

Modelování myšlení

Radek Pelánek

- dosavadní modely měly **statickou** strukturu (pravidla)
- komplexní systémy se **mění**: vývoj, adaptace, učení, ...
- vztah k horizontu modelu
 - vlk: reintrodukce, lov ve smečce
 - epidemie: zásah bioterorismus, antibiotika a resistance

system	rychlost zmen
nervovy system	sekundy az hodiny
imunitni system	hodiny az dny
firma	mesice az roky
zivoctisny druh	dny az století
ekosystem	roky az milenia

Smysl modelů adaptace

- ① pochopení přírody
 - jak funguje myšlení, evoluce, ...
- ② lepší modely komplexních systémů
 - obohacení modelů s agenty
- ③ řešení náročných algoritmických problémů
 - modely jako výpočetní mechanismy

Myšlení, modelování, počítače

(zjednodušeně)

kognitivní modelování

- cíl: zachytit, jak funguje mysl
- důraz na jednoduchost, věrnost, principy
- není úplně zásadní „výkon“

umělá inteligence

- cíl: reprodukovat inteligentní chování
- důraz na „výkon“
- biologické principy jen jako inspirace

Kognitivní modelování

symbolické modelování

- „high-level“
- uvažování na úrovni pravidel, metafora počítače
- manipulace symbolů

konekcionalistické modelování

- „low-level“
- uvažování na úrovni neuronů
- snaha o biologickou věrnost

Kdyby byl mozek tak jednoduchý, že bychom mu mohli rozumět, byli bychom tak jednoduší, že bychom nemohli.
(Lyaal Wattson)

Symbolické modelování

- *computational theory of mind*
- myšlení jako manipulace symbolů
- produkční pravidla: IF X THEN Y
- učení: vytváření nových pravidel (např. spojováním, vyhledáváním)
- kde se vezmou symboly?

Produkční systém

- pravidla typu
IF podmínka THEN akce
- podmínky přichází z prostředí nebo jako následky jiných pravidel
- akce mohou vyvolávat další pravidla nebo měnit prostředí

Příklad:

IF chci čaj THEN uvařit čaj

IF uvařit čaj a nemám horkou vodu THEN dát vodu do konvice

Produkční systémy: kolize

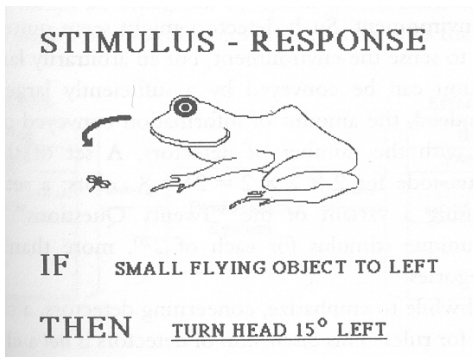
- pravidla mohou být **kolizní** (nedeterministický systém)
- pro danou situaci výběr aplikovatelných, rozhodnutí kolize
- priority, náhoda, ...

Příklad:

IF chci čaj THEN uvařit čaj

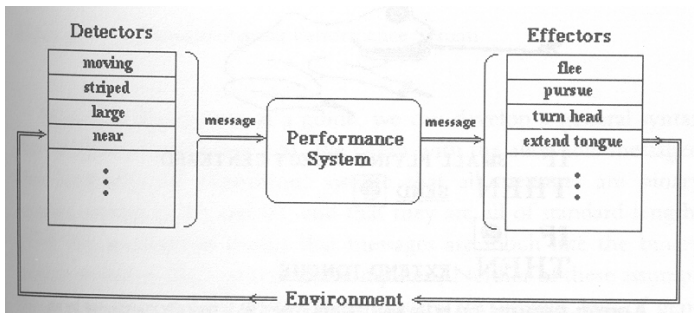
IF chci čaj THEN zajít do hospody a objednat čaj

Příklad žába: základní princip



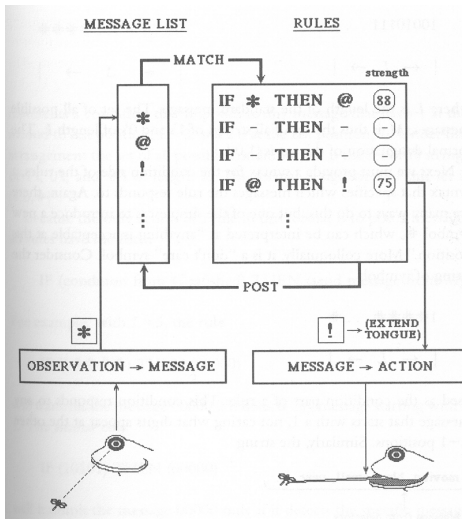
(John Holland, Hidden Order: How Adaptation Builds Complexity)

