

# Šachy, go a umělá inteligence

Martina Šimůnková

Seminář Ko-mix, 25. dubna 2016



# Garry Kasparov vs. Deep Blue

Garry Kasparov vs. Deep Blue

[https://en.wikipedia.org/wiki/Deep\\_Blue\\_versus\\_Garry\\_Kasparov#Summary](https://en.wikipedia.org/wiki/Deep_Blue_versus_Garry_Kasparov#Summary)

Lee Sedol vs. AlphaGo <https://gogameguru.com/>

[lee-sedol-defeats-alphago-masterful-comeback-game-4/](https://gogameguru.com/lee-sedol-defeats-alphago-masterful-comeback-game-4/)

# Hry s úplnou informací

# Hry s úplnou informací

- ▶ Hru popisuje *minimaxový prohledávací strom*. (Ukázka na hře se sirkami.)

# Hry s úplnou informací

- ▶ Hru popisuje *minimaxový prohledávací strom*. (Ukázka na hře se sirkami.)
- ▶ Pro jednoho z hráčů existuje strategie, která mu zajistí vítězství bez ohledu na hru soupeře.  
V případě šachu je možné, že to je strategie pro oba hráče k dosažení remízy.

# Hry s úplnou informací

- ▶ Hru popisuje *minimaxový prohledávací strom*. (Ukázka na hře se sirkami.)
- ▶ Pro jednoho z hráčů existuje strategie, která mu zajistí vítězství bez ohledu na hru soupeře.  
V případě šachu je možné, že to je strategie pro oba hráče k dosažení remízy.
- ▶ Pro nalezení optimální strategie v šachách je třeba projít zhruba  $35^{100} \sim 10^{154}$  pozic.

# Hry s úplnou informací

- ▶ Hru popisuje *minimaxový prohledávací strom*. (Ukázka na hře se sirkami.)
- ▶ Pro jednoho z hráčů existuje strategie, která mu zajistí vítězství bez ohledu na hru soupeře.  
V případě šachu je možné, že to je strategie pro oba hráče k dosažení remízy.
- ▶ Pro nalezení optimální strategie v šachách je třeba projít zhruba  $35^{100} \sim 10^{154}$  pozic.
- ▶ V go zhruba  $150^{200} \sim 10^{435}$  pozic.

# Hry s úplnou informací

- ▶ Hru popisuje *minimaxový prohledávací strom*. (Ukázka na hře se sirkami.)
- ▶ Pro jednoho z hráčů existuje strategie, která mu zajistí vítězství bez ohledu na hru soupeře.  
V případě šachu je možné, že to je strategie pro oba hráče k dosažení remízy.
- ▶ Pro nalezení optimální strategie v šachách je třeba projít zhruba  $35^{100} \sim 10^{154}$  pozic.
- ▶ V go zhruba  $150^{200} \sim 10^{435}$  pozic.
- ▶ Obojí je více než odhad počtu atomů v (pozorovatelném) vesmíru (uvádí se  $10^{78}$  a někdy až  $10^{100}$ ).



# Proč nemůže Deep Blue hrát go

# Proč nemůže Deep Blue hrát go

- ▶ Deep Blue prohledává celou šířku prohledávacího stromu, prořezává pouze jeho hloubku. Nedohrává tedy partie do konce.  
Pozice na listech (seříznutého) stromu ohodnocuje statickou funkcí, která oceňuje figury a jejich postavení na desce.

# Proč nemůže Deep Blue hrát go

- ▶ Deep Blue prohledává celou šířku prohledávacího stromu, prořezává pouze jeho hloubku. Nedohrává tedy partie do konce.  
Pozice na listech (seříznutého) stromu ohodnocuje statickou funkcí, která oceňuje figury a jejich postavení na desce.
- ▶ Omezení šířky prohledávacího stromu řeší prořezem pozic, o kterých „ví“ (řídí se výše zmíněnou ohodnocovací funkcí), že nejsou nejlepší. Používá heuristiky, které toto prořezávání zefektivňují.

