Základy umělé inteligence

Multiagentní systémy, Teorie her

Ing. Tomáš Řehořek

Katedra teoretické informatiky (KTI), Fakulta informačních technologií (FIT)
České vysoké učení technické v Praze (ČVUT)

BI-ZUM, LS 2016/17, 7. přednáška

https://edux.fit.cvut.cz/courses/BI-ZUM/

Osnova přednášky

Multiagentní systémy

Multiagentní systém je kolekce několika **nezávislých agentů** umístěných ve sdíleném **prostředí**, z nichž každý

- **vnímá** okolní prostředí,
- **jedná** (provádí akce) tak, aby dosáhl svých **cílů**,
- **interaguje** s ostatními agenty
  - potenciální **kooperace** a **koordinace akcí**

Aplikační oblasti mutliagentních systémů:

- **distribuované systémy**
  - není možno nebo není efektivní vytvořit a spravovat globální znalost,
- **vysoce dynamická prostředí**
  - je potřeba rychlá reakce a častá změna plánů,
- **nekooperativní rozvržení**
  - prostředí agentů, kteří nechtějí spolupracovat

Definice agenta

Russel & Norwig

Agent je cokoli, co je schopno **vnímat prostředí** pomocí svých **senzorů** a v tomto prostředí **jednat** pomocí svých **efektorů**.

- definice se zaměřuje na **umístění v prostředí** (*situatedness, embodiment*),
- agent může prostředí ovlivňovat, nikoli však plně ovládat
  - selhání senzorů, efektorů, nedeterminismus. . .

Wooldridge & Jennings

Agent je **počítačový systém**, který je **umístěn v nějakém prostředí** a v tomto prostředí je schopen **samostatné činnosti** za účelem dosažení svých (anebo delegovaných) cílů.

- definice zavádí další rozměr: vztah mezi agentem a jeho návrhářem (uživatelem)
  - agent je schopen samostatné činnosti,
  - tato činnost je **účelná** (*purposeful*),
- klíčovou vlastností agenta je **autonomie**

Typy agentů

Agenti mohou být:

- **izolovaní** (*isolated*)
  - ekvivalent single-agentního systému,
- **kooperativní** (*cooperative*)
  - jednající společně za účelem dosažení týmového cíle,
- **self-interested**
  - každý maximalizuje své vlastní dobro, zvažuje ostatní agenty,
- ***kombinací výše uvedených***
  - spolupracující s jedněmi a soutěžící s jinými agenty

Fundamentální vlastnosti agenta

Inteligentní agent je:

- **autonomní**
  - řídí se snahou o dosažení svých cílů, k čemuž nevyžaduje žádné „rady zvenčí“,
  - může si vybrat vlastní cíl i způsob, jak jej dosáhnout,
  - nemáme nad ním přímou kontrolu, jeho chování je determinováno zkušeností,
- **reaktivní**
  - udržuje s prostředím nepřetržitou interakci,
  - reaguje na změny, které se v prostředí odehrávají,
- **proaktivní**
  - generuje cíle a snaží se jich dosáhnout,
  - není řízen pouze událostmi v prostředí, nýbrž i vlastní iniciativou,
- **sociální**
  - interaguje s ostatními agenty
    - **koordiace** – správa vzájemných vazeb mezi akcemi skupiny agentů,
    - **kooperace** – týmová práce za účelem dosažení společného cíle,
    - **vyjednávání** – schopnost dosahovat shod ve věcech společného zájmu

Task environments (PEAS)

Multiagentní problém lze specifikovat pomocí tzv. ***task environment***:

1. **Metrika úspěšnosti** (*Performance measure*),
   - rychlé prohledání bludiště, přesun automobilu z místa *A* do místa *B* rychle a bez nehod, eliminace co nejvíce nepřátel v počítačové hře
2. **Prostředí** (*Environment*),
   - bludiště, silniční síť, herní mapa strategické hry,
3. **Efektory** (*Actuators*),
   - kola, nohy, zbraně, vysílání zpráv ostatním agentům,
4. **Senzory** (*Sensors*),
   - kamera, radar, příjem informací od jiných agentů,

Task environment definuje problém, jehož řešením je návrh odpovídajícího agenta.

Vlastnosti prostředí

Prostředí můžeme popsat různými charakteristikami (známe již z plánování):

- **plně pozorovatelné** vs. **částečně pozorovatelné**
  - mají agenti **kompletní** a správnou informaci o stavu prostředí?
- **deterministické** vs. **stochastické**
  - mají akce prováděné efektory garantovaný a jednoznačně definovaný efekt?
- **epizodální** vs. **sekvenční**
  - mají akce dlouhodobé efekty, anebo se úkoly opakují v nezávislých epizodách?
- **statické** vs. **dynamické**
  - podléhá prostředí procesům, které nemá agent pod kontrolou?
- **diskrétní** vs. **spojité**
  - jsou množiny možných akcí a pozorování pevně dány a jsou konečné?
- **single-agent** vs. **multi-agent**
  - záleží chování jednoho agenta na chování ostatních agentů?

