

Směrem k obecnému inteligentnímu systému

Ing. Vilém Beneš

Západočeská univerzita, Fakulta aplikovaných věd,
Univerzitní 8, Plzeň 32300
shodan@kiv.zcu.cz

Abstrakt. Článek se zabývá koncepcí obecného inteligentního systému. Je zde definována inteligence ve vztahu k prostředí, navržena architektura inteligentního agenta a popsány jeho vnitřní funkční bloky a mechanismy, které se uplatňují během jeho vývoje v neznámém prostředí. Bližší pozornost je věnována procesu vnímání a definování situací.

Klíčová slova. Umělá inteligence, inteligence, agent, animat, prostředí, embodiment, anchoring, reward, reinforcement, action selection, chování, behavior, sensing, perception, vnímání, meaning, evolutionary robotics.

1 Úvod

Pokud má být nějaký systém (agent) inteligentní, musí být schopen naučit se sám efektivně chovat v okolním prostředí. To zahrnuje zjistit jaké stavy prostředí jsou pro agenta nejvíce „příjemné“. Zjistit jak může agent ovlivňovat prostředí a jak se má chovat aby prostředí dostal do cílových („nejpříjemnějších“) stavů. Aby tohoto dosáhl v komplexním prostředí, musí nejdříve prozkoumat které akce má k dispozici, jaký mají jednotlivé akce efekt a kdy je dobré tyto akce provádět v závislosti na datech, které agent přijímá svými senzory.

V tomto článku se budu zabývat koncepcí obecného inteligentního systému. V kapitole 2 budou popsány základní myšlenky – co je obecný inteligentní systém a proč se snažit o jeho vývoj; jaký je vztah prostředí a inteligence; jak by mohla vypadat architektura a vnitřní funkce inteligentního agenta.

Jednou ze základních vlastností inteligentních systémů je schopnost reagovat na stav okolního prostředí – je k tomu potřeba určitým způsobem zpracovat data ze senzorů. Kapitola 3 se zabývá tímto procesem – vnímáním, tj. přijímáním dat z prostředí a jejich zpracováním (převádění na symboly, další použití).

Závěrečná kapitola 4 nabízí celkové shrnutí a dále jsou zde uvedeny otevřené problémy a návrhy na další výzkum.

2 Obecný inteligentní systém

Vytváření systémů, které používají vlastnosti získané člověkem, je velkou pastí moderní umělé inteligence. Je potřeba, aby systém (agent), má-li být inteligentní, zvládnul sám prozkoumat okolní prostředí a zjistil jak funguje. Spojením „Obecný inteligentní systém“ označuji systém který je schopný naučit se chovat efektivně v jakémkoliv prostředí. To znamená že ho bude možné použít v široké škále aplikací. Jeho vývoj navíc poskytne vhled do procesu učení, řešení problémů, způsobu ukládání znalostí a i definování významu pro jednotlivé skutečnosti prostředí a přiřazování symbolů k nim.

V této kapitole bude nejdříve v sekci 2.1 přiblížen vztah inteligence, těla a prostředí. Dále budou v sekci 2.2 popsána různá prostředí použitelná při vývoji inteligentních agentů a faktory které ovlivňují jejich komplexitu. Sekce 2.3 popisuje otázku vnitřního fungování agenta při jeho vývoji v prostředí. Sekce 2.4 přináší návrh vnitřní architektury agenta.

2.1. Vztah prostředí-tělo-inteligence (embodiment)

Ukazuje se, že mentální procesy u lidí – jako třeba pamatování si, myšlení a porozumění řeči – jsou založeny na fyzických interakcích, které lidé mají s okolním prostředím. Kognitivní struktury se vyvíjejí díky tělu – na základě provádění akcí a přijímání dat senzory [1]. V angličtině je tato idea označována pojmem „embodiment“. Pokud máme posuzovat inteligenci, nelze brát tedy v úvahu pouze mentální procesy a izolovat je od těla a od okolního prostředí. To znamená, že inteligence není vlastností pouze mysli, ale celku mysl-tělo-prostředí.

