

Robot založený na ukotveném situovaném modelu

Michal Vavrečka,^{*}† Karla Štěpánová*, Jaroslav Vítků*

*: České Vysoké Učení v Praze, Fakulta Elektrotechnická
Katedra kybernetiky
Technická 2, 166 27, Praha 2

†: Jihočeská Univerzita, Pedagogická fakulta
Katedra pedagogiky a psychologie
Branišovská 31a, České Budějovice

Email: vavrecka@fel.cvut.cz, karlastepanova@gmail.com, vitkujar@fel.cvut.cz

Abstrakt

V příspěvku prezentujeme návrh robotického systému, který vychází z původního softwaru robota Ripley obsahujícího ukotvený model světa. Originální kód robota je postupně převeden do recentní verze operačního systému a standartizovaného robotického komunikačního protokolu (ROS). Zároveň jsme vytvořili novou verzi hardwaru robotického ramene založeného na modifikovatelné sadě motorů, pro snadnější přizpůsobení robota aktuálním podmínkám. Po úspěšné rekompilaci původního softwaru a jejím zprovoznění na novém rameni plánujeme rozšířit robotické moduly o soustavu neuronových sítí, které zajistí adaptabilitu během učení se novým podmínkám. Hlavním cílem projektu je tvorba modulů, které umožní prediktivní schopnosti robota, související s rozšířením ukotveného modelu světa o simulace možných stavů.

1 Úvod

1.1 Kognitivní robotika

Rozvoj robotiky v posledních letech koresponduje s rozvojem v oblasti výpočetní techniky a pokročilých materiálů. Rostoucí výpočetní výkon současných počítačů umožňuje rychlejší zpracování senzorických dat, což je základním požadavkem při tvorbě robotických systémů, a použití nových materiálů při tvorbě robotických těl zlepšuje jejich flexibilitu a funkčnost. Je proto žádoucí, aby také vývoj v oblasti tvorby algoritmů pro ovládání složitých robotických zařízení následoval tento trend a plně využíval výše zmíněných výhod.

Automatizované robotické systémy jsou dnes již samozřejmostí v mnoha průmyslových oblastech, kde nahrazují lidi při vykonávání mechanických pracovních činností, popřípadě pomáhají při nebezpečných úkonech. Jelikož se jedná o předem naprogramované systémy, které dokáží pouze deterministicky plnit fixní úlohu, vede každá změna či nepředvídatelný jev v okolním světě k jejich selhání a neschopnosti pokračovat v činnosti. Směrem, který se snaží tyto obtíže překonat, se v poslední

době stává kognitivní robotika. Jedná se o odvětví, které klade, oproti průmyslové robotice, důraz na tvorbu řídicích mechanismů, které z velké části napodobují způsob zpracování a ukládání senzorických informací u člověka. Z tohoto důvodu se také tvar robota většinou podobá lidskému tělu. Hovoříme pak o humanoidních robotech, vybavených patřičnou kognitivní architekturou.

Oblast kognitivní robotiky se těší v posledních letech zvýšenému zájmu a roste počet humanoidních robotů, kteří jsou vhodní pro výzkumné účely i nasazení v běžném životě. Detailním rozoborem cen a funkcí humanoidních robotů se zabýváme v loňském příspěvku (Plaček a spol., 2014). Mezi špičkové humanoidy patří v Americe robot PR-2 a v Evropě je to robotické dítě iCub. Obě zařízení jsou ale natolik drahá (asi 300.000 Eur), že si je mohou dovolit pouze vybrané robotické laboratoře. V letošním roce se objevilo několik projektů, které nabízejí humanoidní roboty v cenové hladině okolo 10.000 Eur (např. Pepper), které umožňují výzkum v této oblasti také ostatním laboratorím. Přestože existuje snaha vytvořit standardizovaného humanoidního robota pro výzkumné účely, díky cenovým omezením převládá v této komunitě spíše přístup založený na tvorbě vlastních robotických systémů, dle požadavků výzkumníků. V tomto duchu je vytvořeno i robotické rameno prezentované v našem příspěvku.

