navzájom v dvoch funkčných reláciách (pozri obr. 2.16)

$$s^{(t+1)} = f(s^{(t)}, i^{(t)}) \quad (2.27a)$$

$$o^{(t+1)} = g(s^{(t)}, i^{(t)}) \quad (2.27b)$$



Obrázok 2.16. Schématická reprezentácia konečnosťavového stroja reprezentovaného prechodovou a výstupnou funkciou f resp. g (pozri formuly (2.27a-b)).

Prvá formula (2.27a) určuje nasledujúci stav $s^{(t+1)}$ pomocou prechodovej funkcie f , vstupného symbolu $i^{(t)}$ a stavu $s^{(t)}$. Podobne, druhá formula (2.27b) určuje nový výstupný symbol $o^{(t+1)}$ pomocou výstupnej funkcie g , predchádzajúceho stavu $s^{(t)}$ a vstupného symbolu $i^{(t)}$. Hovoríme, že neurónová sieť je *ekvivalentná* s konečnosťavovým strojom, ak odozva výstupom oboch výpočtových zariadení je identická na ľubovoľný rovnaký vstup. Pre ekvivalenciu neurónovej siete s konečnosťavovým strojom je nepodstatné akým spôsobom sa vstupné symboly zobrazujú na výstupné symboly, t. j. akým spôsobom prebiehajú výpočty pri konštrukcii výstupného reťazca, podstatné je rovnosť výstupných reťazcov pre rovnaký vstupný reťazec pre obe zariadenia.

Tabuľka 2.5. Postupnosť vstupných znakov, stavov a výstupných znakov pre konečnosťavový stroj z obr. 2.15.

<i>vstupný symbol</i>	1	0	0	1	1	1	0	1	0	..
<i>stav</i>	s_1	s_1	s_2	s_1	s_1	s_1	s_1	s_2	s_2	..
<i>výstupný symbol</i>	□	a	b	a	a	a	a	b	a	a

Veta 2.5 [4]. Každá neurónová sieť môže byť reprezentovaná ekvivalentným konečnosťavovým strojom s výstupom.

Dôkaz tejto vety je konštruktívny a jednoduchý, ukážeme ako pre neurónovú sieť možno zostrojiť jednotlivé zložky z definície 2.2, $M = (S, I, O, f, g, s_{ini})$.

- (1) Množina stavov S je vytvorená všetkými možnými binárnymi vektormi \mathbf{x}_H , $S = \{\mathbf{x}_H\}$. Nech neurónová sieť má n_H skrytých neurónov a n_O výstupných neurónov, potom kardinalita (počet stavov) množiny S je 2^{n_H} .
- (2) Množina vstupných symbolov je tvorená všetkými možnými binárnymi vektormi \mathbf{x}_I , $I = \{\mathbf{x}_I\}$, kardinalita tejto množiny je 2^{n_I} , kde n_I je počet vstupných neurónov.
- (3) Množina výstupných symbolov je tvorená všetkými možnými binárnymi vektormi \mathbf{x}_O , $O = \{\mathbf{x}_O\}$, kardinalita tejto množiny je 2^{n_O} .
- (4) Funkcia $f : S \times I \rightarrow S$ priradí každému stavu a vstupnému symbolu nový stav. Táto funkcia je špecifikovaná zobrazením (2.23) produkovaného neurónovou sieťou

$$\mathbf{x}_H^{(t+1)} = F(\mathbf{x}_I^{(t)} \oplus \mathbf{x}_H^{(t)}; \mathcal{N})$$

Toto zobrazenie priradí dvojici stavu a vstupu v čase $t-1$ nový stav v čase t .

- (5) Funkcia $g : S \times I \rightarrow O$ priradí každému stavu a vstupu výstupný symbol. Podobne ako aj v predchádzajúcom prípade, aj táto funkcia je plne špecifikovaná zobrazením (2.23)

$$\mathbf{x}_O^{(t+1)} = \tilde{F}(\mathbf{x}_I^{(t)} \oplus \mathbf{x}_H^{(t)}; \mathcal{N})$$

- (6) Počiatočný stav s_{ini} sa obvykle vyberie taký, že aktivity skrytých neurónov sú nulové.

Týmto spôsobom, na základe známej neurónovej siete, sme jednoznačne špecifikovali konečnosťavový stroj, ktorý ekvivalentný s danou neurónovou sieťou, každá neurónová sieť môže byť teda reprezentovaná ekvivalentným konečnosťavovým strojom, Q.E.D.

