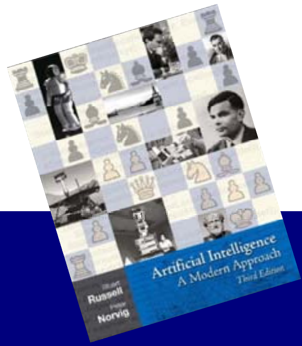


Umělá inteligence II

Roman Barták, KTIML

roman.bartak@mff.cuni.cz
http://ktiml.mff.cuni.cz/~bartak

Na úvod

- Zabýváme se konstrukcí racionálních agentů.
- **Agent** je entita, co vnímá okolní **prostředí** prostřednictvím **senzorů** a ovlivňuje ho prostřednictvím **akčních členů**.
- **Racionální agent** je agent maximalizující očekávanou míru svého zisku.

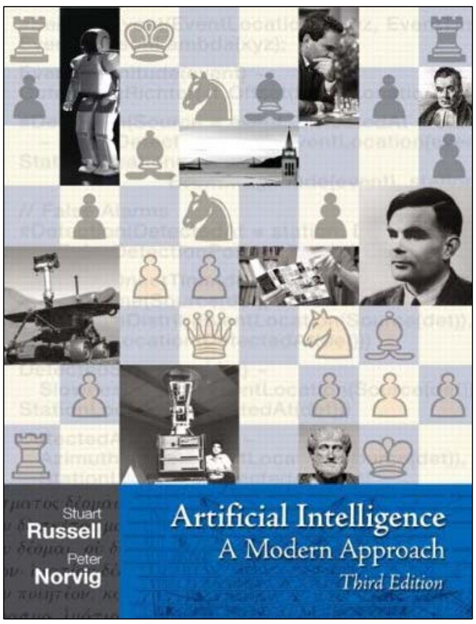
- V UI I jsme se zabývali především logickým přístupem k řešení problému (tvrzení buď platí, neplatí nebo nevíme).
- Ignorovali jsme
  - rozhraní do prostředí (senzory, akční prvky)
  - neurčitost informací
  - schopnost se samostatně zlepšovat (učení)

Obsah přednášky

- **Úvod**
  - motivace, novinky, trochu filozofie
- **Práce s nejistou informací**
  - neurčitost, pravděpodobností uvažování, teorie užitku
- **Strojové učení**
  - rozhodovací stromy, neuronové sítě, zpětnovazební učení
- **Vnímání a chování**
  - zpracování jazyka a obrazu, robotika

Zdroje

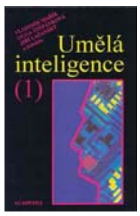
Artificial Intelligence: A Modern Approach

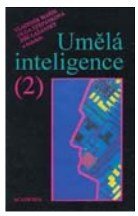


- S. Russell and P. Norvig
- Prentice Hall, 2003 (třetí v.)
- http://aima.cs.berkeley.edu/

Umělá inteligence 1-5

- Vladimír Mařík, Olga Štěpánková, Jiří Lažanský a kol.
- Academia





Web přednášky

http://ktiml.mff.cuni.cz/~bartak/ui2

Zde najdete:

- slajdy k přednášce
- odkazy na zdroje
- informace o zkoušce
- …

Návaznosti

- **Seminář z umělé inteligence II**
  - referativní seminář, kde téma určují účastníci
- **Strojové učení**
  - jak se počítače (sami či s pomocí) učí
- **Neuronové sítě**
  - jak popsat a využít přirozený zdroj inteligence
- **Statistické metody zpracování přirozených jazyků**
  - jak „porozumět" přirozenému jazyku
- **Umělé bytosti**
  - jak to vše spojit dohromady
- …

Aktuálně

- Počítačový systém **Watson** právě porazil lidské šampiony ve hře Jeopardy!
- **O co jde?**
  - odpovídání na otázky položené v přirozeném jazyce s otevřeným obsahem (není omezené téma otázky)
  - systém analyzuje nápovědu v přirozeném jazyce
  - na základě sémantického porovnání s uloženými informacemi vyhledá odpověď
  - pokud má dostatečnou míru jistoty, tak odpoví
- **Použité technologie**
  - zpracování přirozeného jazyka
  - vybavování informací
  - reprezentace znalostí a uvažování o nich
  - strojové učení