1.2 Ukotvení symbolů

Jak již bylo zmíněno výše, cílem aktuálně vytvářených kognitivních architektur je snaha o napodobení lidských kognitivních schopností. Mezi nejobtížnější úlohy v této oblasti patří způsob tvorby reprezentace znalostí ze senzorických dat, jejich použití při predikci budoucích stavů, popřípadě logických inferencí nových znalostí ze starých. V případě reprezentace znalostí převládají v současné době přístupy, které oproti tradičním přístupům založeným na symbolické reprezentaci spojené s GOFAI, využívají více úrovní (např. symbolickou a subsymbolickou) pro zachycení významu těchto znalostí na hlubší sémantické úrovni. Tento přístup vychází z článku z roku

1990 (Harnad, 1990), ve kterém je otázka sémantické reprezentace znalostí prezentována jako problém ukotvení symbolů. Základní otázkou ukotvení je, jak vytvořit funkci a interní mechanismus artefaktu (např. jako interní reprezentaci odkazující na externí objekt) tak, aby byla vlastní vzhledem k samotnému artefaktu, místo toho, aby byla závislá na interpretaci externím pozorovatelem (Ziemke, 1999). Proto je nutné, aby znalost vstupovala do systému z jeho prostředí v ukotvené formě. Na základě těchto požadavků byly vytvořeny základní architektury schopné pomocí zmíněného principu reprezentovat prostorové vztahy (Regier, 1995), vlastnosti objektů (Cangelosi a spol., 2002) popřípadě celé věty (Vavrečka a Farkaš, 2014).

1.3 Model světa

Při splnění požadavku ukotvených reprezentací je dalším požadavkem na kognitivní architekturu schopnost operovat s těmito reprezentacemi a vytvářet závěry a predikovat možné stavy světa. V oblasti umělé inteligence se pro tento typ struktur vžil pojem agenti (roboti) obsahující model světa. Oproti čistě symbolickému přístupu k reprezentaci znalostí, ve kterém se pravidla pro manipulaci s reprezentacemi a tvorby modelu světa řídí na apriorně definovanými pravidly, v případě ukotvených reprezentací již potřebujeme složitější modely, které dokáží zohlednit sémantickou úroveň reprezentací. S teoretickým přístupem k řešení tohoto problému přišel (Barsalou, 1999), který navrhuje mentální simulátory, což jsou interní mechanismy pro práci s ukotvenými reprezentacemi, které umožňují predikovat možné stavy světa. Obtížná formalizace této teorie a nevyjasněnost některých teoretických pojmů se odráží v malém množství robotických systémů, které by dokázaly implementovat ukotvený model světa. Jedním z mála řešení v této oblasti je robotické rameno Ripley (Mavridis, 2007), které obsahuje ukotvený situovaný model světa. sloužící k vyvozování stavů okolního světa mimo robotův senzorický horizont. Právě tento model posloužil jako základ pro rozvoj naší architektury, kterou detailně popisujeme v následujícím textu.

2 Architektura Ripley

Robot Ripley byl postupně vyvinut na MIT jako systém schopný rozpoznat scénu a hovořit o ní v přirozeném jazyce. Robot zde sloužil především pro výzkum v oblastech:

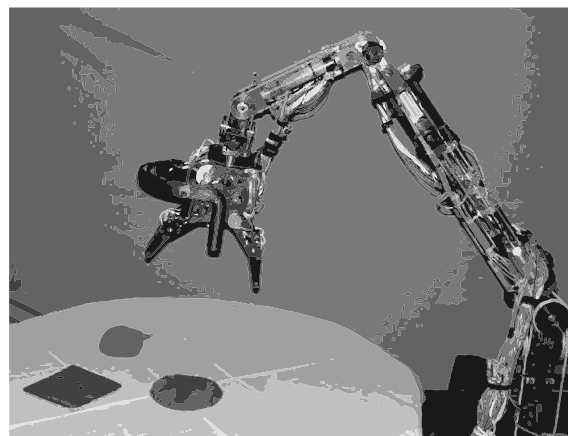
- rozpoznání a interpretace scény
- vnitřní reprezentace světa
- problému ukotvení symbolů
- zpracování přirozeného jazyka