Dôkaz inverznej vety k vete 2.5 (t. j. každý konečnosťavový stroj môže byť reprezentovaný ekvivalentnou neurónovou sieťou) nie je triviálny, prvý kto ju dokázal bol Minsky v r. 1967 v knihe "Computation: Finite and Infinite Machines" [4] pomocou sofistikovaného konštruktívneho dôkazu. Naším cieľom je teda zostrojiť na základe daného konečnosťavového stroja takú neurónovú sieť, ktorá je ekvivalentná s konečnosťavovým strojom.

Veta 2.6 [4]. Každý konečnosťavový stroj s výstupom (Mealyho automat) môže byť reprezentovaný ekvivalentnou rekurentnou neurónovou sieťou.

Príklad 2.2. Ilustračný dôkaz tejto vety vykonáme pomocou jednoduchého príkladu konečnosťavového stroja so stavovým diagramom znázorneným na obr. 2.15. Tento stroj je určený prechodovou a výstupnou funkciou f a g , (pozri tab. 2.4), ktoré môžeme vyjadriť ako Boolove funkcie pomocou dvoch tabuliek:

- (1) Prechodová funkcia $stav_{t+1} = f(stav_t, vstupný\ symbol_t)$:

stav, vstupný znak	prechodová funkcia f
$(s_1, 0) \rightarrow (0, 0)$	$(b) \rightarrow (1)$
$(s_1, 1) \rightarrow (0, 1)$	$(a) \rightarrow (0)$
$(s_2, 0) \rightarrow (1, 0)$	$(a) \rightarrow (0)$
$(s_2, 1) \rightarrow (1, 1)$	$(a) \rightarrow (0)$

(2) Výstupná funkcia $výstupný\ symbol_{t+1} = g(stav_t, vstupný\ symbol_t)$:

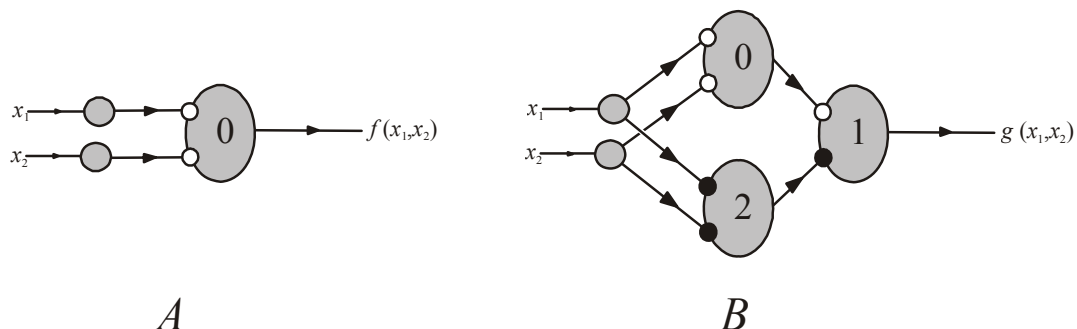
stav, vstupný znak	výstupná funkcia g
$(s_1, 0) \rightarrow (0, 0)$	$(s_2) \rightarrow (1)$
$(s_1, 1) \rightarrow (0, 1)$	$(s_1) \rightarrow (0)$
$(s_2, 0) \rightarrow (1, 0)$	$(s_1) \rightarrow (0)$
$(s_2, 1) \rightarrow (1, 1)$	$(s_2) \rightarrow (1)$

To znamená, že obe funkcie f a g sú špecifikované ako Boolove funkcie

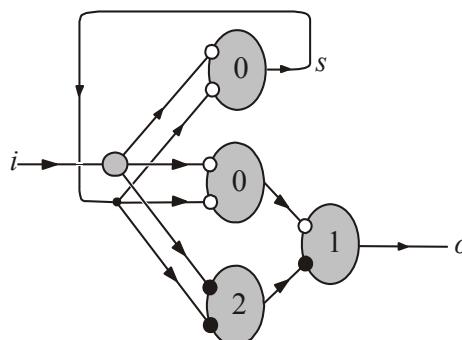
$$f(x_1, x_2) = \neg x_1 \wedge \neg x_2$$

$$g(x_1, x_2) = (\neg x_1 \wedge \neg x_2) \vee (x_1 \wedge x_2)$$

Reprezentácia týchto funkcií pomocou logických neurónov má tvar (pozri kapitolu 2.1) je znázornená na obr. 2.17. Substitúciou týchto neurálnych reprezentácií do schémy z obr. 2.16 dostaneme rekurentnú neurónovú sieť (pozri obr. 2.18), ktorá je ekvivalentná s konečnosťavým strojom so stavovým diagramom znázorneným na obr. 2.15.