Z této úvahy vychází moje definice inteligence: „inteligence je schopnost vyvíjet a provádět chování, které vede k vysokému ohodnocení“. Chováním je označena posloupnost akcí podmíněná stavem okolního prostředí. Ohodnocení (reward, reinforcement) je hodnota závislá na stavu prostředí, agent se snaží hodnotu ohodnocení maximalizovat prováděním akcí (manipulací se stavem prostředí). Výhodou této definice je to že je technická a tím pádem přímo použitelná při vývoji umělé inteligence. Za nevýhodu by se dalo označit to, že pokud bereme v úvahu jednoduchá prostředí, nálepku „inteligentní“ dostanou i věci, které za inteligentní normálně považovány nejsou. Je třeba si povšimnout, že když přejdeme ke komplexním prostředím (a budeme porovnávat agenty se stejným tělem), definice začne být v souladu s naším intuitivním chápáním.

Klasická umělá inteligence (GOFAI) byla založena na názoru, že je možné vytvořit inteligenci čistě jen manipulací se symboly. Pokročilý výzkum se dnes zaměřuje na vztah agenta a jeho prostředí (embodiment, nouvelle AI), na to jak je možné převést skutečnosti okolního světa na vnitřní reprezentace (transduction), či opačně - jak je význam vnitřních struktur závislý na vnějších skutečnostech (grounding meaning, anchoring) [2], [4], [5], [6]. Jak jsem již zmínil dříve, fundamentální pro utvoření použitelného obrazu světa jsou fyzické (senzorimotorické) interakce s prostředím.

2.2 Různá prostředí, faktory jejich komplexity

Inteligentní agent může fungovat v jakémkoliv prostředí – ať už jako robot na cizí planetě, meziplanetární sonda, družice analyzující Zemi, vojenské zařízení, softbot na internetu, motorové vozidlo jedoucí po silnici, robotický hráč fotbalu, diagnostický medicínský systém a jiné systémy asistující lidem v jakékoliv jejich činnosti. Zvláštní pozornost je třeba věnovat agentům ve virtuálních prostředích.

Zdá se, že nejjednodušší způsob jak umožnit agentovi vyvinout si vyšší mentální funkce, je vybrat posloupnost různých prostředí, ve kterých se tento agent bude učit vytvářet a provádět co nejefektivnější chování. Tato posloupnost by měla začínat těmi nejjednoduššími a pokračovat k více složitým. Většina nevirtuálních prostředí je příliš složitá na to, aby v nich mohl být od základu vyvinut inteligentní agent. Do komplexních prostředí je vhodné nasadit až agenta, který se umí sám učit, umí hledat zajímavé skutečnosti a definovat významy, je kreativní a umí si vytvářet jednodušší varianty problémů. Tyto schopnosti je možné získat ve virtuálních prostředích, které

mají tu výhodu, že je možné kompletně zaznamenávat vše, co se v nich děje. Je možné nastavovat všechny vlastnosti prostředí a lze zrychlit čas až na hranici kterou umožňuje výpočetní výkon použitého počítače.

Jednoduché virtuální světy jsou například abstraktní „simulace“ pohybu robotů, ve kterých prostředí sestává z 2D prostoru a ve kterém běží čas a uplatňuje se malý počet pravidel (např. omezená rychlost a oblast pohybu, ap.) a jednoduchá ohodnocující funkce. Dále je možné jako virtuální světy brát klasické hry – dáma, šachy atd., softwarová prostředí – databáze, operační systém [8], atd. Dle mého názoru jsou velmi vhodnou doménou pro vývoj umělé inteligence počítačové hry. Komplexita jejich prostředí začíná na úrovni srovnatelné s klasickou hrou člověče nezlob se – a pokračuje „spojitě“ až k velmi složitým prostředím (quake, counter-strike – 3D prostředí, navigace, pokročilá strategie při boji s lidskými hráči; flashpoint – 3D prostředí, virtuální vojenské bojiště; ultima on-line, everquest – virtuální kopie skutečného světa se stovkami lidských hráčů) [7], [9], [10].

Faktory komplexity prostředí

- průběh je opakovatelný – je možný restart / jsou k dispozici akce které vracejí zpátky stav prostředí
- počet akcí, rozsah jejich parametrů, míra vzájemného ovlivňování mezi akcemi
- míra s kterou je ohodnocení ovlivňováno skrytými skutečnostmi
- počet dimenzí dat přijímaných senzory, kvalita senzorů
- vlastnosti ohodnocení – předáváno po každé akci / pouze při „splnění cíle“; jak dobře lze odhadnout stavy s nejvyšším ohodnocením?
- (fully / partially observable) – můžou senzory zachytit všechny aspekty prostředí které jsou *podstatné* pro výběr akcí?
- (deterministic / stochastic) – prostředí je deterministické když následující stav je úplně determinován aktuálním stavem a provedenou akcí
- (episodic / sequential) – epizodické prostředí sestává z atomických epizod, ve kterých agent provádí akce, epizody jsou nezávislé – následující epizoda není ovlivněna akcemi v předchozí epizodě
- (static / dynamic) – prostředí je dynamické když se může změnit během toho co agent neprovádí žádnou akci
- (discrete / continuous) – toto rozdělení může být aplikováno na stavy prostředí, na způsob jakým funguje čas a na data ze senzorů a na prováděné akce agenta
- (single / multi agent) – jsou v prostředí další agenti snažící se maximalizovat podobné/stejně ohodnocení?

