

Umělá inteligence (2. přednáška)

Agenti

Anna Vernerová ¹ Jonathan L. Verner ²

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

²Katedra logiky
Filozofická fakulta
<http://jonathan.temno.eu/teaching/ALG500008.html>
jonathan.verner@ff.cuni.cz

Návrh racionálního agenta

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

Návrh racionálního agenta

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

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

Návrh racionálního agenta

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.

Návrh racionálního agenta

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 $\mathcal{P}$)

Návrh racionálního agenta

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 $\mathcal{A}$)

Návrh racionálního agenta

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

Návrh racionálního agenta

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

Agent

Vjemy

Akce

Prostředí

Úspěšnost

Příklady

Agent	Vjemy	Akce	Prostředí	Úspěšnost
hráč šachů				

Příklady

Agent	Vjemy	Akce	Prostředí	Úspěšnost
hráč šachů	tahy pro- tihráče, stav šachovnice			

Příklady

Agent	Vjemy	Akce	Prostředí	Úspěšnost
hráč šachů	tahy protivníka, stav šachovnice	vlastní tahy		

Příklady

Agent	Vjemy	Akce	Prostředí	Úspěšnost
hráč šachů	tahy protivníka, stav šachovnice	vlastní tahy	šachovnice, pravidla	

Příklady

Agent	Vjemy	Akce	Prostředí	Úspěšnost
hráč šachů	tahy pro- tihráče, stav šachovnice	vlastní tahy	šachovnice, pravidla	Elo

Příklady

Agent	Vjemy	Akce	Prostředí	Úspěšnost
hráč šachů	tahy protivráče, stav šachovnice	vlastní tahy	šachovnice, pravidla	Elo
automatický vysavač				

Příklady

Agent	Vjemy	Akce	Prostředí	Úspěšnost
hráč šachů	tahy protivráče, stav šachovnice	vlastní tahy	šachovnice, pravidla	Elo
automatický vysavač	senzor špíny			

Příklady

Agent	Vjemy	Akce	Prostředí	Úspěšnost
hráč šachů	tahy protivráče, stav šachovnice	vlastní tahy	šachovnice, pravidla	Elo
automatický vysavač	senzor špíny	vysávání, pohyb		

Příklady

Agent	Vjemy	Akce	Prostředí	Úspěšnost
hráč šachů	tahy protivráče, stav šachovnice	vlastní tahy	šachovnice, pravidla	Elo
automatický vysavač	senzor špíny	vysávání, pohyb	byt, kancelář	

Příklady

Agent	Vjemy	Akce	Prostředí	Úspěšnost
hráč šachů	tahy protivráče, stav šachovnice	vlastní tahy	šachovnice, pravidla	Elo
automatický vysavač	senzor špíny	vysávání, pohyb	byt, kancelář	čistota, hlučnost, vysátá špína

Příklady

Agent	Vjemy	Akce	Prostředí	Úspěšnost
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				

Příklady

Agent	Vjemy	Akce	Prostředí	Úspěšnost
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			

Příklady

Agent	Vjemy	Akce	Prostředí	Úspěšnost
hráč šachů	tahy protivníka, 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			

Příklady

Agent	Vjemy	Akce	Prostředí	Úspěšnost
hráč šachů	tahy protivníka, 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			

Příklady

Agent	Vjemy	Akce	Prostředí	Úspěšnost
hráč šachů	tahy protivníka, 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, ...			

Příklady

Agent	Vjemy	Akce	Prostředí	Úspěšnost
hráč šachů	tahy protihráče, stav šachovnice	vlastní tahy	šachovnice, pravidla	Elo
automatický vysavač	senzor špíny	vysávání, pohyb	byt, kancelář	čistota, hluchnost, vysátá špína
autonomní vozidlo	video obraz vozovky, rychlost, GPS, ...	přidání plynu		

Příklady

Agent	Vjemy	Akce	Prostředí	Úspěšnost
hráč šachů	tahy protivníka, stav šachovnice	vlastní tahy	šachovnice, pravidla	Elo
automatický vysavač	senzor špíny	vysávání, pohyb	byt, kancelář	čistota, hluchost, vysátá špína
autonomní vozidlo	video obraz vozovky, rychlost, GPS, ...	přidání plynu, otočení volantem		

Příklady

Agent	Vjemy	Akce	Prostředí	Úspěšnost
hráč šachů	tahy protihráče, stav šachovnice	vlastní tahy	šachovnice, pravidla	Elo
automatický vysavač	senzor špíny	vysávání, pohyb	byt, kancelář	čistota, hluchost, vysátá špína
autonomní vozidlo	video obraz vozovky, rychlost, GPS, ...	přidání plynu, otočení volantem, ...		

Příklady

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

Příklady

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

Příklady

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

Příklady

Agent	Vjemy	Akce	Prostředí	Úspěšnost
hráč šachů	tahy protihráče, stav šachovnice	vlastní tahy	šachovnice, pravidla	Elo
automatický vysavač	senzor špíny	vysávání, pohyb	byt, kancelář	čistota, hluchost, 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

Klasifikace prostředí

- plná vs. částečná pozorovatelnost
 - šachy

Klasifikace prostředí

- plná vs. částečná pozorovatelnost
 - šachy
 - mariáš

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

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

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

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

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

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

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

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

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)

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

- diskrétní vs. spojité
 - šachy, mariáš

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

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

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

- 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

- 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

- 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

- 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.

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

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...

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

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

Racionální neznamená

- vševědoucí — vjemy nemusí poskytnout všechny relevantní informace

Racionální neznamená

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

Racionální neznamená

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

Tudíž racionální $\neq$ úspěšný.

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*)

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

Základní typy agentů

- reflexivní agenti
 - reagují pouze na aktuální vjem

Základní typy agentů

- reflexivní agenti
 - reagují pouze na aktuální vjem
 - pokud vjemy neposkytují dostatečný obraz světa, agent selže

Základní typy agentů

- reflexivní agenti
 - reagují pouze na aktuální vjem
 - pokud vjemy neposkytují dostatečný obraz světa, agent selže
 - částečně lze řešit randomizací akcí

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

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

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

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

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**)

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

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

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

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