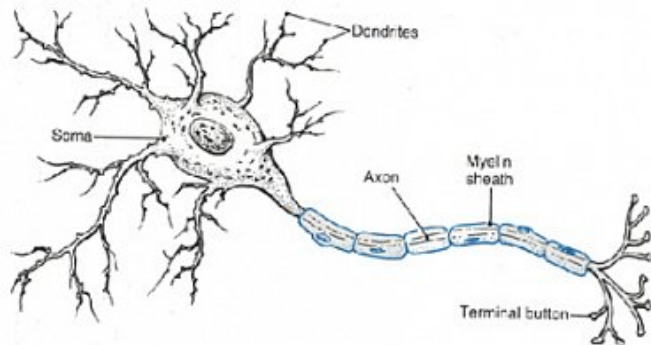
# Proč nemůže Deep Blue hrát go

- ▶ Deep Blue prohledává celou šířku prohledávacího stromu, prořezává pouze jeho hloubku. Nedohrává tedy partie do konce.  
Pozice na listech (seříznutého) stromu ohodnocuje statickou funkcí, která oceňuje figury a jejich postavení na desce.
- ▶ Omezení šířky prohledávacího stromu řeší prořezem pozic, o kterých „ví“ (řídí se výše zmíněnou ohodnocovací funkcí), že nejsou nejlepší. Používá heuristiky, které toto prořezávání zefektivňují.
- ▶ V go toto nejde: srovnej zhruba 35 možností tahů v průměrné šachové pozici a 150 tahů v pozici hry go. V go je nutné uvažovat jen některé „perspektivní tahy“.

# Proč nemůže Deep Blue hrát go

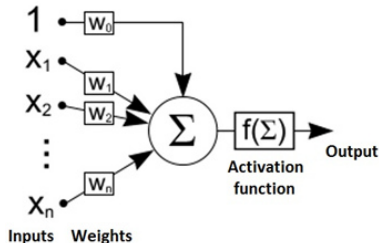
- ▶ Deep Blue prohledává celou šířku prohledávacího stromu, prořezává pouze jeho hloubku. Nedohrává tedy partie do konce.  
Pozice na listech (seříznutého) stromu ohodnocuje statickou funkcí, která oceňuje figury a jejich postavení na desce.
- ▶ Omezení šířky prohledávacího stromu řeší prořezem pozic, o kterých „ví“ (řídí se výše zmíněnou ohodnocovací funkcí), že nejsou nejlepší. Používá heuristiky, které toto prořezávání zefektivňují.
- ▶ V go toto nejde: srovnej zhruba 35 možností tahů v průměrné šachové pozici a 150 tahů v pozici hry go. V go je nutné uvažovat jen některé „perspektivní tahy“.
- ▶ Podobně: není jednoduché sestavit statickou ohodnocovací funkci.

## Neuronová buňka



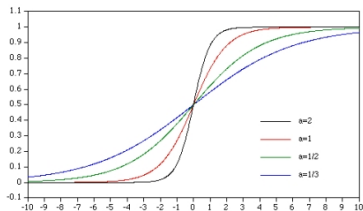
<http://www.root.cz/clanky/biologicke-algoritmy-4-neuronove-site>

## Umělý neuron



<http://www.root.cz/clanky/biologicke-algoritmy-4-neuronove-site>

## Aktivační funkce



[www.beigebag.com/case\\_sigmoid\\_half.htm](http://www.beigebag.com/case_sigmoid_half.htm)

## Potenciál umělého neuronu

$$P = \sum_{k=1}^n w_k x_k$$

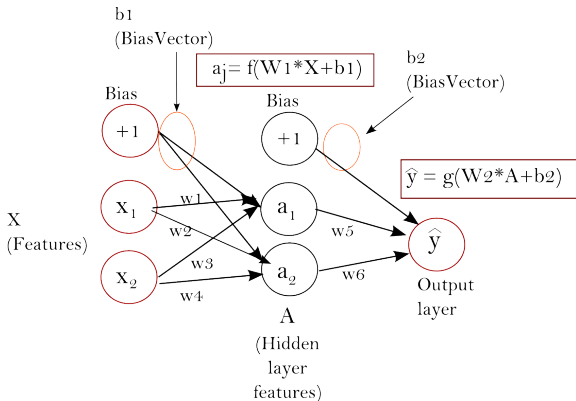
## Příklad aktivační funkce

$$f(\Sigma) = \frac{1}{1 + \exp(-a\Sigma)}$$

$$f'(\Sigma) = a f(\Sigma) (1 - f(\Sigma))$$

$\Sigma$  je rozdíl potenciálu  $P$  a prahové hodnoty  $w_0$ :  $\Sigma = P - w_0$ .