Tabulka 1. Faktory ovlivňující komplexitu prostředí ve kterém agent provádí akce.

V tabulce 1 jsou uvedeny faktory, které definují komplexitu daného prostředí. Je třeba zkoumat jaký mají tyto faktory vliv na mentální procesy agenta, který se vyvíjí v příslušném prostředí.

2.3 Vnitřní funkce agenta, mechanismy, vývoj v prostředí

Vývoj agenta

Agent pracující v neznámém prostředí by měl být schopen uplatnit znalosti získané z jiných prostředí. Zároveň by měl být schopný přizpůsobit se specifickým skutečnostem v aktuálním prostředí. Je třeba aby architektura agenta byla maximálně obecná a umožnila agentovi co nejlépe poznat nové prostředí a vytvořit a provádět

chování, které mu bude co nejvíce odpovídat.

Mechanismy spolupracující při vývoji agenta

1. provádění náhodných akcí
2. častější provádění lepších akcí
3. hledání cílů (stavů s maximálním ohodnocení)
4. vytváření makroakcí (abstraktní akce)
5. zjištění efektů akcí
6. definování situací (vnímání – abstraktní senzory)
7. odvozování pravidel pro efekty akcí a situace (model)
8. plánování
9. odvození použitelných strategií chování v závislosti na situacích
10. definování vnitřních akcí a vnitřních senzorů agenta => prostředí pro kopii agenta (vzniká kontrolní hierarchie)
11. korekční mechanismy – detekce zacyklení, verifikace modelu, detekce inkonzistencí v modelu
12. mechanismy podporující učení – abstrakce, zjednodušování modelu, hledání jednodušších případů (snížení dimenzionality, sandbox, fishpool)

Tabulka 2. Mechanismy spolupracující při vývoji (vytváření mysli) agenta.

V tabulce 2 jsou uvedeny mechanismy, které by měly fungovat „v“ agentovi při jeho vývoji v prostředí. Pokud agent funguje v naprosto neznámém prostředí, nezbyvá nic jiného, než náhodně provádět akce – a uchovávat záznamy o tom co se v prostředí dělo (mechanismus 1). Po nasbírání jistého objemu statistického materiálu je možné určit které akce jsou v průměru lépe ohodnocené – a ty je zpravidla dobré vykonávat častěji (2). Všimněme si, že v tomto okamžiku ještě data ze senzorů nedávají žádný smysl. To se změní působením následujících mechanismů. Ve stavovém prostoru senzorů jsou hledány stavy do kterých je dobré se dostávat (3), zjišťuje se jaký efekt mají jednotlivé akce (co v prostředí mění) (5), jsou určovány podobné akce a jsou vytvářeny kompozitní akce (makroakce, abstraktní akce) (4), které mají větší vliv na prostředí. Jsou definovány situace. To jsou regiony ve stavovém prostoru senzorů, které mají pro agenta stejný význam (tj. jeví se agentovi stejné). (6) [12]. Ze sensorimotorických (akce-senzory-ohodnocení) interakcí se světem jsou, jako součást vnímání, odvozovány symbolické reprezentace skutečností a fungování prostředí – model (7). Pokud agent umí zjistit ve které situaci se nachází, ví jak fungují jeho akce a zná ohodnocení různých stavů prostředí – může začít plánovat – hledat sekvence akcí, které mu umožní dostat prostředí do cílových stavů (8). Poté co plánování úspěšně funguje je možné zjišťovat jaké obecné vlastnosti má správné chování v určitých situacích – a odvodit tak pravidla pro efektivní (reaktivní) chování. Tato pravidla sniží náročnost plánování (9). Vnitřní „prostředí“ agenta (tj. jaké funkce používá, jaké jsou jejich parametry, řízení vnitřních ohodnocovacích funkcí, aj.) je možné označit za nové prostředí ve kterém je možné nasadit novou kopii agenta – tento agent tedy bude pracovat v jakémsi abstraktním světě a bude řídit původního agenta – tím se vytvoří kontrolní hierarchie s kladnou zpětnou vazbou (obecné způsoby řešení problémů objevené na nízké úrovni se uplatní i ve vyšších úrovních mysli agenta) (10). V tabulce 2 jsou dále ještě uvedeny mechanismy které sledují vnitřní reprezentace a funkce agenta a předcházejí chybám (11) a podporují učení (12).