Watson inside

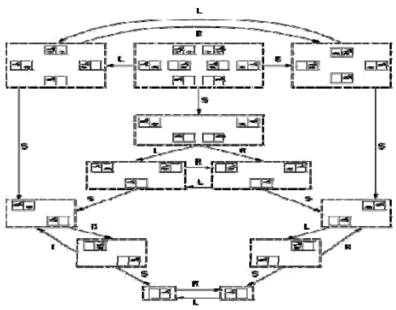
- **Rozumí** Watson kladeným otázkám?
  - ne
- Dělá Watson to samé, co **webové vyhledávače** typu Google?
  - ne (Watson vrací jedinou odpověď a musí si s ní být dostatečně jistý)
- Jak tedy Watson **funguje**?
  - rozloží otázku
  - vygeneruje kandidáty/hypotézy na odpověď
  - hledá podporu pro své hypotézy (masivní paralelismus)
- Odkud Watson čerpal **informace**?
  - přímo z textových dokumentů (včetně celé Wikipedie)

Trochu filozofie

- Watson je reprezentantem **slabé UI**
  - stroje se chovají jako inteligentní bytosti
- **Silná UI**
  - stroje skutečně myslí
- Většina výzkumníků považuje hypotézu slabé UI za potvrzenou a nestarají se o to, jestli se hovoří o simulaci inteligence nebo o skutečnou inteligenci.
  - postupně jsou bourány argumenty typu „stroje nikdy nebudou umět X"

Neurčitost dosud

- Jak jsme se vyrovnali s neurčitostí v čistě logickém přístupu?
- **domnělý (belief) stav**
  - reprezentuje množinu stavů, ve který se agent může nacházet



Problémy

- musíme uvažovat každou interpretaci senzorického vstupu nezávisle na její pravděpodobnosti (velké a složité reprezentace)
- plány akcí musí zachycovat každou eventualitu (velké plány)
- někdy žádný plán negarantuje dosažení cíle, přesto agent musí nějak jednat

Příklad

Diagnostika zubního kazu

Pokusme se ji udělat čistě logicky.

□ když pacienta bolí zuby, má kaz

Toothache $\Rightarrow$ Cavity

Platí to ale vždy?

□ ne každá bolest zubu je způsobena kazem, je tedy potřeba vypsat všechny důvody

Toothache $\Rightarrow$ Cavity $\vee$ GumProblem $\vee$ Abscess $\vee$ …

A co to zkusit obráceně?

Cavity $\Rightarrow$ Toothache

□ to ale také není pravda – ne každý kaz způsobuje bolest

Abychom pravidlo měli správně, musí pokrývat úplný výčet problémů i symptomů.

Použití logiky

Proč použití logiky selhává v oblastech jako je medicínská (a jiná) diagnostika?

- **lenost**: je moc pracné sepsat přesně všechna pravidla a taková pravidla se špatně používají
- **teoretická ignorance**: medicína nemá přesnou teorii pro danou oblast
- **praktická ignorance**: i kdybychom měli úplnou teorii, nemůžeme si být u daného pacienta jisti, dokud neprovedeme všechny diagnostické testy.

Budeme potřebovat jiný nástroj, který umí pracovat se silou domněnek – **teorii pravděpodobnosti**.

Logický agent o každém tvrzení rozhoduje zda platí, neplatí nebo na něj nemá názor, teorie pravděpodobnosti nám umožní kvantifikovat víru v dané tvrzení.

Pravděpodobnost

základní pojmy

- Podobně jako v logice máme možné světy, tak v teorii pravděpodobnosti máme množinu možných světů (říká se jim **elementární jevy**) – **sample space**.
  - elementární jevy jsou **vzájemně výlučné** (nemohou nastat zároveň)
  - prostor světů Ω je **úplný** (obsahuje všechny možné jevy)
- Každý elementární jev ω má přiřazenu pravděpodobnost $P(\omega)$ tak, že platí:
  - $0 \leq P(\omega) \leq 1$
  - $\Sigma_{\omega \in \Omega} P(\omega) = 1$

Příklad:

- elementární jevy jsou dvojice hodnot, které mohou padnout na dvou kostkách (je jich celkem 36)
- $P(\omega) = 1/36$

Jevy

- Dotazy a tvrzení jsou zpravidla o množinách elementárních jevů – tzv. **jevech**.

  např. „padla dvě stejná čísla" je jev
- Pravděpodobnost jevu zjistíme součtem pravděpodobností elementárních jevů, ze kterých se skládá.
  - $P(\phi) = \Sigma_{\omega \in \phi} P(\omega)$

  např.

  $P(doubles) = 1/36+1/36+1/36+1/36+1/36+1/36 = 1/6.$

Těmto pravděpodobnostem se říká **apriorní (nepodmíněné) pravděpodobnosti**.

Podmíněná pravděpodobnost

- Většinou máme o světě nějakou informaci (**evidence**) a zajímá nás pravděpodobnost nějakého jevu.
  - Jaká je pravděpodobnost páru, pokud víme, že na kostce 1 padla 5?