# Vícevrstvé síť

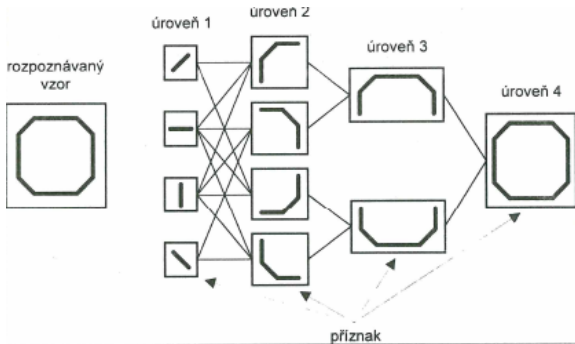


<http://i.stack.imgur.com/5SkqL.png>



# Neocognitron

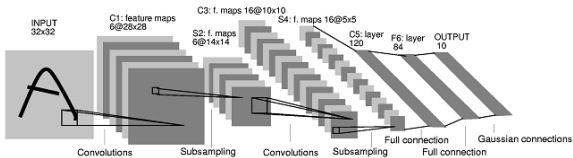
Fukushima Funihoko, 80. léta, rozpoznávání psaných číslic, vzor pro konvoluční neuronové síť.



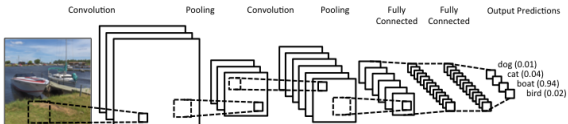
Šnorek, M.: Neuronové sítě a neuropočítače, FEL ČVUT

<https://www.youtube.com/watch?v=Qil4kmvm2Sw>

Síť je navržena tak, aby rozpoznala „vybočené“ tvary. Učí se bez učitele: dostává vzory číslic a učí se je vzájemně rozpoznávat.



<http://i.stack.imgur.com/oUwMk.png>



[https://www.clarifai.com/static/img\\_ours/cnn.png](https://www.clarifai.com/static/img_ours/cnn.png)

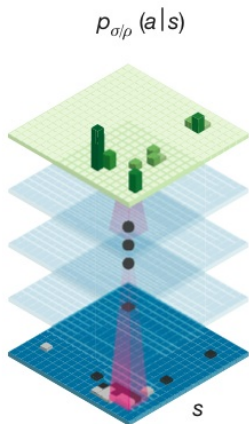
Firma **Deep Mind Technology** byla založena v roce 2011 v Londýně a v roce 2014 koupena firmou Google (akvizicí) a přejmenována na **Google Deep Mind**.

V tomtéž roce začali vyvíjet AlphaGo: začali neuronovou sítí, kterou učili odhadnout příští tah v pozici.

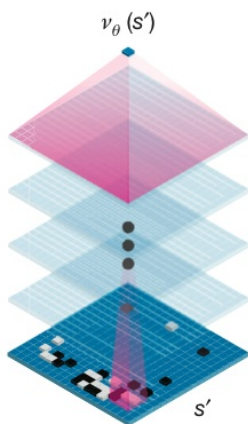
V první fázi použili *supervised learning* s 30 milióny pozic z KGS go serveru, většinou z partií silných amatérských hráčů. Takto vytvořili *policy network*, která pozici přiřadí pravděpodobnostní rozdělení dalšího tahu. Skutečně zahráný tah v pozici měl v tomto rozdělení v průměru 57%.

Dvě varianty *policy network* pak nechali hrát partie proti sobě a tím je jednak dále učili a zároveň s tím vytvářeli *value network* ohodnocující pozici pravděpodobností výhry. Použitá metoda se nazývá *reinforcement learning*.