Model

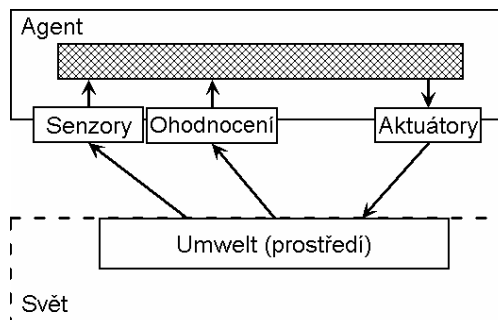
Vrcholným cílem agenta je vytvořit si funkční model okolního světa. Tento model mu umožňuje předvídat budoucí vývoj, předvídat jaký budou mít dopad akce a odhadovat kde ve stavovém prostoru dat ze senzorů se nacházejí místa s vysokým ohodnocením a jak je možné se do nich dostat. Model světa umožňuje agentovi plánovat a produkovat efektivní chování. Způsob vytvoření modelu světa musí být maximálně obecný – aby bylo možné postihnout jakékoliv skutečnosti (objekty, procesy, jevy, vztahy) v okolním prostředí. Jak jsem zmiňoval v sekci 2.1, veškeré vnitřní reprezentace agenta i celý proces vnímání závisí na akcích, které agent může provádět. Je tedy třeba navrhnout architekturu, ve které budou fungovat procesy, které vytvoří model světa na základě interakcí (akce-ohodnocení-senzor) agenta s prostředím.

Vnímání

Ve většině „použitelných“ aplikací je stavový prostor senzorů extrémně veliký (více dimenzí, spojitý), podobně tak i prostor všech možných akcí (více různých akcí, některé se spojitými parametry). Aby bylo možné „se vyznat“ v tomto množství různých možností je třeba v těchto prostorech ohraničit oblasti, které jsou si podobné (stejně). U akcí ty oblasti, které mají podobný (stejný¹) efekt v prostředí, u dat ze senzorů ty oblasti, které mají pro agenta stejný význam (pro všechny stavy senzorů z této oblasti agent bude vybírat stejnou akci). Příkladem akcí s podobným efektem může být třeba: 1. otočit se o 90 stupňů doleva; 2. otočit se o 90.002 stupňů doleva; 3. otočit se o 270 stupňů doprava – tyto akce převedou prostředí do tří stavů, které jsou pro agenta stejné ...v případě, že pro něj nemá význam uvažovat potřebný čas ani průběh otáčení a malé odchylky. O tom jak vymezit stejné oblasti ve stavových prostorech se budu zabývat v kapitole 3 – Vnímání.

2.4 Architektura obecného inteligentního systému

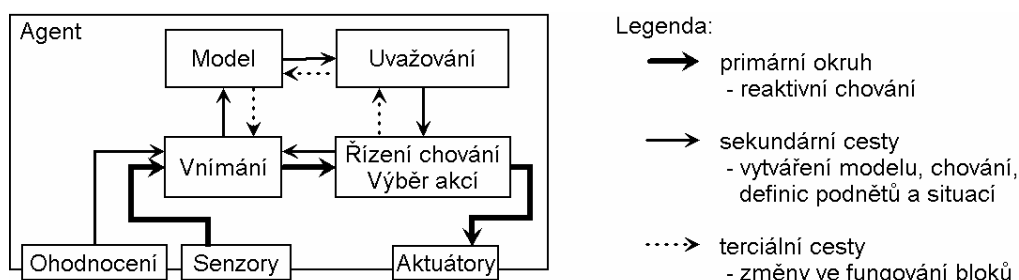
Inteligentní agent by měl být schopný přizpůsobit se jakémukoliv prostředí. To znamená, že jeho vnitřní architektura by měla umožnit co nejobecnější způsob zpracování vstupních dat a co největší volnost ve vytvářeném a prováděném chování. Tak, aby „v“ agentovi mohly fungovat mechanismy tak, jak byly popsány v sekci 2.3 a tabulce 2.