Obrázok 2.17. Znázornenie Boolových funkcií f a g z príkladu 2.1, ktoré špecifikujú binárnu reprezentáciu prechodovej resp. výstupnej funkcie.



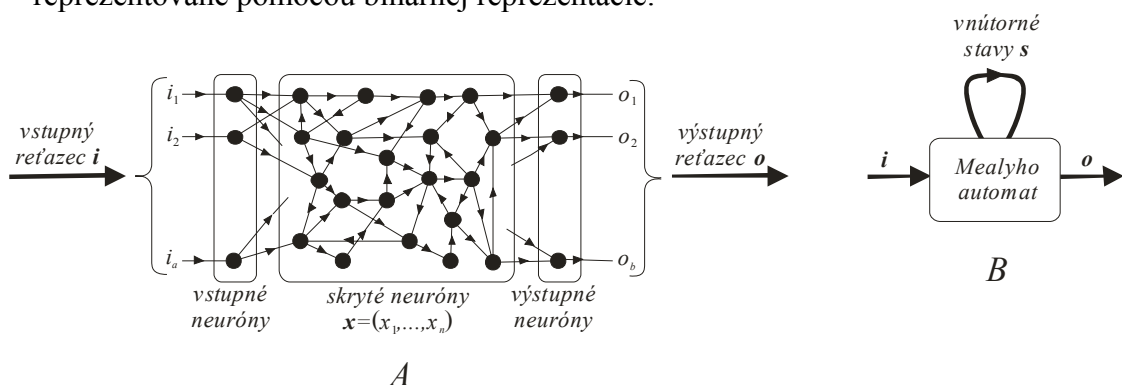
Obrázok 2.18. Rekurentná neurónová sieť reprezentujúca konečnosťavý stroj z obr. 2.15. Táto sieť vznikla substitúciou neurónových reprezentácií Boolových funkcií f a g z obr. 2.17 do diagramu na obr. 2.16, tento diagram reprezentuje všeobecný konečnosťavý stroj so zadanými funkciami f a g .

Poznamenajme, že tento jednoduchý príklad nám dostatočne podrobne reprezentuje všeobecný postup konštrukcie rekurentnej neurónovej siete, ktorá je ekvivalentná s daným konečnosťavým strojom, ktorý je špecifikovaný funkciami f a g . V prvom kroku zostrojíme neurálnu reprezentáciu týchto funkcií metódou špecifikovanou pre Boolove funkcie v kapitole 2.1 a v druhom kroku tieto neurálne reprezentácie substituujeme do všeobecného diagramu z obr. 2.16 konečnosťavého stroja špecifikovaného funkciami f a g . Týmto príkladom sme zavŕšili dôkaz vety 2.6, ukázali sme konštruktívnu metódu ako zostrojiť rekurentnú neurónovú sieť, ktorá je ekvivalentná s požadovaným konečnosťavým strojom.

Záverom je potrebné zdôrazniť, že neurónové siete zložené z logických neurónov sú silným výpočtovým prostriedkom: (1) neurónové siete reprezentované acyklickým grafom sú univerzálnym aproximátorom Boolových funkcií a (2) medzi konečnosťavým strojom a neurónovou sieťou existuje princíp ekvivalencie, ľubovoľný konečný automat môže byť simulovaný rekurentnou neurónovou sieťou, a podobne, ľubovoľná neurónová sieť môže byť simulovaná konečnosťavým strojom. Tieto dve vlastnosti sú prototypom výpočtovej mohutnosti pre iné typy neurónových sietí používajúce neuróny so zložitejšou aktivačnou funkciou. Podstatným ohraničením neurónových sietí je fakt, že spoje medzi neurónmi a ich rozdelenie na excitačné a inhibičné a tiež aj prahové faktory sú určené voľbou architektúry siete. Ináč povedané, siete s logickými neurónmi nie sú schopné adaptačného procesu (učenia), Boolova funkcia (alebo Boolove funkcie v prípade viacerých výstupných neurónov) určená sieťou je nemenná v celom priebehu používania siete.