V neposlední řadě sloužil i pro prezentační účely. Robot byl schopen výše popsané schopnosti integrovat

do jednoho funkčního celku. Dokázal rozpoznat scénu a udržovat si její reprezentaci v paměti (Mavridis, 2007). Dále pak konverzovat o scéně v mluvené řeči s člověkem, tj. odpovídat na otázky a plnit jednodušší příkazy typu: *"bring me the red ball"*. Ilustrační videa jsou k dispozici například online¹. V následujících odstavcích zevrubně popíšeme hardware, celkovou architekturu Ripley robota a vybrané části architektury.

2.1 Popis hardware

Původní robot měl formu 6-DOF manipulátoru s gripperem, viz Obr. 1. Kamera byla umístěna přímo na gripperu, což umožňovalo změnu perspektivy pohledu na scénu. Motory robota a byly schopny měřit sílu aplikovanou na táhla ovládající robotovo tělo. Ripley měl schopnost rozlišovat předměty podle hmotnosti. Navíc byl vybaven tzv. "compliant módem", ve kterém motory pouze kompenzovaly gravitaci a robot se učil motorická primitiva od učitele.



Obr. 1: Původní hardware robota Ripley. Klouby v základně, lokti a na hlavě dovolují robotu pohybovat se nad scénou. Kamera je umístěna na gripperu (chapadle), které umožňuje učení s učitelem v podobě přímé manipulace s robotem. Robot má celkem 6 stupňů volnosti. Motory umístěné v základně ovládají rameno pomocí táhel a jsou schopny měřit aktuálně aplikovanou sílu (převzato z (Hsiao a spol., 2003)).

Mezi základní schopnosti robota patří:

- vidění: předzpracování obrazu, segmentace obrazu na objekty dle tvaru, barvy a pozice; perzistence objektů apod.
- ostatní senzory: schopnost vnímat svoji konfiguraci (polohu), hmotnost objektů a taktilní informaci (např. zda robot drží objekt)
- komunikace v přirozeném jazyce: robot dále obsahuje moduly pro zpracování a syntézu řeči

¹Ripley the Robot online: <http://goo.gl/mCiRG0>

- motorika: robot je ručně naučen set základních motorických primitiv ("jak se dostat na pozici X,Y,Z") pomocí učitele. Robot je v "compliant módu" přímo naváděn za rukojeť na danou pozici. Dopředná kinematika robota je pak řešena pomocí interpolace mezi vhodnými, takto naučenými konfiguracemi kloubů. Více například v (Roy a spol., 2003).

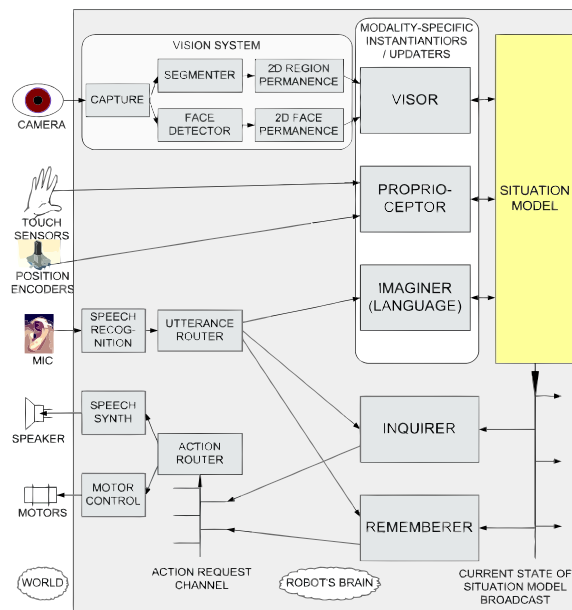
Získaná sensorická data, která jsou takto předzpracovaná, se dále ukládají v ukotveném situovaném modelu, který je popsán níže.