Obrázek 1. Agent a jeho vztah k okolnímu světu. „Umwelt“ označuje část světa která má vliv na agenta. Vnitřní funkční celky agenta jsou znázorněny jako black-box.

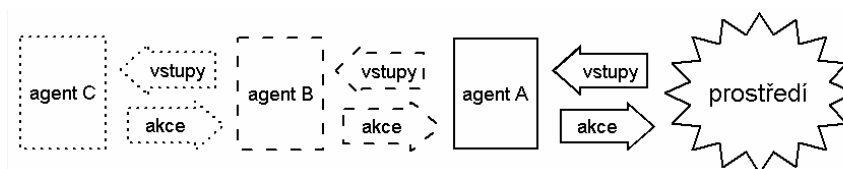
¹ Když dvě akce mají „stejný efekt“, znamená to, že tyto dvě akce převedou prostředí do stavů, které bude agent vnímat jako stejné.

Základním typem chování je chování reaktivní – na podnět, který je vnímán, agent přímo reaguje nějakým chováním. Propojení bloků vnímání a řízení chování umožňuje přímou reakci na podnět. To jak je podnět definován (jak mají vypadat data senzorů, aby daly vzniknout tomuto podnětu) se vyvíjí během fungování agenta v prostředí, tato definice je dána strukturou a funkcí bloku vnímání. Reaktivní chování mohou být odvozena přímo z minulých sekvencí akcí, prováděných agentem a jejich ohodnocení.



Obrázek 2. Vnitřní funkční bloky architektury inteligentního agenta.

Pro vyvinutí sofistikovanějších chování používá agent další funkční bloky – model (world model) a uvažování (deliberative mechanisms) – viz obrázek 2. Součástí modelu je současný stav prostředí (tak jak je vnímán), pravidla popisující jak prostředí funguje (včetně toho jak funguje a jaký je vliv samotného agenta), historie provedených akcí, získávaného ohodnocení a dat ze senzorů. Blok uvažování pracuje s modelem a snaží se najít cesty do cílových stavů (problem solving, planning, scheduling) a nová efektivní chování, která je dále možno převést na reaktivní. Ve všech funkčních blocích architektury probíhá učení. Více o strukturách inteligentních agentů viz [11].



Obrázek 3. Hierarchie řízení. Agent A (A-brain [14]) operuje v prostředí. Vnitřní procesy agenta A jsou prostředím, ve kterém operuje agent B (B-brain).

K tomu aby bylo možné vytvořit jednoduše agenta, který bude disponovat vyššími mentálními funkcemi, navrhuji využít kontrolní hierarchii, která bude sestávat z více kopií původního agenta – viz obrázek 3 a mechanismus 10 z tabulky 2. Vnější prostředí agenta B je „vnitřní prostředí“ agenta A – tedy vlastnosti, parametry, úspěšnost jeho vnitřních procesů.

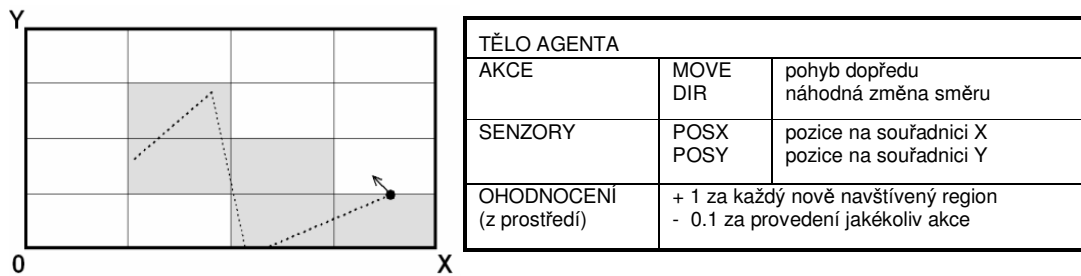
3 Vnímání

Vnímání je proces získávání dat z okolí a jejich zpracování. Vzhledem k tomu, že agent musí být schopný adaptovat se na jakékoliv prostředí, je nutné aby jeho vnímání byl maximálně obecný proces. V prostoru všech možných stavů dat ze senzorů je vhodné

definovat oblasti, které jsou nějakým způsobem pro agenta zajímavé (nebezpečí, vysoce ohodnocený stav, výskyt objektu, nacházíme se v situaci kde je výhodné zvolit netradiční chování, aj.). V průběhu práce agenta v prostředí je pak zjišťováno, jestli se data ze senzorů nacházejí v zajímavých oblastech. Primární jsou pro definování oblastí tyto věci: ohodnocení (vysoké / nízké ohodnocení; případně průměrné / extrémní) a akce (obvyklý / neobvyklý efekt; popřípadě – lze / nelze provést). Pokud se podaří najít zajímavou oblast – říkám, že byla definována situace. Situaci tvoří rozdělení stavového prostoru na oblast(i) a zbytek stavového prostoru.