Vety 2.5 a 2.6 umožňujú študovať klasický problém konečnosťavosti (neurónové siete), aký je vzťah medzi **subsymbólickou reprezentáciou** (neurálnou, ktorá je reprezentovaná obrazcami zloženými z aktivít neurónov) a **symbólickou reprezentáciou** (na ktorej je založená klasická umelá inteligencia):

- (1) Pomocou vety 2.5 každá subsymbólická neurónová sieť môže byť pretransformovaná na symbólický konečnosťavý stroj (Mealyho automat, pozri obr. 2.19), pričom symboly môžu byť tvorené pomocou prirodzených čísel, ktoré sú priradené binárnym vektorom vstupu, výstupu a aktivitám skrytých neurónov.
- (2) Pomocou vety 2.6 každý symbólický konečnosťavý stroj (Mealyho automat) môže byť pretransformovaný na subsymbólickú neurónovú sieť, ktorá je ekvivalentná s daným konečnosťavým strojom. Symboly z reprezentácie konečnosťavého stroja sú reprezentované pomocou binárnej reprezentácie.



Obrázok 2.19. Diagram A znázorňuje neurónovú sieť, ktorá pomocou vrstvy skrytých neurónov transformuje vstupné binárne signály na výstupné binárne signály. Stav skrytých neurónov je reprezentovaný binárnym vektorom x , ktorý obsahuje aktivity jednotlivých skrytých neurónov. Poznamenajme, že vrstva skrytých neurónov je reprezentovaná cyklickým grafom, t. j. má rekurentný charakter, ktorý umožňuje študovať aj procesy vyžadujúce pamäť. Tento diagram tvorí základom *subsymbólickej reprezentácie*, ktorá je synonymom pre konečnosťavý (neurálny) prístup k štúdiu procesov (kognitívnych) prebiehajúcich v mozgu. Diagram B reprezentuje *symbólický prístup* k štúdiu procesov v mozgu, ktorý je „prepisom“ konečnosťavého schémy znázornenej na diagrame A pomocou konečnosťavého stroja nazývanej Mealyho automat. Vstupné i a výstupné

σ veličiny a vnútorné stavy s sú v tomto prístupe reprezentované symbolmi, ktoré v najjednoduchšom prístupe môžu byť reprezentované prirodzenými číslami, ktoré sú priradené binárnym vektorom i , σ , resp. x , z diagramu A.

Naznačený postup pre prechod od subsymbolickej k symbolickej reprezentácii a naopak pomocou neurónových sietí s logickými neurónmi a konečnostavových strojov, môže slúžiť štúdiu vlastností týchto modelov a prechodu medzi nimi.

2.3 Konštrukcia binárnych obvodov

Logické neuróny môžu byť použité na konštrukciu jednoduchých binárnych obvodov [4], ich kombináciou môžeme zostrojiť už zložité zariadenia, akými sú napr. rôzne sčítačky, násobičky a podobne.

2.2.1 Časový posunovač

Časový posunovač je taký binárny obvod, ktorý posunie tok signálu o niekoľko vopred zvolených jednotiek. Študujme jednoduchý obvod znázornený na obr. 2.12, diagram A, ktorý realizuje posun o jednotku (jednotkový posunovač). Aktivitu neurónu v čase t môžeme vyjadriť takto

$$y^{(t)} = t(x^{(t-1)} - 1) \quad (2.27)$$

Vstupný signál v čase t sa transformuje na výstupný signál v čase $t+1$

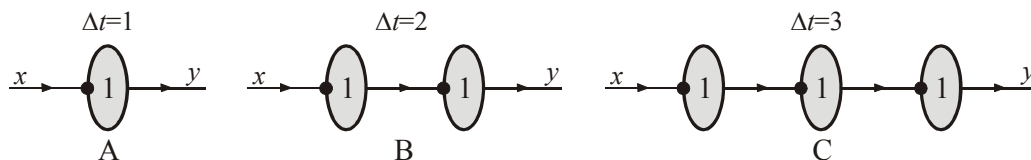
$$\begin{aligned} x^{(t)} = 0 &\Rightarrow y^{(t+1)} = t(-1) = 0 \\ x^{(t)} = 1 &\Rightarrow y^{(t+1)} = t(0) = 1 \end{aligned} \quad (2.28)$$

Nech sekvencie (dĺžky 10) vstupných signálov má tvar (0011000111), potom výstupný symbol má tvar sekvencie (dĺžky 11) (#0011000111), kde prvý znak # je prázdny znak, pretože v počiatočnom čase $t = 1$ nevieme špecifikovať výstupný signál, pozri Tab. 2.4

Tabuľka 2.4. Vstupná a výstupná sekvencia pre jednotkový časový posunovač

čas	1	2	3	4	5	6	7	8	9	10	11	12
vstup x	0	0	1	1	0	0	0	1	1	1
výstup y	#	0	0	1	1	0	0	0	1	1	1	..