Příklad žába: kontext

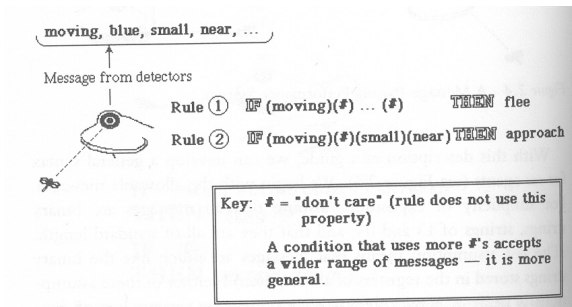


(John Holland, *Hidden Order: How Adaptation Builds Complexity*)

Příklad žába: produkční systém

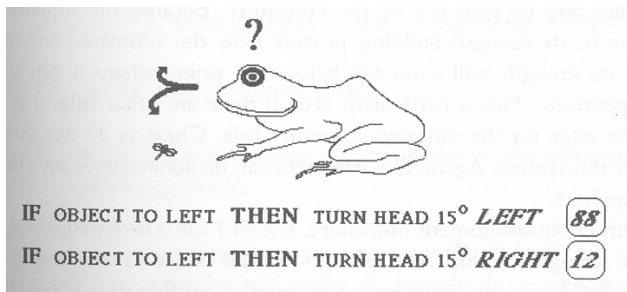


Příklad žába: výběr použitelných pravidel



(John Holland, *Hidden Order: How Adaptation Builds Complexity*)

Příklad žába: řešení kolize



(John Holland, *Hidden Order: How Adaptation Builds Complexity*)

Klasifikační systémy

(Learning classifier systems)

Konkrétní příklad architektury:

- 1 produkční systém
- 2 ohodnocení pravidel (bucket brigade)
- 3 objevování pravidel pomocí genetického algoritmu

Ohodnocení pravidel: bucket brigade

- pravidla konfliktní
- pravidlo má **kapitál**
- výběr pravidla podle aktuálního kapitálu
- pravidlo **platí** za to, že může být použito
- úspěšná aplikace $\Rightarrow$ zisk **odměny** od prostředí
- část odměny **redistribuována** pravidlům, které umožnily použití pravidla, které vedlo k odměně

Vývoj nových pravidel

- pravidla jsou reprezentována jako binární řetězce
- přímočaré použití genetického algoritmu
- zdatnost = kapitál

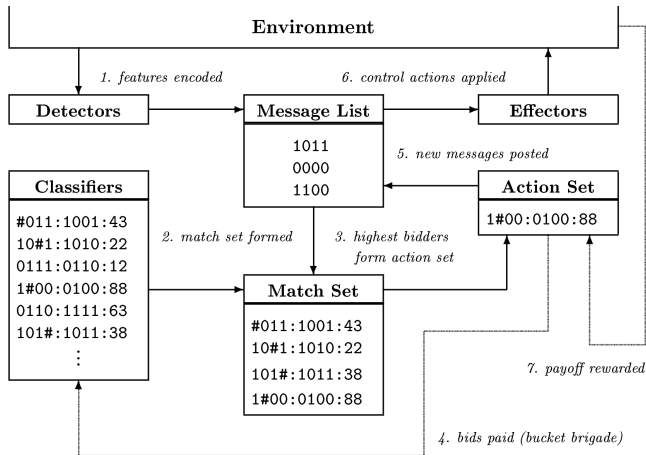


Figure 21.2 A classifier system interacting with its environment

(G. W. Flake, *The Computational Beauty of Nature*)

Příklad Woods

- zvíře na čtverečkovaném poli
 - '*' zvíře
 - '.' volné pole
 - 'O' skála
 - 'F' jídlo
- zvíře vidí okolních 8 pozic, tj. detektory lze zakódovat pomocí 16 bitového řetězce ($.$ = 00, O = 10, F =11)
- cílem je najít co nejrychleji jídlo

Příklad Woods

```

.....
.00F..00F..00F..00F..00F..00F..00F..00F..00F..00F..
.000..000..000..000..000..000..000..000..000..000..
.000..000..000..000..000..000..000..000..000..000..
.....
.....
.00F..00F..00F..00F..00F..00F..00F..00F..00F..00F..
.000..000..000..000*.000..000..000..000..000..000..
.000..000..000..000..000..000..000..000..000..000..
.....
.....
.00F..00F..00F..00F..00F..00F..00F..00F..00F..00F..
.000..000..000..000..000..000..000..000..000..000..
.000..000..000..000..000..000..000..000..000..000..
.....