    $P(doubles \mid Die_1 = 5) = 1/36 / (6*1/36) = 1/6$
- **Podmíněná pravděpodobnost**

  $P(a \mid b) = P(a \wedge b) / P(b)$, pokud $P(B) \neq 0$

  často se zapisuje v podobě **součinového pravidla**

  $P(a \wedge b) = P(a \mid b) . P(b)$

Pozor! Pokud máme více vstupních informací, zajímá nás jiná podmíněná pravděpodobnost.
$P(doubles \mid Die_1 = 5, Die_2 = 5) = 1$

Náhodná proměnná

- Jak kompaktně popsat možné světy (elementární jevy)?
- Použijeme **náhodné proměnné** popisující vlastnosti světa. Každá proměnná může nabývat hodnot ze své domény (podobně jako u CSP, uvažujeme diskrétní domény).
  - Die_1 – hodnota, která padla na kostce 1 (1,…,6)
  - Cavity – pacient má či nemá kaz (true, false)
- Elementární jev je potom charakterizován hodnotami všech náhodných proměnných.

  $P(Die_1 = 5, Die_2 = 5)$

Sdružená distribuce

- Pravděpodobnosti elementárních jevů můžeme popsat tabulkou, tzv. **úplnou sdruženou distribucí**, kde jednotlivé prvky jsou indexovány hodnotami náhodných proměnných.

	toothache		*¬ toothache*	
	catch	*¬ catch*	*catch*	*¬ catch*
cavity	.108	.012	.072	.008
¬ cavity	.016	.064	.144	.576

- Z tabulky snadno vypočteme pravděpodobnosti jednotlivých hodnot náhodných proměnných
  - P(toothache=true) = 0.108 + 0.012 + 0.016 + 0.064 = 0.2
  - P(toothache=false) = 0.072+ 0.008 + 0.144 + 0.576 = 0.8

  Zkráceně budeme zapisovat takto:
  - **P**(toothache) = ⟨0.2, 0.8⟩

Vlastnosti

- P(¬a) = 1 – P(a)
- **princip inkluze a exkluze**

  P(a ∨ b) = P(a) + P(b) – P(a ∧ b)
- **řetězcové pravidlo**

  P(A,B,C,D)

  = **P**(A|B,C,D) **P**(B,C,D)

  = **P**(A|B,C,D) **P**(B|C,D) **P**(C,D)

  = **P**(A|B,C,D) **P**(B|C,D) **P**(C|D) **P**(D)

Odvozování
ze sdružené distribuce

- Jak odpovídat na otázky?
- Báze znalostí je reprezentována úplnou sdruženou distribucí.

	toothache		*¬ toothache*	
	catch	*¬ catch*	*catch*	*¬ catch*
cavity	.108	.012	.072	.008
¬ cavity	.016	.064	.144	.576

- Chceme-li znát pravděpodobnost nějakého tvrzení, sečteme pravděpodobnosti všech „světů", kde tvrzení platí (**marginalizace**).

$$P(\phi) = \Sigma_{\omega:\omega|=\phi} P(\omega)$$
$$\mathbf{P}(\mathbf{Y}) = \Sigma_{z\in\mathbf{Z}} \mathbf{P}(\mathbf{Y},z)$$

Příklady odvozování
z úplné sdružené distribuce

	toothache		*¬ toothache*	
	catch	*¬ catch*	*catch*	*¬ catch*
cavity	.108	.012	.072	.008
¬ cavity	.016	.064	.144	.576

$$P(\phi) = \Sigma_{\omega:\omega|=\phi} P(\omega)$$
$$\mathbf{P}(\mathbf{Y}) = \Sigma_{z\in\mathbf{Z}} \mathbf{P}(\mathbf{Y},z)$$

P(toothache) (= P(Toothache=true))
$= 0.108 + 0.012 + 0.016 + 0.064 = 0.2$

P(cavity $\vee$ toothache)
$= 0.108+0.012+0.072+0.008+0.016+0.064 = 0.28$

P($\neg$cavity|toothache)
$= P(\neg\text{cavity} \wedge \text{toothache}) / P(\text{toothache})$
$= (0.016 + 0.064) / (0.108 + 0.012 + 0.016 + 0.064)$
$= 0.4$

Normalizace

	toothache		¬ toothache	
	catch	¬ catch	catch	¬ catch
cavity	.108	.012	.072	.008
¬ cavity	.016	.064	.144	.576

P(¬cavity|toothache)
= P(¬cavity ∧ toothache) / P(toothache)
= (0.016 + 0.064) / (0.108 + 0.012 + 0.016 + 0.064) = 0.4

P(cavity|toothache)
= P(cavity ∧ toothache) / P(toothache)
= (0.108 + 0.012) / (0.108 + 0.012 + 0.016 + 0.064) = 0.6

V obou případech je jmenovatel stejný!