Agentní funkce

Chování agenta je popsáno tzv. **agentní funkcí** (*agent function*) mapující **vnímané sekvence** na **sekvence akcí**:

$$f\colon \mathcal{P} \to \mathcal{A}.$$

Tato funkce může být realizována **agentním programem**, který běží na nějaké **fyzické architektuře**.

Klíčové otázky:

- která funkce je správná?
- lze tuto funkci implementovat malým agentním programem?

Typy agentů z hlediska agentních funkcí

Rozlišujeme 4 úrovně agentů z hlediska komplexnosti:

1. **Jednoduchý reflexní agent**
    - reaguje vždy pouze na **aktuální vjemy**, nemá vnitřní stav,
    - řízen sadou pevně zakódovaných **if-then** pravidel,
2. **Reflexní agent se stavem**
    - má **paměť**, která umožňuje ukládat **historii vjemů**,
    - stále řízen pevně zakódovanými pravidly,
3. **Goal-based agent**
    - umí přijímat cíle, které má splnit (chování není pevně zakódováno),
    - akce jsou voleny s ohledem na **cíle** pomocí **plánování**,
4. **Utility-based agent**
    - zavádí tzv. **utilitní funkci**
        - mapuje stav prostředí na reálné číslo,
        - vyjadřuje kvalitu stavu,
        - umožňuje řešit konflikty mezi více cíli či možnými způsoby splnění cílů

Utilitní funkce

Aby mělo smysl jakkoli jednat (provádět v prostředí akce), musí být některé stavy považovány za lepší než stavy jiné.

Dva základní přístupy k porovnávání stavů:

- explicitní **relace preference** $\succcurlyeq$ na množině stavů,
- **utilitní funkce** $u\colon S \to \mathbb{R}$, která numericky popisuje „kvalitu“ stavu
  - $u(s_1) \geq u(s_2) \Leftrightarrow s_1 \succcurlyeq s_2$.

Příklad: Výhra $\succcurlyeq$ Remíza, Výhra $\succcurlyeq$ Prohra, Remíza $\succcurlyeq$ Prohra

- $u(\text{Výhra}) = 1$,
- $u(\text{Remíza}) = 0$,
- $u(\text{Prohra}) = -1$.

Racionalita

Agent může měnit stav prostředí pomocí svých efektorů, mnohdy je však výsledek provedení akcí nejistý.

Definice: Loterie

Loterie je pravděpodobnostní rozdělení na množině možných výsledků. Značíme:

$$[p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k],$$

kde o_i jsou možné výsledky a $p_i > 0$ jsou pravděpodobnosti takové, že

$$\sum_{i=1}^{k} p_i = 1.$$

Loterie slouží ke specifikaci faktu, že akce skončí výsledkem o_i s pravděpodobností p_i.

Nadále budeme uvažovat, že **samotná loterie je výsledkem akce**.

Racionalita

Uvažujme množinu agentů $\mathcal{A} = \{A_1, \ldots, A_n\}$, kde každý agent volí mezi akcemi, jejichž provedením dojde ke spuštění nějaké loterie $\ell \in \mathcal{L}$. Ztožněme akce s odpovídajícími loteriemi, tj. převeďme volbu vhodné akce na volbu vhodné loterie.

Uvažujme dále, že každý agent A_i má přiřazenu utilitní funkci u_i, takovou, že $u_i(o_j)$ vyjadřuje utilitu výsledku o_j pro agenta A_i.

Definice: Self-interested racionální agent

Self-interested racionální agent je takový agent A_i, který v každém okamžiku volí akci spouštějící loterii ℓ^* maximalizující očekávanou hodnotu své osobní utility u_i:

$$\ell^* \in \arg\max_{\ell \in \mathcal{L}} \sum_{(p_j :\ o_j) \in \ell} p_j \cdot u_i(o_j)$$

Definice: Kooperativní racionální agent

Kooperativní racionální agent je takový agent A_i, který v každém okamžiku volí akci spouštějící loterii ℓ^* maximalizující očekávanou hodnotu sumy utilit všech agentů z $\mathcal{A}$:

$$\ell^* \in \arg\max_{\ell \in \mathcal{L}} \left(\sum_{A_k \in \mathcal{A} \setminus \{A_i\}} \left(\sum_{(p_j :\ o_j) \in \ell} p_j \cdot u_k(o_j) \right) + \sum_{(p_j :\ o_j) \in \ell} p_j \cdot u_i(o_j) \right).$$