```

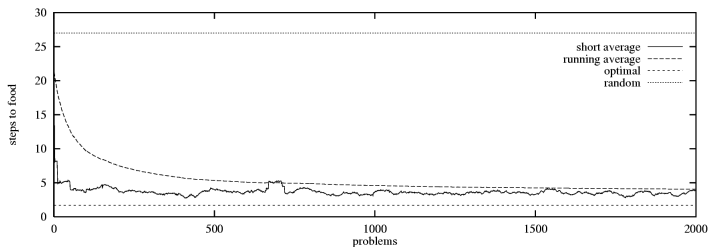
(G. W. Flake, *The Computational Beauty of Nature*)

Příklad Woods

Realizace experimentu (učení):

- náhodně umístit zvíře
- nechat běžat dokud nenajde jídlo
- a pak znova od začátku

Příklad Woods: výsledky



(G. W. Flake, *The Computational Beauty of Nature*)

Příklad Woods: složitější prostředí

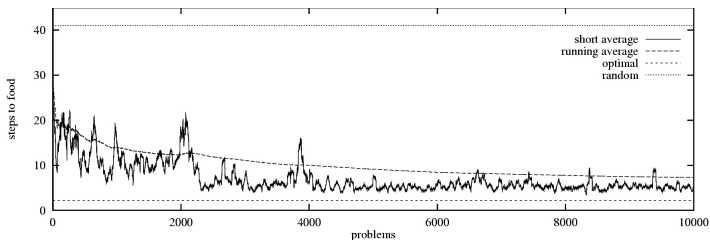
```

.....0......00.....0.....
.OFO.....F.....F.....0.....F.....FO.....
.....0.....00.....F.....
.....0.....0.....F.....0.....
...F.....OFO.....OFO.....F.....00.....F.....
...00.....0.....0.....
.....00.....0.....00.....0.....
.OFO.....F.....OF.....F.....OFO.....F.....
.....0.....
...00.....0.....0.....00.....0.....
...F.....F.....0.....FO.....F.....OF.....OFO.....
.....0.....OF.....
...0.....0.....0.....0.....0.....0.....
...F.....F.....F.....FO.....F.....OF.....
...0.....00.....0.....0.....0.....
.....0.....0.....
...F.....OFO.....F.....F.....F.....OF.....
...00.....0.....00.....0.....0.....

```

(G. W. Flake, *The Computational Beauty of Nature*)

Příklad Woods: výsledky II



(G. W. Flake, *The Computational Beauty of Nature*)

Deduktivní a induktivní uvažování

- **deduktivní** uvažování
 - z daných předpokladů se dle zákonů logiky odvozují logicky platné závěry (konkretizace)
 - modelování není velký problém (deduktivní uvažování lze vyjádřit formálně), hojně používáno např. v ekonomii
 - ovšem neodpovídá realitě – lidé používají deduktivní uvažování jen v omezené míře
- **induktivní** uvažování
 - zevšeobecňování, odvozování obecných zákonů z konkrétních příkladů, odhadování vývoje, ...
 - často a úspěšně používáno lidmi
 - jak modelovat?

El Farol Bar

El Farol je bar v Santa Fe, kde hrají ve čtvrtek večer irskou hudbu

- v okolí baru žije 100 lidí
- každý by rád zašel do baru, ale když je tam moc narváno, tak to za nic nestojí:
 - v baru méně jak 60 lidí $\Rightarrow$ lepší být v baru než doma
 - v baru více jak 60 lidí $\Rightarrow$ lepší být v doma než v baru
- všichni se **rozhodují současně**, žádná domluva

Deduktivní model rozhodování

deduktivní řešení:

- „hodit si kostkou“, tj. rozhodovat se náhodnostně podle vypočítané pravděpodobnosti
- tak se ale lidé nerozhodují...