Modulární software architektura robota byla implementována především v jazyce C++ (některé moduly v Javě). Kvůli velké výpočetní náročnosti (především zpracování obrazu) bylo pro spuštění celé architektury potřeba několik počítačů, včetně dvou rack serverů s procesory Intel Xeon. Potřebná komunikace je v architektuře realizována pomocí Parallel Virtual Machine (PVM)², umožňující přenos serializovaných dat mezi moduly a počítači po síti.

2.2 Popis modulů a princip funkce

Kognitivní architektura robota Ripley se skládá ze specializovaných modulů, které zpracovávají data ze sensorických modalit a integrují je do ukotveného kognitivního modelu. Detailní schéma propojení je zobrazeno na Obr. 2. Vstupy z kamerových senzorů jsou zpracovány v samostatném systému pro vidění (*Vision System*). Ten se skládá z modulů pro zachycení aktuálního snímku (*Capture*), následuje detekce tváře (*Face Detector*) v případě interakce s člověkem, nebo ohraničení objektu ve scéně (*Segmenter*). Tyto moduly předzpracovávají vstupní data do ukotveného modelu světa, resp. jeho stochastické vrstvy *Visor*. Auditivní vstupy jsou zpracovávány v modulu *Speech recognition*, kde dochází k jeho převodu na text a odeslání do *Utterance routeru*. Odtud je jazykový vstup (text) poslán do kategorické vrstvy *Imaginer*, která je, podobně jako stochastická vrstva, součástí ukotveného situovaného modelu. Poslední modalitou je motorický vstup a výstup, sloužící k ovládání jednotlivých kloubů robota. Propriocepční a taktilní systémy nemají specializované moduly pro předzpracování a jsou napojeny přímo na situovaný ukotvený model. Ten přijímá informace ze senzorů, což slouží k vytvoření představy o aktuální poloze robota v prostoru a uchopeném předmětu. Konkrétně se jedná o informaci z tlakových senzorů, umístěných na robotickém chapadle, a senzorů polohy umístěných v kloubech. V ukotveném modelu je za sensorické vstupy z těla zodpovědný modul *Proprioceptor*, který odpovídá spojitě reprezentací vrstvě. Za vykonávání akcí je zodpovědný modul *Inquirer*, který je přímo napojený na jazykový vstup a situovaný model. Poté, co zpracuje požadavek (příkaz nebo dotaz), provede ve spolupráci s ukotveným kognitivním modelem jeho zpracování. Následně je výstup ve formě odpovědi

zaslán do *Action routeru*, kde je převeden na motorický výstup (*Motor Control*), sloužící k ovládání robotického ramene, nebo jazykový výstup (*Speech synth*). Pro uchování minulých stavů a akcí slouží modul *Rememberer*, který zastupuje robotickou paměť. Jsou v ní uchovány předchozí dotazy a konverzace a také minulé stavy ukotveného modelu, takže je robot schopen odpovídat také na otázky týkající se jeho předchozích akcí, popřípadě popsat situace v minulosti. Modul *Remember* je navíc vybaven elementární formou epizodické paměti, takže je schopný oddělovat jednotlivé posloupnosti akcí do samostatných celků.



Obr. 2: Celkové schéma původní SW architektury. Základem je situační model uchovávající představu o stavu světa. Model je ukotven pomocí sensorických dat z kamery, dotykových senzorů a zpracování řeči. Robot je (pomocí ramene a syntetizéru řeči) schopen sám aktualizovat nejisté/staré informace o světě. Modul pro zpracování řeči je použit buď pro aktualizaci dat v situačním modelu, nebo pro zpracování příkazů (převzato z (Mavridis, 2007)).