Osnova přednášky

Teorie her

- Matematická disciplína studující komplikované situace, kdy jsou utilitní funkce jednotlivých agentů ve vzájemné interakci,
- Zabývá se jazyky pro zápis struktury her, jejich analýzou a zkoumáním strategií,
- Základní kategorie:
  - **kooperativní hry** – modelovací jednotkou je tým hráčů,
  - **nekooperativní hry** – modelovací jednotkou je jeden hráč,
- Hry mohou být popsány různými teoretickými aparáty:
  - **normální forma** – hra vyjádřena **maticí**,
    - náplň zbytku dnešní přednášky,
  - **extenzivní forma** – hra vyjádřena **stromem**
    - náplň příští přednášky

Hry v normální formě

Definice: Konečná hra v normální formě

Konečná hra v normální formě pro n hráčů je trojice $(\mathcal{N}, \mathcal{A}, u)$, kde

- $\mathcal{N} = \{N_1, \ldots, N_n\}$ je konečná množina **hráčů**,
- $\mathcal{A} = A_1 \times \ldots \times A_n$, kde A_i je množina akcí, které má k dipozici hráč N_i,
  - $\mathbf{a} \in \mathcal{A}$ je tzv. **akční profil**, vyjadřující jednu konkrétní volbu akcí provedenou nezávisle **všemi hráči** ve hře,
- $u = (u_1, \ldots, u_n)$ je posloupnost **utilitních funkcí** pro jednotlivé hráče
  - $u_i : \mathcal{A} \to \mathbb{R}$,
  - $u_i(\mathbf{a})$ vyjadřuje utilitu hráče N_i z akčního profilu $\mathbf{a} \in \mathcal{A}$

Příklad: Kámen-nůžky-papír

Rock-paper-scissors

$N_1 \backslash N_2$	K	N	P
K	**0**, **0**	**1**, **-1**	**-1**, **1**
N	**-1**, **1**	**0**, **0**	**1**, **-1**
P	**1**, **-1**	**-1**, **1**	**0**, **0**

$$\mathcal{N} = \{N_1, N_2\}$$

$$\begin{aligned}\mathcal{A} = \{K, N, P\} \times \{K, N, P\} = \{&(K, K), (K, N), (K, P),\\ &(N, K), (N, N), (N, P),\\ &(P, K), (P, N), (P, P)\,\}\end{aligned}$$

$$\begin{array}{llll} u_1: & (K,K) \mapsto 0, & (K,N) \mapsto 1, & (K,P) \mapsto -1,\\ & (N,K) \mapsto -1, & (N,N) \mapsto 0, & (N,P) \mapsto 1,\\ & (P,K) \mapsto 1, & (P,N) \mapsto -1, & (P,P) \mapsto 0 \end{array}$$

$$\begin{array}{llll} u_2: & (K,K) \mapsto 0, & (K,N) \mapsto -1, & (K,P) \mapsto 1,\\ & (N,K) \mapsto 1, & (N,N) \mapsto 0, & (N,P) \mapsto -1,\\ & (P,K) \mapsto -1, & (P,N) \mapsto 1, & (P,P) \mapsto 0 \end{array}$$

Kanonické hry v normální formě

Vězňovo dilema

Dva vězni jsou zatčeni a oddeleně vyslýcháni. Mají se rozhodnout, zda udat komplice, přičemž platí:

- pokud ani jeden vězeň komplice neudá, dostanou oba trest 1 rok vězení,
- pokud *A* udá a *B* neudá, bude *B* potrestán 5 lety vězení a *A* propuštěn,
- pokud *A* neudá a *B* udá, bude *A* potrestán 5 lety vězení a *B* propuštěn,
- pokud oba dva vězni komplice udají, dostane každý trest 3 roky vězení.

$N_1 \backslash N_2$	C	D
C	1, 1	5, 0
D	0, 5	3, 3

Zobecněná varianta:

$N_1 \backslash N_2$	C	D
C	a, a	c, b
D	b, c	d, d

$b < a < d < c$

Kanonické hry v normální formě

Common-payoff

Definice: Common-payoff hra

Uvažujme hru v normální formě $G = (\mathcal{N}, \mathcal{A}, u)$, kde $u = (u_i, \ldots, u_n)$. Řekneme, že G je **common-payoff** hra, jestliže

$$\forall \mathbf{a} \in \mathcal{A} \colon u_1(\mathbf{a}) = u_2(\mathbf{a}) = \ldots = u_n(\mathbf{a}).$$

Příklady common-payoff her:

- Po které straně silnice jezdit?
- Kde společně strávit večer?