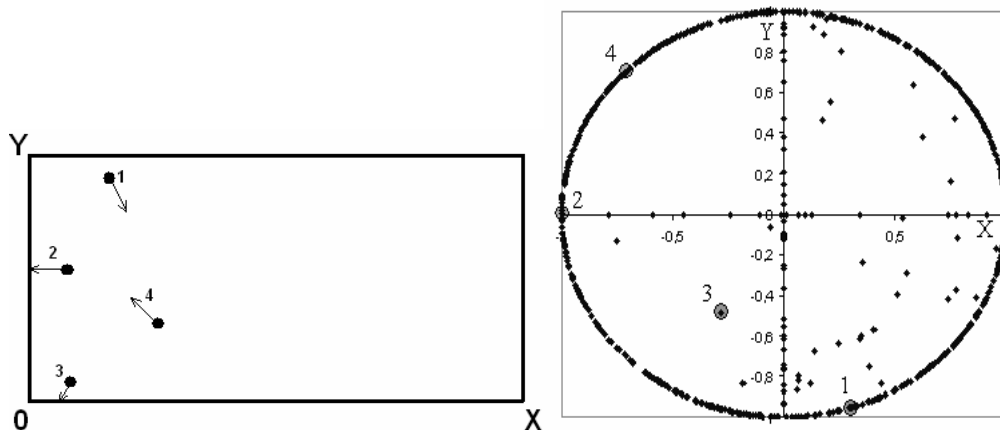
Poté, co agent definoval různé situace v datech ze senzorů, tedy v tom jak „vidí“ okolní prostředí, může začít zjišťovat jak prostředí funguje (vytvářet model) – jak je možné mezi jednotlivými situacemi přecházet, jaké situace znamenají vysoké ohodnocení pro agenta a jaké je nejefektivnější chování v závislosti na situaci v jaké se nachází.

Popsané principy definování situací budu ilustrovat na jednoduchém příkladu prostředí, ve kterém se pohybuje agent. Jedná se o 2D obdélníkovou plochu, rozdělenou na menší regiony. Účel agenta je definován ohodnocovací funkcí, která má vysokou kladnou hodnotu za navštívení dlouho nenavštíveného regionu a nízkou zápornou při provedení každé akce. Popsán bude případ definování nové situace po zjištění neobvyklého efektu akce.



Obrázek 3. Vlevo: Virtuální prostředí. Vpravo: Virtuální tělo agenta.

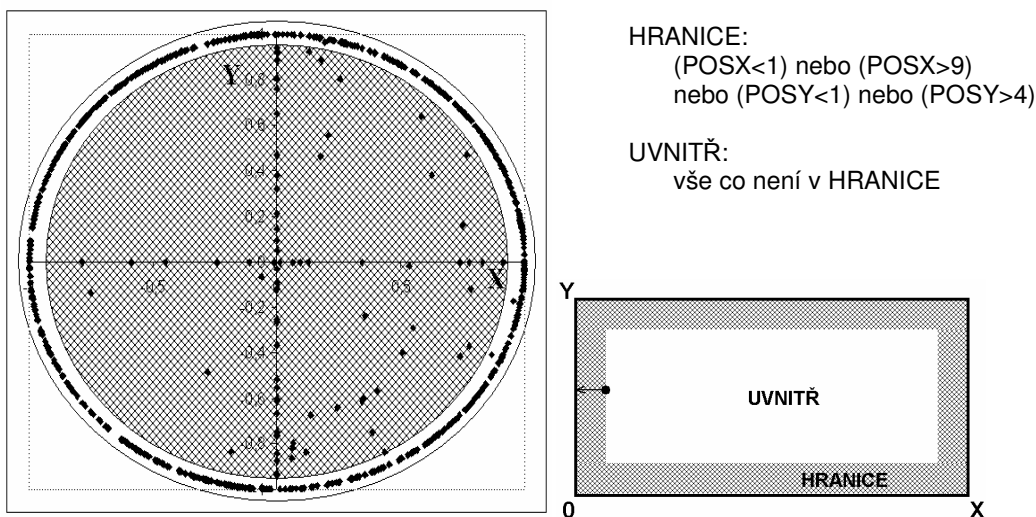
Na obrázku 3 vlevo je znázorněno virtuální prostředí, vpravo je popsáno tělo agenta. Intuitivně chápeme, že maximálně efektivní chování agenta bude spočívat v opakování akce MOVE za občasného provedení akce DIR – aby agent přešel dlouhému setrvávání na hranici obdélníkové oblasti. Co bude agent v takovém prostředí vnímat?