2.3 Ukotvený situovaný model

Srdcem celé architektury je ukotvený situovaný model, které slouží k sémantické reprezentaci okolního prostředí a umožňuje robotovi komunikaci v přirozeném jazyce, která je založena na ukotvených reprezentacích. Jak již bylo zmíněno v úvodu, pouze symbolická vrstva reprezentace znalostí je nedostatečující pro dostatečné sémantické zpracování informací z okolního prostředí a proto je potřeba vytvářet systémy, které obsahují více vrstev, ve kterých je znalost o okolním světě reprezentována v různých úrovních. Právě tyto požadavky splňuje situovaný ukotvený model. Jak vyplývá z předchozího sekce,

²PVM online: <http://www.csm.ornl.gov/pvm/>.

jsou jednotlivé senzory napojeny do specifických vrstev ukotveného modelu, konkrétně se jedná o stochastickou, spojitou a kategorickou reprezentační vrstvu. Jejich vlastnosti korespondují se symbolickou a subsymbolickou reprezentační vrstvou tak, jak byla navržena Harnadem v (Harnad, 1990), a následně použita v podobných architekturách pro ukotvení symbolů (Cangelosi a Riga, 2006; Vavrečka a Farkaš, 2014). Hlavní výhodou takových architektur je možnost reprezentovat senzory vlastnosti v jejich přirozeném formátu, neztrácíme tedy informaci o podobnosti mezi objekty či situacemi, což je typické pro symbolické reprezentace. Ukotvený model světa tedy slouží k reprezentaci objektů, jejich vlastností a vzájemných vztahů ve třech vzájemně propojených vrstvách, přičemž popisují model světa v různých úrovních detailů. Kategorická vrstva operuje v jazykovém symbolickém kódu a slouží jednoznačné identifikaci stavů světa. Na její tvorbě se podílí stochastická vrstva, která se snaží ze senzorních dat odvodit, na základě pravděpodobnosti, aktuální stavy okolního světa. Podobně funguje spojitá reprezentační vrstva. Komunikace mezi jednotlivými vrstvami zajišťuje jejich vzájemnou konzistenci. To znamená, že nové vizuální vstupy, vedou k modifikaci nejen stochastické vrstvy, ale také obou ostatních vrstev. Podobně jazykový vstup obsahující informaci o aktuálním stavu světa modifikuje subsymbolické vrstvy. Vzájemná kooperace jednotlivých vrstev vytváří ukotvený reprezentační systém, ze kterého je možné vytvářet aktuální model světa a ve spolupráci s modulem *Rememberer* také rekonstruovat předchozí stavy světa (Mavridis a Roy, 2006).

3 Náš model

Námi navržený model kognitivní architektury je nazvaný Predicon a staví na výše popsané architektuře a SW robota Ripley. Predicon přejme stávající schopnosti robota Ripley, stávající architekturu budeme dále vylepšovat a rozšiřovat o nové schopnosti/moduly.

Jednotlivé fáze projektu lze popsat následovně:

- stavba robotického ramene,
- kompilace SW originální architektury pod novým operačním systémem,
- adaptace architektury na nové robotické rameno,
- postupná integrace architektury s Robotickým Operačním Systémem (ROS),
- úprava stávajících/přidávání nových modulů do architektury podle aktuálních cílů výzkumu.

Predicon se v první fázi bude zaměřovat na vyšší adaptivitu jednotlivých částí, například sub-systémy pro vidění schopné adaptivního rozpoznání barev apod. Nová rozšíření pak budou realizována dle aktuálního směru

výzkumu, jeden z možných cílů je schopnost autonomně se učit dynamiku prostředí.

3.1 Popis rekompilace

Od Dr. Mavridise jsme získali originální zdrojové kódy Ripley architektury na jednom z původních počítačů. Originální robot je již bohužel umístěn v MIT muzeu, takže jej nelze využít. V první fázi je tedy naším cílem adaptovat původní architekturu tak, aby byla použitelná na současný OS (Linux Ubuntu) a na nový robotický HW (viz další kapitola).

SW implementace Ripleye kopíruje architekturu z Obr.2. Často tedy platí, že jeden blok odpovídá jedné spustitelné aplikaci. Aplikace mezi sebou komunikují pomocí PVM protokolu po síti.

Celá architektura byla původně spouštěna na různém počítačovém HW (PC, notebook, rack) s různými operačními systémy (Linux, RT Linux, Windows). Ke konci vývoje Ripley architektury se posledním výzkumníkům již nedařilo bez problému zkompilevat celou architekturu. Proto byly problematické části spouštěny na původních počítačích a upravovány byly jen určité části (například Situation Model).