N_1 \ N_2	L	R
L	1, 1	0, 0
R	0, 0	1, 1

N_1 \ N_2	home	party
home	2, 2	0, 0
party	0, 0	5, 5

Kanonické hry v normální formě

Constant-sum

Definice: Constant-sum hra

Uvažujme hru v normální formě $G = (\mathcal{N}, \mathcal{A}, u)$, kde $u = (u_i, \dots, u_n)$. Řekneme, že G je **constant-sum** hra, jestliže

$$\exists c \in \mathbb{R} : \forall \mathbf{a} \in \mathcal{A} : \sum_{i=1}^{n} u_i(\mathbf{a}) = c.$$

Speciální případ, kdy $c = 0$, nazýváme **zero-sum** hra.

Příklady:

- Kámen-nůžky-papír,
- Matching pennies

$N_1 \backslash N_2$	H	T
H	1, -1	-1, 1
T	-1, 1	1, -1

Analýza her v normální formě

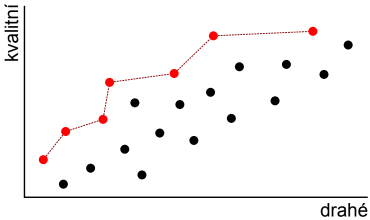
- Hry popsané v normální formě poskytují pohled „shora“
  - hráči jsou rovnocenní,
  - nenahlížíme optikou „hráč 1 má vyhrát, hráč 2 má prohrát“,
  - zajímavé je hledální akčních profilů optimálních z jistých „globálních“ hledisek.
- Které akční profily jsou „zajímavé“?
  - **Paretova optima**
    - ★ změnou akčního profilu již žádný hráč nemůže zvýšit svou utilitu, aniž se by snížila utilita jiných hráčů,
  - **Nashovo equilibria**
    - ★ žádný hráč změnou své akce (s ohledem na kompletní akční profil skupiny hráčů) nemůže zvýšit svou utilitu,
    - ★ akční profil je **stabilní** – žádný hráč není motivován změnit své rozhodnutí, ani tohoto rozhodnutí litovat

Paretovo optimum

- pojem z **multikriteriální optimalizace**,
- z oblasti optimalizace (hill climbing, evoluční algoritmy. . .) dosud známe problémy s jedinou kriteriální funkcí
  - optimalizační problém jsme definovali jako hledání $\arg\max_{\mathbf{x} \in X} f(\mathbf{x})$ pro jednu kriteriální funkci $f \colon X \to \mathbb{R}$,
- mnoho praktických problémů lze chápat jako hledání **kompromisu** mezi **vzájemně si odporujícími** kritérii
  - např. nákup levných, kvalitních výrobků:
    - levné výrobky bývají nekvalitní,
    - kvalitní výrobky bývají drahé

Paretovo optimum: Příklad

Uvažujme jistou kategorii výrobků na trhu, které lze pořídit v různých kvalitách za různé ceny:



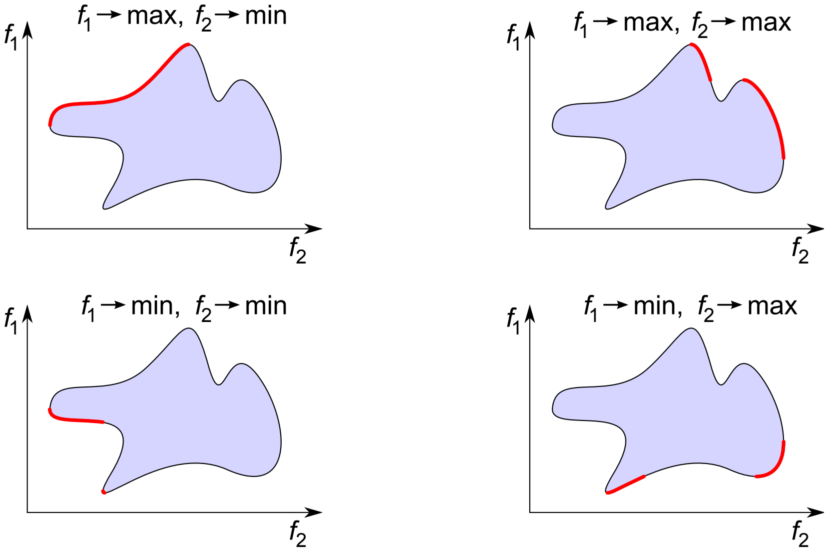
Výrobky označené červeně **jsou** paretovsky optimální, neboť:

- neexistuje výrobek o stejné nebo nižší ceně, který by byl kvalitnější,
- neexistuje výrobek o stejné nebo vyšší kvalitě, který by byl levnější.