Obrázek 4. Vlevo: Čtyři příklady provedení akce MOVE. Vpravo: Efekt akce MOVE jako změna senzorů PSX a PSY (5000 případů), čtyři příklady z levého obrázku jsou zvýrazněny.

Agent provádí náhodně akce a zkoumá jaký mají vliv na data ze senzorů a na jejich změny. Na obrázku 4 jsou znázorněny příklady pohybu agenta a graf změn senzorů polohy po provedení akce MOVE. Pokud agent nenarazí během provádění akce MOVE na okraj, jeho poloha se změní o 1 (body tvořící kružnici na obrázku vpravo).

Agent použije shlukování (clustering) efektů akcí pro zjištění obvyklého / neobvyklého efektu akce MOVE a definování nové situace. Je třeba zjistit do jakých oblastí spadá v poměru větší množství případů. To se provádí grafovým shlukováním – shlukováním podle frekvence výskytu. Nejdříve je vytvořen použitím Delaunayovy triangulace graf. Je spočítána průměrná délka hran a hrany kratší než průměrná délka hrany jsou z grafu odstraněny. Vzniklé oblasti v grafu jsou obarveny a je nalezena oblast ve které je frekvence výskytu případů nejvyšší. Na obrázku 5 nahoře bude oblast nejvyšší frekvence mezikružící zahrnující případy na kružnici s poloměrem 1 a blízké okolí – to je obvyklý efekt akce MOVE, zbytek stavového prostoru je efekt neobvyklý.



HRANICE:
 $(POSX < 1)$ nebo $(POSX > 9)$
 nebo $(POSY < 1)$ nebo $(POSY > 4)$

UVNITŘ:
 vše co není v HRANICE

Obrázek 5. Definování nové situace. Vlevo: Efekt akce MOVE rozdělen na obvyklé a neobvyklé případy. Vpravo dole: Zobrazení stavů senzorů PSX a PSY pro které nastává neobvyklý efekt (místo pojmenování „HRANICE“ a „UVNITŘ“ agent použije pro člověka nesmyslné označení). Vpravo nahoře: odvozená pravidla pro místa kde mívá akce MOVE neobvyklý efekt.

Nyní agent může začít zkoumat co způsobí že efekt akce MOVE je obvyklý / neobvyklý. Po analýze je zjištěno, že akce MOVE má neobvyklý efekt pouze v určitých místech vnějšího prostředí. Použitím dalšího shlukování je možné určit kdy se agent v těchto oblastech nachází. Dále je možné odvodit pravidla která tyto oblasti popisují (viz obrázek 5 vpravo nahoře).

Akce MOVE má vliv na ohodnocení, byl zkoumán její efekt a definována nová situace UVNITŘ / HRANICE. Agent nyní pozná jestli je v MOVE-situaci uvnitř oblasti nebo na hranici – a může pro tyto dva případy vyvinout odlišná chování. Výsledné kompozitní chování přináší lepší ohodnocení než chování které neuvažuje definovanou situaci.

4 Další výzkum

V tabulce 2 jsou uvedeny nezodpovězené otázky, které vyvstávají z našeho přístupu, který uvažuje inteligenci pouze ve vztahu k prostředí a který uvažuje jako nutnost nevytvářet agenta ručně, ale nechat ho vyvinout se pouze na základě prováděných akcí, přijímaných dat ze senzorů a přijímané ohodnocovací funkce.

Otevřené problémy

- jakým způsobem vytvářet vnitřní ohodnocovací funkce? – některá prostředí předávají ohodnocení pouze sporadicky, toto ohodnocení si agent musí sám rozložit na provedené akce, aby mohl dále volit nejefektivnější chování
- jak zjistit efekty akcí, které se projeví ve (velmi) vzdálené budoucnosti?
- jak má vypadat architektura a vnitřní funkce inteligentního agenta?
- jaká je ideální sekvence prostředí pro vývoj agenta?
- které dovednosti a principy (statistika, neuronové sítě, algebra, geometrie, pravidlové systémy, plánování, shlukování) jsou nezbytné pro inteligenci?
- jaké vlastnosti mají metaznalosti? – tj. jaké jsou principy získávání znalostí? a jaké znalosti jsou „přenositelné“ mezi prostředími?
- lze najít zřejmou souvislost mezi rozdíly v aktuátorech a rozdíly ve vnímání lidí a (virtuálních) agentů? ...tj. jak rozdíly v aktuátorech ovlivňují rozdíly ve vnímání

Tabulka 2. Otevřené problémy.