Výsledný zdrojový kód (obsahující množství různých verzí různých podsystémů, naměřených a/nebo testovacích dat apod.) čítal úctyhodných 14.5GB dat. Původní velikost se nám postupně povedlo redukovat na 5GB dat. Zdrojové kódy jsou postupně rekompileovány pro Linux Ubuntu 12.04. Dalším hlavním problémem byla různorodost verzí jednotlivých pod-systémů a jejich vzájemná nekompatibilita (například různá struktura PVM zpráv). V době psaní příspěvku se již podařilo zkompilevat většinu SW. Zbývají pravděpodobně již jen některé části vidění (viz Obr.2 nahoře), využívající OpenCV.

Po úspěšné kompilaci všech pod-systémů otestujeme schopnost jejich vzájemné interakce a začneme SW integrovat s modernějším Robotickým Operačním Systémem (ROS)³. Jednotlivé PVM uzly (části Ripley architektury - aplikace) odpovídají ROS uzlům. Proto budeme potřebné části SW postupně převádět pod ROS, čímž docílíme kompatibility se současným robotickým HW a zefektivníme komunikaci mezi sub-systémy. V další fázi budou do ROSu převedeny především senzory a motorické sub-systémy. Námi použitý robotický HW je popsán v následující kapitole.

3.2 Popis hardware

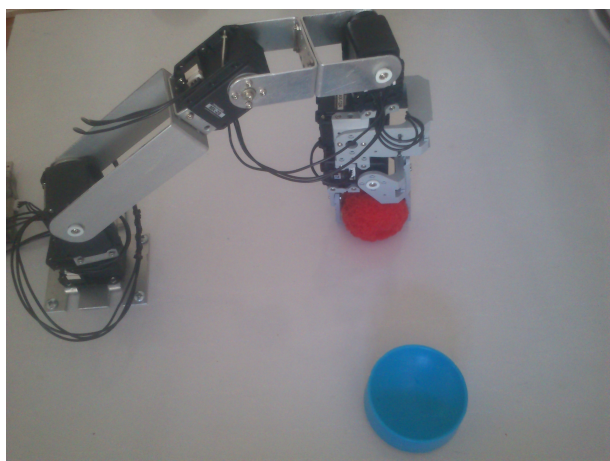
Pro sestavení robotického ramene jsme využili díly z robotické stavebnice Bioloid Premium Kit od firmy Robotis. Jedná se o stavebnici ze servo motorů, plastových dílů, senzorů a řídicí elektroniky. Základní balení obsahuje vše potřebné pro postavení humanoida se 16 stupni volnosti. Díky konstrukci z univerzálních dílů je ale možné robota

³ROS: <http://www.ros.org/core-components/>

bez větších obtíží modifikovat a přizpůsobit konkrétní situaci. Vzhledem k tomu, že servomotory obsažené v základní verzi stavebnice nebyly schopné udržet námi navržené rameno, museli jsme je vyměnit za silnější motory, pro které jsme museli následně vyrobit i odpovídající propojující díly.

Výsledné rameno (viz. Obr.3) má 6-DOF a je sestaveno z následujících modulárních servomotorů: 1x MX-64, 2x MX-28 a 3x MX-12A. Tyto servomotory sestávají z DC motoru, převodů z ozubených kol, obvodů pro řízení a obvodů pro digitální komunikaci po sběrnici Dynamixel. Všechna serva a také další periferie jsou připojena k řídicí jednotce CM530, která funguje jako digitální centrum robota. Řídicí jednotka v sobě obsahuje vlastní mikrokontrolér a pomocné obvody pro komunikaci s periferiemi a počítačem.

Jednotlivé Dynamixel motory, ze kterých je robotické rameno postaveno, jsou propojeny pod ROS za pomoci *dynamixel_motor* package, tudíž lze přímo v ROS ovládat jejich nastavení a sledovat jejich otočení, teplotu, napětí a zátěž. USB web-kamera Creative s rozlišením 720p je propojena s ROS pomocí *usb_cam* package (při rozlišení 640x480 px je frekvence přenosu cca 25 Hz a při rozlišení 1280x720 je frekvence cca 20 Hz).



