

# Umělá inteligence (2. přednáška)

## Agenti

Anna Vernerová <sup>1</sup>    Jonathan L. Verner <sup>2</sup>

<sup>1</sup>Ústav formální a aplikované lingvistiky  
Matematicko-fyzikální fakulta  
<http://ufal.mff.cuni.cz/anna-vernerova/umela-inteligence>  
[vernerova@ufal.mff.cuni.cz](mailto:vernerova@ufal.mff.cuni.cz)

<sup>2</sup>Katedra logiky  
Filozofická fakulta  
<http://jonathan.temno.eu/teaching/ALG500008.html>  
[jonathan.verner@ff.cuni.cz](mailto:jonathan.verner@ff.cuni.cz)

Racionální chování je takové, které na základě dostupných informací volí akce nejpravděpodobněji vedoucí k maximalizaci “užitku”.

$$f : \mathcal{P}^* \rightarrow \mathcal{A}$$

Agent = architektura + program.

Je třeba specifikovat

- možné vjemy (množina  $\mathcal{P}$ )
- možné akce (množina  $\mathcal{A}$ )
- prostředí
- měřítko úspěšnosti

# Příklady

<b>Agent</b>	<b>Vjemy</b>	<b>Akce</b>	<b>Prostředí</b>	<b>Úspěšnost</b>
hráč šachů	tahy protivráče, stav šachovnice	vlastní tahy	šachovnice, pravidla	Elo
automatický vysavač	senzor špíny	vysávání, pohyb	byt, kancelář	čistota, hlučnost, vysátá špína
autonomní vozidlo	video obraz vozovky, rychlost, GPS, ...	přidání plynu, otočení volantem, ...	vozovka, ostatní vozidla	bezpečnost, spolehlivost, rychlost

# Klasifikace prostředí

- plná vs. částečná pozorovatelnost
  - šachy
  - mariáš
- deterministické vs. stochastické vs. nedeterministické vs. strategické
  - deterministické — následující stav prostředí je plně určen předchozím stavem a akcí agenta
  - strategické — stav prostředí mění pouze (další) agenti
  - šachy, mariáš, vysavač
  - autonomní vozidlo
- epizodické vs. sekvenční
  - epizodické — následující vjemy nezávislé na předchozích vjemech a akcích
  - výstupní kontrola kvality
- statické vs. dynamické
  - mariáš, šachy (bez hodin)
  - autonomní vozidlo, šachy (s hodinami)

- diskrétní vs. spojité
  - šachy, mariáš
  - autonomní vozidlo

(Pozor, počítače vždy pracují s *digitálním*, t.j. diskrétním vstupem; je otázka, zda svět samotný není diskrétní.)

- jedno agentní vs. multiagentní
  - jako agenty modelujeme ty objekty v prostředí, které se snaží maximalizovat nějakou užitkovou funkci
  - spolupracující agenti: komunikace
  - kompetitivní multiagentní prostředí: randomizované jednání zabrání předvídatelnosti

Racionální agent se snaží maximalizovat hodnotící funkci.

Chceme hodnotit chování nebo stav prostředí?

- dobrý vysavač vysaje hodně špíny, hodnotící funkce dává body za vysátou špínu  
ideální strategií je vysát špínu, pak ji vysypat, pak zase vysát a tak dokola. . .
- dobrý vysavač udržuje místnost čistou

Obecně je lepší hodnotit stav prostředí, často totiž není jasné, jaké chování je dobré.

Racionální neznamená

- vševědoucí — vjemy nemusí poskytnout všechny relevantní informace
- vidící do budoucnosti — nemůže předvídat výsledky některých akcí

Tudíž racionální  $\neq$  úspěšný.

Racionální chování zahrnuje akce vedoucí k získávání informací (*information gathering*) a proces učení (*autonomní agent* je díky procesu učení schopen kompenzovat neúplnou nebo i chybnou znalost svého tvůrce).

# Základní typy agentů

- reflexivní agenti
  - reagují pouze na aktuální vjem
  - pokud vjemy neposkytují dostatečný obraz světa, agent selže
  - částečně lze řešit randomizací akcí
- reflexivní agenti s vnitřním stavem
  - buduje si model vnějšího světa
- „goal-based“ agenti
  - agent uvažuje o budoucnosti
  - **hledá** posloupnost akcí, které vedou k dosažení cíle (**plánování**)
- „utility-based“ agenti
  - cíle jsou pouze binární
  - cíle jsou často konfliktní
  - agent se proto nesnaží dosáhnout cíle, nýbrž maximalizovat “žádoucnost” stavu