Pro další postup ve vývoji inteligentních agentů navrhuji vytvářet další virtuální prostředí, nechat v nich vyvíjet se agenty s odlišnými architekturami a vnitřním fungováním a srovnávat jak vývoj jednotlivých agentů probíhá v závislosti na architektuře, tělu a prostředích. Dále pak hledat mechanismy, které umožňují další vývoj inteligentních agentů a které podporují jejich efektivitu v prostředí.

Získaný vhled umožní vytvořit agenta, který bude efektivně pracovat ve složitém prostředí, bude schopný učit se, správně reagovat na nezvyklé situace a dále bude možné, díky pochopení mechanismů myšlení agentů a i lidí, rozšířit obohacující propojení nás lidí a námi vytvořených umělých artefaktů.

5 Závěr

Tento článek přináší nový pohled na vytváření inteligentních agentů. Vycházel jsem z představy s anglickým názvem „embodiment“ – názoru, že kognitivní funkce, vnitřní reprezentace a i inteligence mají základ v senzomotorických interakcích s prostředím. Prezentoval jsem taxonomii prostředí, architekturu inteligentního agenta, mechanismy fungující při jeho vývoji a detailněji jsem popsal určité aspekty vnímání, které jsou odvozeny z možných akcí agenta a jsou primárním způsobem jak dát skutečnostem v prostředí význam.

Literatura

- [1] Pecher, D., Zwaan, R. A.: *Grounding Cognition*. Cambridge University Press, ISBN-10: 0521834643.
- [2] Brooks, R. A.: *Achieving artificial intelligence through building robots*. Technical Report A.I. Memo 899, Massachusetts Institute of Technology, May 1986.
- [3] Brooks, R. A.: *Intelligence without reason*. Proceedings of the 12th International Joint Conference on AI (IJCAI-91) (Sydney, Australia) (John Myopoulos and Ray Reiter, eds.), Morgan Kaufmann publishers Inc.: San Mateo, CA, USA, 1991, pp. 569-595.
- [4] Dautenhahn, K.: *Embodied Cognition in Animals and Artifacts*. In M. Mataric (ed), *Embodied Cognition and Action: Papers from the 1996 AAAI Fall Symposium*. AAAI Technical Report FS--96--02. Menlo Park, CA: The AAAI Press, 27--32.
- [5] Quick, T., Dautenhahn, K., Nehaniv, C., and Roberts, G.: *The essence of embodiment: A framework for understanding and exploiting structural coupling between system and environment*. 3rd Int. Conf. Comp. Anticipatory Syst. Liege, Belgium (1999).
- [6] Coradeschi, S., Saffiotti, A.: *Anchoring symbols to sensor data: preliminary report*. In Proc. of the 17th AAAI Conf., pages 129–135, Menlo Park, CA, 2000. AAAI Press.
- [7] Laird, J. E., van Lent, M.: *Human-level AI's killer application: Interactive computer games*. *Computer*, **34**(7) (2001) 70–75.
- [8] Etzioni, O.: *Intelligence without robots: A reply to brooks*. *AI Magazine*, vol. 14, no. 4, pp. 7 -- 13, Winter 1993.
- [9] Leonard T.: *Building an ai sensory system: Examining the design of thief: The dark project*. In Game Development Convergence (GDC 2003), 2003.
- [10] Woodcock, S., Laird, J. E., Pottinger, D.: *Game AI: The state of the industry*. *Game Developer Magazine*, August, 2000.
- [11] Franklin S.: *Autonomous Agents as Embodied AI*. *Cybernetics and Systems*, special issue on Epistemological Issues in Embodied AI, 28:6 (1997) 499-520.
- [12] King G., Oates T.: *The Importance of Being Discrete: Learning Classes of Actions and Outcomes through Interaction*. *AI '01: Proceedings of the 14th Biennial Conference of the Canadian Society on Computational Studies of Intelligence*, 2001.
- [13] <http://www.cogsci.rpi.edu/~rsun/embodied.html>
- [14] Minsky M.: *Society of Mind*. Simon & Schuster, 1985.