Obr. 3: Nový hardware pro výzkum a vývoj architektury predikon. Robot je složen z dvou typů motorů Dynamixel a USB web-kamery Creative s rozlišením 720p. Komunikace s počítačem (Linux Ubuntu) je realizována pomocí ROS infrastruktury (Quigley a spol., 2009).

3.3 Popis úlohy

Fáze 1: Základní interakce - kolize Trénovací fáze: Robot pozoruje scénu se 4 typy základních kolizí. Na scéně se nachází jeden objekt uprostřed scény a jeden objekt, který se pohybuje směrem k němu. Robot pozoruje tuto kolizi a detekuje objekty (pozornostní modul) a jejich vlastnosti (barvu, tvar, velikost, prostorovou polohu, rychlost a směr). Tyto vlastnosti jsou reprezentovány v samostatných

modulech. Ze začátku budou tyto vlastnosti sjednoceny s proprioceptivní informací, aby bylo možné jednoduše vytvořit kanonickou scénu, která zahrnuje pozici objektů a aktuátorů ve světě. Všechny tyto vlastnosti jsou následně projektovány do dynamického modelu, který se bude učit, jak jednotlivé vlastnosti (příznaky přispívají k jednotlivým typům kolizí (typ kolize je definovaný z jazykového modulu). Cílem dynamických modulů bude naučit se rozpoznat ze senzorických příznaků o jaký typ kolize se jedná.

Testovací fáze: Po trénovací fázi následuje fáze, kdy budeme robotovi prezentovat pouze část kolize (např. do chvíle kontaktu objektů). Následně je měřena odpověď systému na jednotlivých úrovních. Z jazykového výstupu můžeme vyčíst jaký typ kolize systém rozpoznal. Motorický výstup zajistí posun ramene do takové pozice, která bude odpovídat budoucí trajektorii interagujících objektů tak, aby je bylo schopno zachytit. Je možné také pomocí vizualizačních nástrojů analyzovat vnitřní svět agenta (mentální model), který reprezentuje pozorovanou kolizi.

Fáze 2: Více interakcí, řetězení a kompozicionalita

V následující fázi budeme prezentovat jiný typ interakcí. Naší snahou bude ukotvit jednotlivá akční slovesa pomocí sady základních primitiv - např. chodit (bez kolize, bez interakce), padat (pohyb dolů, zastavení). Může se jednat i o reprezentaci více komplexních akcí - např. popis odpalu při tenisu jako sady základních pohybů.

3.4 Návrh prvotních vylepšení

Adaptivita Naším prvotním cílem je, aby byl celý systém adaptivní, tudíž budeme postupně předkládat jednotlivé moduly tak, aby nebyly hodnoty pevně zadané, ale aby reagovaly na okolní prostředí a přizpůsobovaly se aktuálnímu stavu učení. Budeme postupovat od jednodušších modulů až po moduly komplexnější, například od adaptivního modulu pro rozpoznávání barev a tvarů až po adaptivní moduly pro pozornost, paměť či řeč.

Jazyk Dalším naším cílem je výzkum akvizice jazyka. Z toho důvodu provedem postupně následující kroky:

1. pro propagační i výzkumné účely převedeme současný slovník robota do ČJ (budeme dále pracovat jak s ČJ, tak s AJ verzí)
2. Rozšíření slovní zásoby a gramatických struktur robota
3. Schopnost rozlišit druh věty (oznamovací, tázací ...) a její konkrétní typ (otázka na pozici, vlastnost, dotaz na minulý stav) nejen dle prvního slova (jak je to ve stávající architektuře)

4. Schopnost komunikace pomocí variabilní délky věty
5. Učení se nových slov přímo z jazykového vstupu nezávisle na senzorických datech
6. Učení se slovům popisujícím nejen statické objekty (podstatná jména, přídavná jména), ale i akce (slovesa - např. dotknout se, naražit, odrazit se, posunout atp.) a vztahy objektů (předložky a příslovce - např. před, nad, skrz, více atp.)