Policy network



Value network



# Učení s učitelem, backpropagation

Síť předkládáme data z tréninkové množiny a ona si po každém vzorku přenastaví váhy svých neuronů. Snaží se minimalizovat součet čtverců odchylek svých výstupních hodnot od předepsaných hodnot.

$$M = \frac{1}{2} \sum_i (y_i - d_i)^2$$

Metoda největšího spádu:  $\Delta w_{kl, \text{new}} = -\eta_1 \frac{\partial M}{\partial w_{kl}} + \eta_2 \Delta w_{kl, \text{old}}$ .

Výstupní vrstva

$$\frac{\partial M}{\partial w_{ki}} = \frac{\partial M}{\partial y_i} \frac{\partial y_i}{\partial w_{ki}}$$

$$\frac{\partial M}{\partial y_i} = y_i - d_i$$

$$\frac{\partial y_i}{\partial w_{ki}} = a y_i (1 - y_i) x_k$$

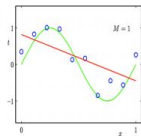
Skrytá vrstva: vstup výstupní vrstvy  $x_k$  je výstupem skryté vrstvy.

$$\frac{\partial M}{\partial w_{lk}} = \sum_i \frac{\partial M}{\partial y_i} \frac{\partial y_i}{\partial x_k} \frac{\partial x_k}{\partial w_{lk}}$$

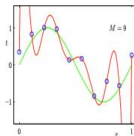
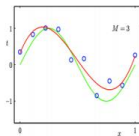
Vzory obsahují šum: je třeba dát pozor na „přeučení“ sítě.

## Under- and Over-fitting examples

Regression:

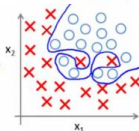
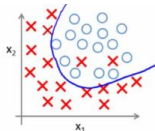
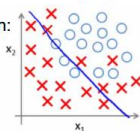


predictor too inflexible:  
cannot capture pattern



predictor too flexible:  
fits noise in the data

Classification:



[https://www.researchgate.net/post/How\\_to\\_Avoid\\_Overfitting](https://www.researchgate.net/post/How_to_Avoid_Overfitting)

# Co umějí neuronové sítě

Ukážeme si, jak hraje neuronová síť ping pong. Vstup neuronové sítě je to, co uvidíte na obrazovce; neuronová síť se skládá z neuronů, vazby mezi neurony jsou reprezentovány parametry, které neuronová síť metodou pokus omyl během učení se přenastavuje podobně jako když se malé dítě učí chodit. Přitom se snaží maximalizovat skóre. Topologii neuronové sítě navrhli vývojáři, neuložili do ní žádné informace o konkrétní hře.

Výsledek vypadá takto:

<https://www.youtube.com/watch?v=V1eYniJ0Rnk>

Pattern recognition – rozpoznávání tvarů, například spamu mezi e-maily, čísel na obrázku, dalšího tahu při hře go, pravděpodobnosti vítězství z pozice hry go.

Chytrost neuronové sítě závisí na inteligenci jejích tvůrců.  
Průměrná pravděpodobnost skutečného tahu v partii byla u  
Google deep mind 57%, u ostatních týmů nejvíce 44%.

Rutinní činnosti vykonává neuronová síť bezchybněji a rychleji.

Vložené principy umí aplikovat v analogických situacích lépe  
než člověk. Jako příklad může sloužit tah 37 ve druhé partii  
AlphaGo a Lee Sedola.

[https://gogameguru.com/  
alphago-races-ahead-2-0-lee-sedol/](https://gogameguru.com/alphago-races-ahead-2-0-lee-sedol/)

Zajímavé úvahy o možné budoucí nezaměstnanosti a o  
pohledu na tvořivost z perspektivy umělé inteligence:  
<https://hdbennett.wordpress.com>



## K dalšímu studiu

Mařík, Štěpánková, Lažanský a kol.: Umělá inteligence,  
Academia

Šnorek, M.: Neuronové sítě a neuropočítače, vydavatelství  
ČVUT

Karel Ha, přednáška na jarní škole kombinatoriky o článku v  
časopise Nature „Mastering the game of Go with deep neural  
networks and tree search“: <http://www.slideshare.net/KarelHa1/>

mastering-the-game-of-go-with-deep-neural-networks-and-tree-search-presentation