Sériovým zapojením jednotkových posunovačov dostaneme posunovače o dve, tri, .. časové jednotky, pozri diagramy B a C na obr. 2.12.

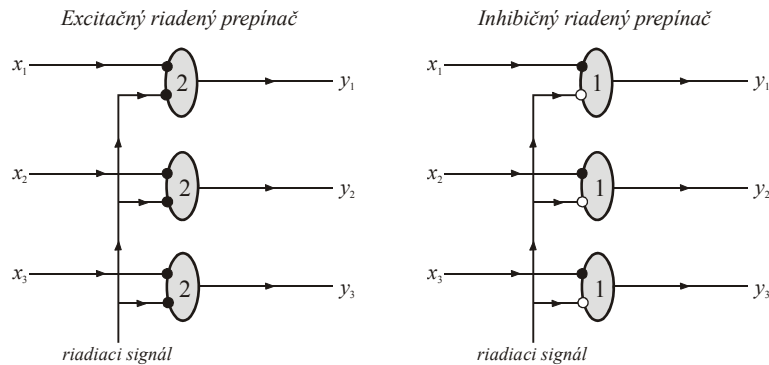


Obrázok 2.12. Neurónové siete realizujúce posun vstupného signálu o 1, 2 a 3 jednotky.

2.3.2 Riadený prepínač

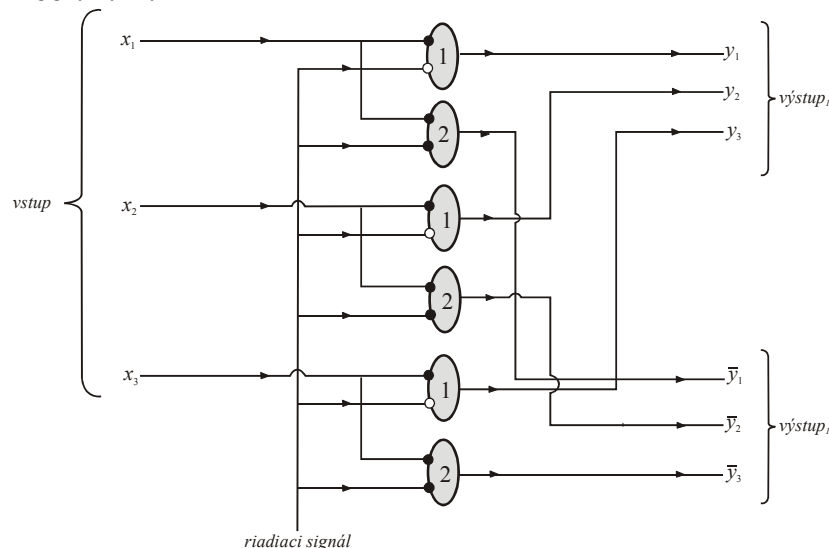
Riadený prepínač pomocou riadiaceho signálu zapína alebo vypína tok signálu. Tak ak v excitačnom riadenom prepínači je riadiaci signál jednotkový, potom prepínač prepúšťa

signál na výstup prepínača. V alternatívnom prípade, pre inhibične riadený prepínač, nulový riadiaci signál spôsobuje prenos signálu na výstup. Tieto dva prepínače sú znázornené na obr. 2.13.



Obrázok 2.13. Znážornenie neurónových sietí, ktoré sú excitačný resp. inhibičný riadiaci prepínač.

Kombináciou excitačného a inhibičného prepínača zostrojíme riadené rozdeľovacie zariadenie, pozri obr. 2.14.



Obrázok 2.14. Neurónová sieť riadeného rozdeľovacieho zariadenia. Ak riadiaci signál je jednotkový (nulový), potom vstup je riadený na 1. výstup (2.výstup). Obrázne môžeme povedať, že riadiaci signál pôsobí ako „prepínač“ do 1. alebo 2. polohy.