Induktivní model rozhodování

Každý agent má několik **hypotéz** pro předpovídání počtu návštěvníků. Příklad dosavadní sekvence a hypotéz:

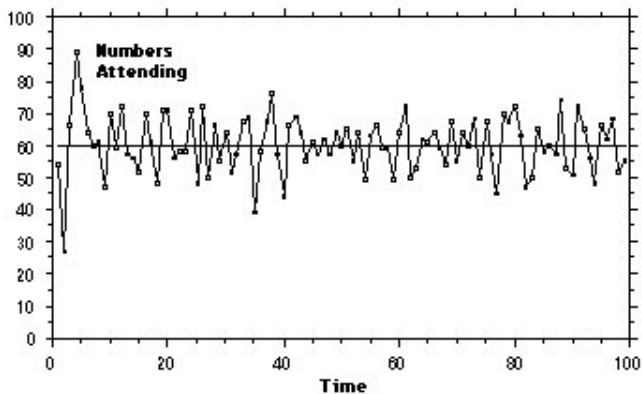
44, 78, 56, 15, 23, 67, 84, 34, 45, 76, 40, 56, 22, 35

- stejně jako minulý týden [35]
- zaokrouhlený průměr za minulé 4 týdny [49]
- stejně jako před dvěma týdny [22]
- zrcadlový obraz okolo 50 z minulého týdne [65]
- 42 [42]

Induktivní model rozhodování

- pro každou hypotézu si agent pamatuje její **úspěšnost**
- vždy se rozhoduje podle aktuální nejúspěšnější hypotézy (případně průměruje několik nejúspěšnějších)
- hypotézy vygenerovány „ručně“, každému agentu náhodně přiřazeno několik z nich

Simulace



Výsledky a rozšíření

- počet návštěvníků osciluje okolo hodnoty 60
- induktivně uvažující agenti tedy řeší problém poměrně dobře
- možná rozšíření: tvorba nových hypotéz (genetický algoritmus), komunikace mezi agenty, ...

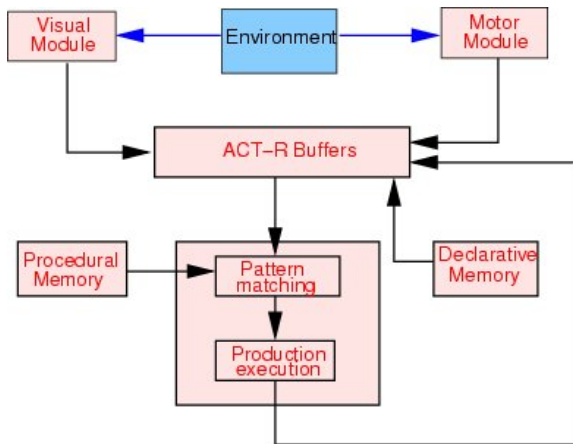
podobným způsobem můžeme udělat model trhu (induktivně uvažující obchodníci)

ACT-R (Adaptive Control of Thought–Rational)

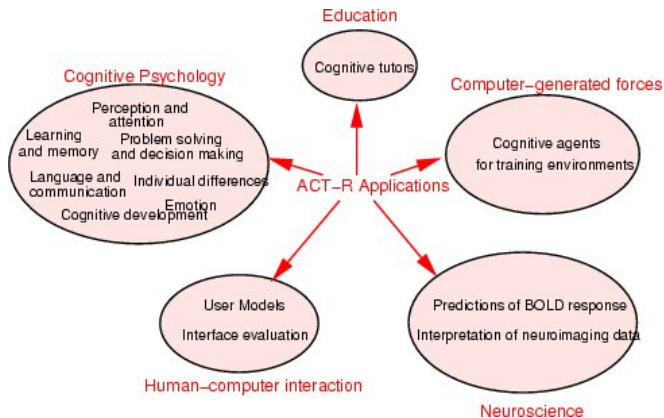
- kognitivní architektura
- založeno na produkčních systémech
- rozsáhlé, mnoho aspektů vnímání, myšlení, paměti
- mnoho aplikací

<http://act-r.psy.cmu.edu/>

ACT-R (Adaptive Control of Thought–Rational)



ACT-R: aplikace



Aplikace ve výuce

- inteligentní výukové systémy
- produkční systém pro určitou oblast (např. aritmetika)
- model „sleduje“ řešení studenta $\Rightarrow$ ohodnocení schopností studenta, identifikace silných/slabých míst
- pravidla pro typické chyby $\Rightarrow$ možnost cílené zpětné vazby

příklad: sčítání zlomků

Neuronové sítě

- velmi aktuální téma v oblasti strojového učení
- zde pouze stručné souvislosti s modelováním
- PV021 Neural Networks

Neuronové sítě

- konekcionistické modelování
- model založený na zjednodušené imitaci fungování mozku
- neurony, spojení
- **induktivní učení** prostřednictvím trénovacích dat
- „znalosti“ jsou v síti uloženy jako váhy **spojení** mezi neurony