Nepotřebujeme ho dokonce ani znát, protože víme, že P(¬cavity|toothache) + P(cavity|toothache) = 1

Můžeme tedy uvažovat **normalizační konstantu** α, jejíž hodnotu dopočteme tak, aby příslušná distribuce byla v součtu 1.

P(Cavity|toothache) = α **P**(Cavity,toothache)
= α [**P**(Cavity,toothache,catch) + α **P**(Cavity,toothache,¬catch)]
= α [⟨0.108, 0.016⟩ + ⟨0.012, 0.064⟩]
= α [⟨0.12, 0.08⟩] = [⟨0.6, 0.4⟩]

Enumerační odvozování

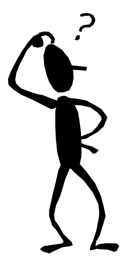
krátké shrnutí

- Typicky známe hodnoty e proměnných **E** z **pozorování** a zajímá nás pravděpodobností distribuce pro hodnoty proměnných **Y** z **dotazu**.
  Zbylé proměnné jsou **skryté H** = **X** – **Y** – **E**.

$\mathbf{P}(\mathbf{Y} \mid \mathbf{E}=e) = \alpha\, \mathbf{P}(\mathbf{Y}, \mathbf{E}=e) = \alpha\, \Sigma_h\, \mathbf{P}(\mathbf{Y},\mathbf{E}=e,\mathbf{H}=h)$

Zřejmé problémy:

- časová složitost v nejhorším případě $O(d^n)$, kde d je max. počet hodnot náhodné proměnné
- pro uložení úplné sdružené distribuce potřebujeme prostor $O(d^n)$
- jak zjistíme hodnoty všech prvků úplné sdružené distribuce?



Nezávislost

- Uvažujme v naší sdružené distribuci ještě čtvrtou proměnnou Weather se čtyřmi hodnotami {cloudy, sunny, rain, snow}, tj. úplná sdružená distribuce má 2x2x2x4 = 32 prvků.

  P(toothache,catch,cavity,cloudy)
  = P(cloudy|toothache,catch,cavity) * P(toothache,catch,cavity)

Opravdu problémy se zuby ovlivňují počasí?
  P(cloudy|toothache,catch,cavity) = P(cloudy)

Můžeme tedy obecně psát:

  P(Toothache,Catch,Cavity,Weather)
  = **P**(Toothache,Catch,Cavity) * **P**(Weather)

Úplná sdružená distribuce tak lze popsat dvěma tabulkami, jedna s 8 prvky a druhá se 4 prvky.

Této vlastnosti se říká **(úplná) nezávislost**:

P(X|Y) = **P**(X) nebo **P**(Y|X) = **P**(Y) nebo **P**(X,Y) = **P**(X).**P**(Y)

Podmíněná nezávislost

- Úplná nezávislost nám umožňuje **zmenšit reprezentaci** sdružené distribuce, bohužel je málo častá a vlastní nezávislé množiny mohou být pořád velké.
- Můžeme jít ale dál.
- Pokud máme kaz (cavity), záleží zachycení nástroje (catch) na bolesti (toothache)?
  P(catch | toothache, cavity) = P(catch | cavity)
  P(catch | toothache, ¬ cavity) = P(catch | ¬ cavity)
- Proměnná Catch je podmíněně nezávislá na proměnné Toothache při znalosti Cavity.
  P(Catch | Toothache, Cavity) = **P**(Catch | Cavity)

Této vlastnosti se říká **podmíněná nezávislost**:

P(X|Y,Z) = **P**(X|Y) nebo **P**(Z|X,Y) = **P**(Z|Y) nebo
P(Z,X|Y) = **P**(Z|Y) **P**(X|Y)

Podmíněná nezávislost

využití

- Podmíněnou nezávislost můžeme dále využít pro redukci velikosti reprezentace sdružené distribuce.

  P(Toothache,Catch,Cavity)
  = **P**(Toothache|Catch,Cavity) **P**(Catch,Cavity)
  = **P**(Toothache|Catch,Cavity) **P**(Catch|Cavity) **P**(Cavity)
  = **P**(Toothache|Cavity) **P**(Catch|Cavity) **P**(Cavity)

- Pro reprezentaci teď potřebujeme tři tabulky o velikosti 2 + 2 + 1 = 5 (stačí reprezentovat nezávislé prvky).