2.3.3. Sumátor dvoch binárnych čísel

Sumátor binárnych čísel rieši úlohu sčítania dvoch binárnych čísel – reťazcov rovnakej dĺžky

$$\begin{array}{r} 110011 \\ 101010 \\ \hline 1011101 \end{array}$$

Realizácia tohto problému sa môže uskutočniť úplne analogický procesu sčítania dvoch dekadických čísel, t. j. musíme používať techniku prenosu „zostatku“ (angl. *carry*) do nasledujúcej polohy. Tento princíp si vysvetlíme na všeobecnom príklade

$$\frac{\alpha_n \dots \alpha_2 \alpha_1}{\beta_n \dots \beta_2 \beta_1} \quad (2.29)$$

$$\gamma_{n+1} \gamma_n \dots \gamma_2 \gamma_1$$

sčítania dvoch binárnych čísel $\alpha_n \dots \alpha_2 \alpha_1$ a $\beta_n \dots \beta_2 \beta_1$ dĺžky n . Zostatok δ_i je rekurentne definovaný takto:

$$\delta_1 = 0 \quad (2.30a)$$

$$\delta_2 = \begin{cases} 1 & (\text{ak } \alpha_1 + \beta_1 = 2) \\ 0 & (\text{ináč}) \end{cases} \quad (2.30b)$$

$$\delta_{i+1} = \begin{cases} 1 & (\text{ak } \alpha_i + \beta_i + \delta_i \geq 2) \\ 0 & (\text{ak } \alpha_i + \beta_i + \delta_i \leq 1) \end{cases} \quad (i = 2, 3, \dots, n) \quad (2.30c)$$

kde (2.20a-b) je inicializácia postupného výpočtu zostatkov. Pre ilustračný príklad zo začiatku tejto podkapitoly zostatky sú špecifikované binárnym reťazcom dĺžky 7 tvaru $\delta = 1000100$. Schému (2.19) prepíšeme do nového tvaru

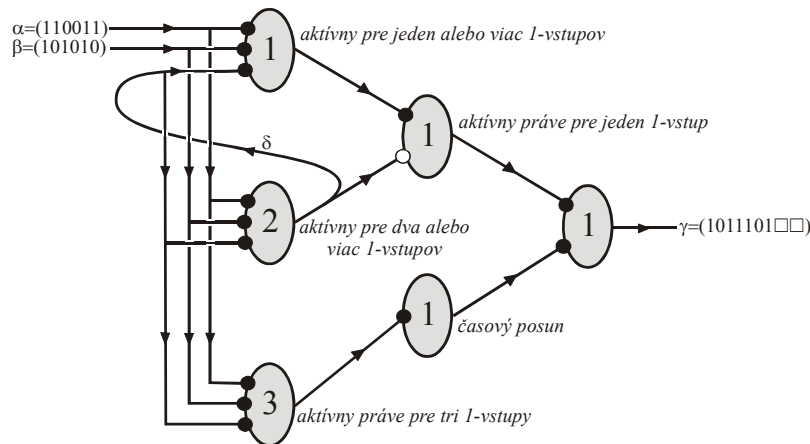
$$\frac{\alpha_n \dots \alpha_2 \alpha_1}{\beta_n \dots \beta_2 \beta_1} \quad (2.31)$$

$$\frac{\delta_{n+1} \delta_n \dots \delta_2 \delta_1}{\gamma_{n+1} \gamma_n \dots \gamma_2 \gamma_1}$$

Potom výsledok $\gamma = \gamma_{n+1} \dots \gamma_2 \gamma_1$ je určený vzťahmi

$$\gamma_i = \begin{cases} 1 & (\text{ak } \alpha_i + \beta_i + \delta_i = 1 \vee 3) \\ 0 & (\text{ak } \alpha_i + \beta_i + \delta_i = 0 \vee 2) \end{cases} \quad (i = 1, 2, \dots, n) \quad (2.32a)$$

$$\gamma_{n+1} = \delta_{n+1} \quad (2.32b)$$



Obrázok 2.15. Neurónová sieť sumátora dvoch binárnych čísel, ktorá obsahuje šesť prahových neurónov.