Neuron

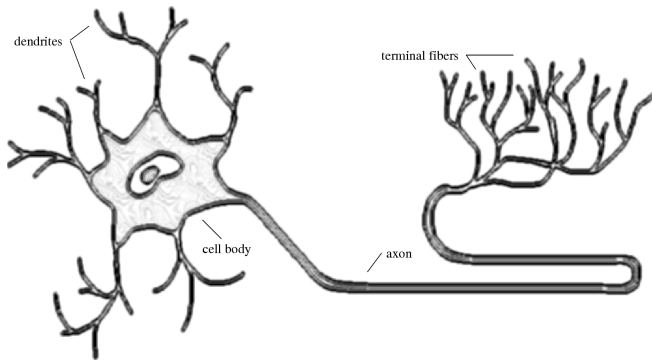
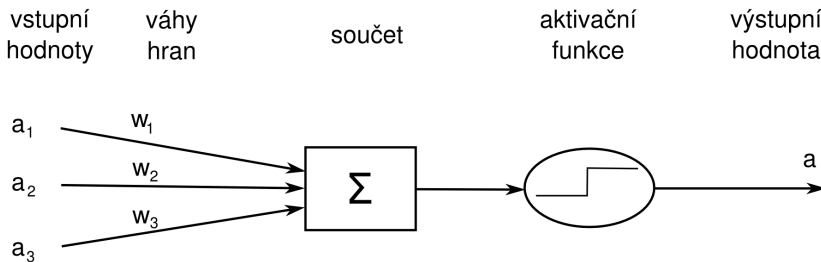


Figure 18.1 A “typical” neuron with major components identified

(G. W. Flake, The Computational Beauty of Nature)

Model neuronu

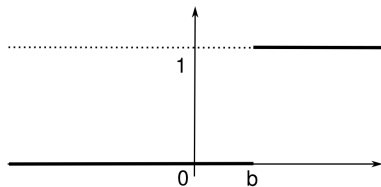


$$a(t + 1) = \Phi\left(\sum_{i=1}^n w_i \times a_i(t) - b\right)$$

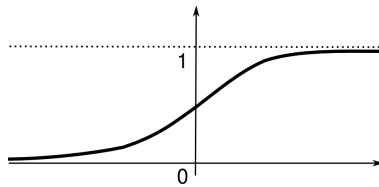
Model neuronu: aktivační funkce

dříve používané (biologicky motivované):

diskrétní aktivační funkce



spojitá aktivační funkce



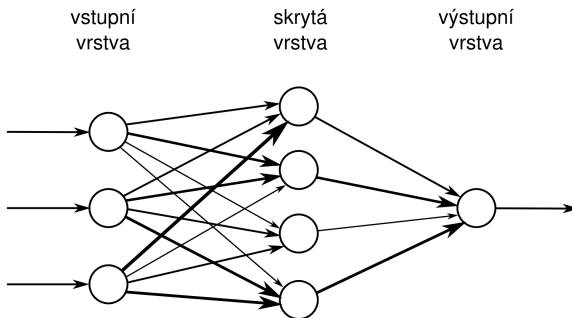
aktuálně populární: „rectifier“ $\max(x, 0)$ (dobrý praktický výkon)

Umělá neuronová síť

... připomíná reálnou neuronovou síť:

- neurony spojené váženými vazbami
- schopná se učit na základě předkládaných dat
- „znalosti“ uloženy jako váhy spojení mezi neurony

Příklad architektury – dopředná síť



Dopředná síť: učení

- předkládáme sadu příkladů: vstup + žádaný výstup
- upravujeme váhy spojení mezi neurony (backpropagation algoritmus)
- učení = korigující negativní zpětná vazba

Architektura

- dopředná síť
 - hrany vedou jedním směrem (žádné cykly)
 - snadnější na naučení, použití
 - neodpovídá biologické realitě
 - dominantně používáno v aktuálních aplikacích
- zpětnovazebná síť
 - reálnější z pohledu modelování
 - náročnější na zpracování

Deep learning

dopředná síť, mnoho úrovní:

- úrovně mohou sloužit jako „stupně abstrakce“ (feature learning)
- problém: náročné síť naučit
- použité techniky nemají biologickou motivaci

Applikace v modelování: Vývoj sdíleného slovníku

příklad kombinace ABM + neuronové sítě

- model **vytváření sdíleného slovníku** ve skupině agentů
- pojmenování reprezentováno pomocí neuronové sítě
- interakce = agenti si vymění svoje názory na název objektu, trochu se přeučí
- díky interakci vzniká sdílený slovník

Shrnutí

- účel modelování myšlení:
 - pochopení, jak myšlení funguje
 - zlepšení modelů (s agenty)
 - aplikace pro řešení „inženýrských“ problémů
- typy modelů:
 - symbolické
 - konekcionistické
 - hybridní (neprobíráno)