Výrobky označené černě **nejsou** paretovsky optimální, neboť:

- existuje kvalitnější výrobek o stejné nebo nižší ceně nebo (or)
- existuje levnější výrobek o stejné nebo vyšší kvalitě.

Paretovsky optimální linie: Možné případy



Paretovo optimum v teorii her

Definice: Paretovská dominance

Uvažujme hru v normální formě $(\mathcal{N}, \mathcal{A}, u)$, kde $u = (u_1, u_2, \dots, u_n)$. Řekneme, že akční profil $\mathbf{a}' = (a'_{N_1}, a'_{N_2}, \dots, a'_{N_n}) \in \mathcal{A}$ **paretovsky dominuje** akční profil $\mathbf{a} = (a_{N_1}, \dots, a_{N_n}) \in \mathcal{A}$, jestliže jsou splněny následující podmínky:

1. $\forall i \in \{1, \dots, n\}\colon u_i(\mathbf{a}') \geq u_i(\mathbf{a})$,
2. $\exists i \in \{1, \dots, n\}\colon u_i(\mathbf{a}') > u_i(\mathbf{a})$.

Definice: Paretovo optimum

Uvažujme hru v normální formě $(\mathcal{N}, \mathcal{A}, u)$. Řekneme, že akční profil $\mathbf{a}^* \in \mathcal{A}$ je paretovsky optimální, jestliže neexistuje žádný akční profil $\mathbf{a}' \in \mathcal{A}$, který jej paretovsky dominuje.

Příklad: Paretovsky optimální akční profily

Uvažujme hru v normální formě vyjádřenou následující herní maticí.

N_1 \ N_2	A	B	C	D
E	6, 3	8, 2	8, 3	7, 1
F	3, 2	4, 5	6, 4	6, 5
G	4, 5	0, 8	5, 7	6, 1

Paretovsky optimální akční profily této hry jsou zvýrazněny zeleně.

Nashovo equilibrium

Best response

Hry v normální formě předpokládají **omezenou pozorovatelnost**

- hráči volí akce nezávisle na sobě,
- kdyby jeden hráč věděl, jaké akce zvolili ostatní hráči, byla by volba optimální akce triviální,
- to vede na tzv. **best response**

Definice: Best response

Uvažujme hru v normální formě $(\mathcal{N}, \mathcal{A}, u)$, akční profil $\mathbf{a} = (a_{N_1}, a_{N_2}, \dots, a_{N_n})$, jednoho konkrétního hráče $N_i \in \mathcal{N}$ a jeho utilitní funkci u_i. Označme

$$\mathbf{a}_{-i} = (a_{N_1}, \dots, a_{N_{i-1}}, a_{N_{i+1}}, \dots, a_{N_n})$$

jako redukovaný akční profil definující akce všech hráčů kromě hráče N_i.

Potom **best response** na $\mathbf{a}_{-i}$ je množina

$$BR(\mathbf{a}_{-i}) = \arg \max_{\hat{a}_{N_i} \in A_i} u_i((a_{N_1}, \dots, a_{N_{i-1}}, \hat{a}_{N_i}, a_{N_{i-1}}, \dots, a_{N_n}))$$

Nashovo equilibrium

Definice: Nashovo equilibrium

Uvažujme hru v normální formě $(\mathcal{N}, \mathcal{A}, u)$ a akční profil $\mathbf{a} = (a_{N_1}, a_{N_2}, \ldots, a_{N_n})$. Řekneme, že **a** je **Nashovo equilibrium**, jestliže

$$\forall i \in \{1, \ldots, n\} \colon a_{N_i} \in BR(\mathbf{a}_{-i}).$$

Nashovo equilibrium je takový akční profil, kde akce každého hráče představuje best response na akce hráčů ostatních

- za znalosti akcí zvolených ostatními hráči jsou všichni hráči spokojeni s akcí, kterou sami zvolili,
- equilibrium = rovnováha, stabilita; žádný hráč nechce svou strategii měnit

Příklad: Nashova equilibria

Uvažujme hru v normální formě vyjádřenou následující herní maticí.

$N_1 \backslash N_2$	A	B	C	D
E	6, 3	8, 2	8, 3	9, 8
F	7, 9	4, 5	6, 4	6, 5
G	4, 5	9, 9	5, 7	6, 1

Nashova equilibria této hry jsou zvýrazněna zlatou barvou.