Príklad zo začiatku tejto podkapitoly môže byť prepísaný do tvaru

$$\begin{array}{r} 110011 \\ 101010 \\ 1000100 \\ \hline 1011101 \end{array}$$

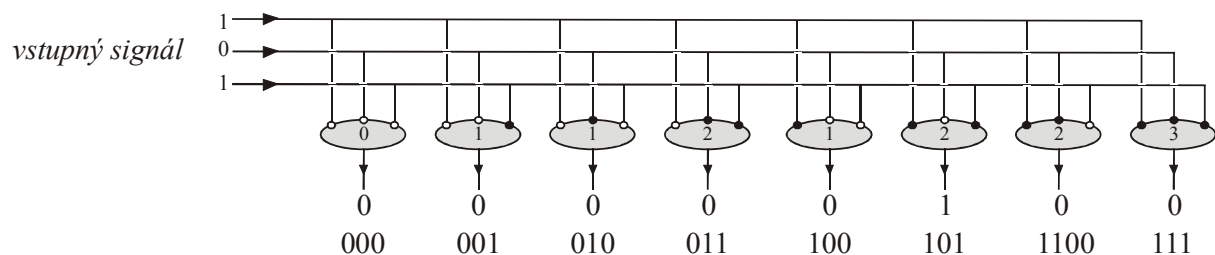
kde výsledný reťazec bol určený formulami (2.32). Jednotlivé kroky z tohto príkladu sú rozpísané takto:

0. krok. $\delta_1 = 0$
1. krok. $\alpha_1 = 1, \beta_1 = 0, \delta_2 = 0, \boxed{\gamma_1 = 1}$
2. krok. $\alpha_2 = 1, \beta_2 = 0, \delta_3 = 1, \boxed{\gamma_2 = 0}$
3. krok. $\alpha_3 = 0, \beta_3 = 0, \delta_4 = 0, \boxed{\gamma_3 = 1}$
4. krok. $\alpha_4 = 0, \beta_4 = 1, \delta_5 = 0, \boxed{\gamma_4 = 1}$
5. krok. $\alpha_5 = 1, \beta_5 = 0, \delta_6 = 0, \boxed{\gamma_5 = 1}$
6. krok. $\alpha_6 = 1, \beta_6 = 1, \delta_7 = 1, \boxed{\gamma_6 = 0}$
7. krok. $\alpha_7 = 0, \beta_7 = 0, \boxed{\gamma_7 = 1}$

Implementácia sumátora je znázornená na obr. 2.15.

2.4.4. Binárny paralelný dekóder

Budeme študovať paralelný dekóder pre tri binárne číslice. Predpokladáme, že informácia do dekódera prichádza prostredníctvom impulzov cez tri paralelne vlákna, pozri obr. 2.16. Dekóder obsahuje osem nezávislých elementov – neurónov, z ktorých každý dekóduje jednu možnú binárnu kombináciu prichádzajúcich signálov.



Obrázok 2.16. Binárny paralelný dekóder, ktorý obsahuje osem prahových neurónov, ktoré dekódujú separátne vstupný signál od kombinácie 000 až po kombináciu 111.

2.4 Záver

Neurónové siete zložené z logických neurónov sú silným výpočtovým prostriedkom: (1) neurónové siete reprezentované acyklickým grafom sú univerzálnym aproximátorom Boolových funkcií a (2) ľubovoľný konečný automat môže byť simulovaný rekurentnou neurónovou sieťou. Vlastnosť (1) môže byť podstatne zosilnená tvrdením, že každá Boolova funkcia môže byť reprezentovaná doprednou 3-vrstvovou neurónovou sieťou. Podstatným ohrozením logických neurónov je, že klasifikujú len Boolove funkcie, ktoré sú lineárne separovateľné. Táto skutočnosť môže byť prekonaná pomocou logických neurónov vyššieho rádu alebo 3-vrstvovými neurónovými sieťami.

Literatúra

- [1] Kolář, J., Štěpánková, O., Chytil, M.: *Logika, algebry a grafy*. SNTL, Praha, 1989.
- [2] Kvasnička, V., Pospíchal, J.: *Matematická logika*, Vydavateľstvo STU, Bratislava, 2006.
- [3] McCulloch, W. S., Pitts, W. H.: A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, **5** (1943), 115-133.
- [4] Minsky, M. L.: *Computation. Finite and Infinite Machines*. Prentice-Hall, Englewood Cliffs, NJ, 1967.
- [5] Minsky, M. and Papert, S.: *Perceptrons. An Introduction to Computational Geometry*. MIT Press, Cambridge, MA, 1969.
- [6] Molnár L., Češka, M., Melichar, B.: *Gramatiky a jazyky*. Bratislava, Alfa, 1987.