Změna reprezentace scény V aktuální verzi je si sice robot schopný vytvořit mentální model člověka a správně převést danou informaci (např. o pozici objektu - vlevo/vpravo) do své reprezentace, ale není schopný si například uvědomit, že některý z objektů může být pro člověka zakrytý. Jednou ze schopností Prediconu by mělo být takovéto situace pochopit a popsat.

Dynamické scény Postupně se chceme přesunout od statických objektů a scén k dynamickým scénám a k objektům, které spolu budou interagovat. Predicon by měl být schopný odvodit budoucí stav věcí z minulých pozorování. Například z pozorování pohybů těles bude schopný odvodit jejich budoucí pohyb a to, zda dojde ke kolizi a jaký bude výsledek jejich interakce (v závislosti na vlastnostech sledovaných objektů).

4 Závěr

Cílem našeho projektu je rozvoj kognitivní architektury založené na ukotveném situovaném modelu. Oproti přístupu, založenému na postupné vývoji systému, jsme zvolili variantu, která staví na ověřeném robotickém systému Ripley, který má podobné vlastnosti, jako navržený systém Predicon. Tento postup nám umožní zrychlit vývoj systému, jelikož jsou základní moduly již vytvořeny. Nově vytvářený systém by měl v praxi prověřit současné teorie ukotvených reprezentací, umožňující rozvoj umělé inteligence, která je schopná plnohodnotně pracovat z informacemi z okolního světa.

Poděkování

Tento projekt byl financován z grantu SGS "New approaches to collection, evaluation and utilization of biological data II", No. SGS13/203/OHK3/3T/13 a grantu RPPJU No. CZ.1.07/2.3.00/30.0049.

Literatura

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22(04):577–660.

Cangelosi, A., Greco, A. a Harnad, S. (2002). Symbol grounding and the symbolic theft hypothesis. V *Simulating the evolution of language*, str. 191–210. Springer London.

Cangelosi, A. a Riga, T. (2006). An embodied model for sensorimotor grounding and grounding transfer: Experiments with epigenetic robots. *Cognitive Science*, 30(4):673–689.

Harnad, S. (1990). The symbol grounding problem. *Physica D*, 42:335–346.

Hsiao, K., Mavridis, N. a Roy, D. (2003). Coupling perception and simulation: steps towards conversational robotics. V *Intelligent Robots and Systems, 2003. (IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on*, vol. 1, str. 928–933 vol.1.

Mavridis, N. (2007). *Grounded Situation Models for Situated Conversational Assistants*. Dizertační práce, Massachusetts Institute of Technology. Dept. of Architecture. Program In Media Arts and Sciences.

Mavridis, N. a Roy, D. (2006). Grounded situation models for robots: Where words and percepts meet. V *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, str. 4690–4697. IEEE.

Plaček, M., Vavrečka, M. a Lhotská, L. (2014). Matlab toolbox pro ovládání robota bioloid. V *In Kognitivní Věda a Umělý Život II*, str. 233–241, Zaječí u Břeclavi: Slezská univerzita v Opavě.

Quigley, M., Conley, K., Gerkey, B., Faust, J., Foote, T., Leibs, J., Wheeler, R. a Ng, A. (2009). Ros: an open-source robot operating system. V *ICRA Workshop on Open Source Software*.

Regier, T. (1995). A model of the human capacity for categorizing spatial relations.

Roy, D., Hsiao, K.-Y., Mavridis, N. a Gorniak, P. (2003). Ripley, hand me the cup! (sensorimotor representations for grounding word meaning). V *IEEE Workshop on Automatic Speech Recognition and Understanding*.

Vavrečka, M. a Farkaš, I. (2014). A multimodal connectionist architecture for unsupervised grounding of spatial language. *Cognitive Computation*, 6(1):101–112.

Ziemke, T. (1999). Rethinking grounding. V *Understanding representation in the cognitive sciences*, str. 177–190. Springer US.