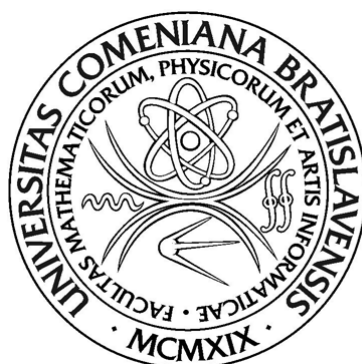


UNIVERZITA KOMENSKÉHO, BRATISLAVA
FAKULTA MATEMATIKY, FYZIKY
A INFORMATIKY
KATEDRA APLIKOVANEJ INFORMATIKY



Modelovanie zdieľanej vizuálnej
pozornosti pomocou neurónových sietí
Diplomová práca



Bratislava, 2010

Bc. Martin Bezák

UNIVERZITA KOMENSKÉHO, BRATISLAVA
FAKULTA MATEMATIKY, FYZIKY
A INFORMATIKY
KATEDRA APLIKOVANEJ INFORMATIKY

**Modelovanie zdieľanej vizuálnej
pozornosti pomocou neurónových sietí**

Diplomová práca

Bc. Martin Bezák

Študijný program: Kognitívna veda

Študijný odbor: 3.1.9 Psychológia a 9.2.8 Umelá inteligencia

Vedúci diplomovej práce:
doc. Ing. Igor Farkaš PhD.

Bratislava, 2010

Prehlásenie

Prehlasujem, že som túto diplomovú prácu vypracoval samostatne len s použitím uvedenej literatúry.

V Bratislave dňa _____

podpis

Podakovanie

Týmto ďakujem svojmu vedúcemu diplomovej práce doc. Ing. Igorovi Farkašovi PhD. za odborné vedenie, pomoc, trpezlivosť a ochotu pri spracovaní uvedenej témy.

Abstrakt

Bezák, Martin. Modelovanie zdieľanej vizuálnej pozornosti pomocou neurónových sietí [diplomová práca]. Univerzita Komenského v Bratislave. Fakulta matematiky, fyziky a informatiky; Katedra aplikovanej informatiky.

Vedúci práce doc. Ing. Igor Farkaš PhD. Bratislava: FMFI UK, 2010.

Táto práca sa zaoberá modelovaním a simuláciou zdieľanej vizuálnej pozornosti u agenta. Navrhuje model založený na učení posilňovaním pomocou CACLA algoritmu využívajúceho dopredné neurónové siete ako funkčné aproximátory. Pomocou tohoto algoritmu dokáže tento model pracovať v spojitom priestore a vykonávať v ňom spojité akcie. Model sa vie naučiť len s danou informáciou o smere pohľadu rodiča otáčať za objektami v plne trojrozmernom priestore. Vykonané testy ukazujú dobré výsledky v úspešnosti nachádzania objektov a otáčaní sa za týmito objektami v priestore.

Kľúčová slová učenie posilňovaním, neurónová sieť, zdieľaná pozornosť, Continuous Actor Critic Learning Automaton, CACLA

Abstract

Bezák, Martin. Modeling joint visual attention using neural networks. [diploma thesis]. Comenius University in Bratislava. Faculty of Mathematics, Physics and Informatics; Department of Applied Informatics.

Supervisor doc. Ing. Igor Farkaš PhD.Bratislava: FMFI UK, 2010.

This thesis deals with modeling and simulation of joint visual attention in the agent. We propose a model based on reinforcement learning by CACLA algorithm, using feedforward neural networks as function approximators. With this algorithm, the model can work in continuous and continuous action spaces. The model can learn how to face objects in a fully three-dimensional space, only by the given information about the direction of a caregivers perspective. Experimental results shows that model achieve very good results in facing objects.

Keywords reinforcement learning, neural networks, joint attention, Continuous Actor Critic Learning Automaton, CACLA

Predhovor

Modelovanie kognitívnych funkcií je jedným z možných spôsobov ako spoznávať človeka, zároveň umožňuje vytvárať napríklad interaktívnych humanoidných robotov, ktorí nám môžu poslúžiť v bežnom živote, tak aj v profesionálnej či vedeckej sfére. Navrhovaním a skúmaním rôznych implementácií a modelov s využitím skutočných experimentálnych poznatkov, niekedy odhalí prekvapivé fakty a zároveň môže vzniknúť množstvo nových špekulácií. Jednou z ľudských schopností, ktorá je pod drobnohľadom nielen kognitívnych vedcov, je schopnosť zdieľať vizuálnu pozornosť. Táto je jedným z dôležitých faktorov pri vývine rôznych komunikačných, sociálnych a kognitívnych schopností. Táto práca navrhuje výpočtový model zdieľanej vizuálnej pozornosti, postavený na učení posilňovaním s použitím neurónových sietí. Úlohou tohoto modelu je nadobudnúť schopnosť zdieľanej vizuálnej pozornosti. Pre túto úlohu bol zvolený tzv. CACLA algoritmus, ktorý pracuje s doprednými neurónovými sieťami a má vhodné predpoklady na implementáciu problému v spojitom prostredí.

Obsah

| | |
|---|-------------|
| Zoznam obrázkov | viii |
| Zoznam tabuliek | ix |
| 1 Úvod | 1 |
| 2 Zdieľaná pozornosť | 3 |
| 2.1 Základné prvky pozornosti | 3 |
| 2.2 Vývin zdieľanej pozornosti | 6 |
| 2.3 Butterworthov model vývinu zdieľanej vizuálnej pozornosti | 8 |
| 2.4 Baron-Cohenove moduly zdieľanej vizuálnej pozornosti | 10 |
| 2.5 Model zahŕňajúci štruktúry a mechanizmy podieľajúce sa na vývine zdieľanej pozornosti | 11 |
| 2.5.1 Vnemové schopnosti | 11 |
| 2.5.2 Štrukturované sociálne prostredie | 12 |
| 2.5.3 Učenie posilňovaním | 12 |
| 2.5.4 Habituácia | 13 |
| 3 Výpočtové modelovanie | 14 |
| 3.1 Umelé neurónové siete | 14 |
| 3.1.1 Viacvrstvový perceptrón | 15 |
| 3.1.2 Učenie so spätným šírením chyby | 17 |
| 3.2 Učenie posilňovaním | 19 |
| 3.2.1 Prvky učenia posilňovaním | 20 |
| 3.2.2 Markovovský rozhodovací proces | 21 |
| 3.2.3 Definícia problému učenia posilňovaním | 22 |

| | | |
|----------|--------------------------------------|-----------|
| 3.2.4 | Explorácia a exploatácia | 24 |
| 3.2.4.1 | ε -greedy | 25 |
| 3.2.4.2 | Softmax výberové pravidlo | 25 |
| 3.2.4.3 | Gaussovská explorácia | 26 |
| 3.2.5 | Diskrétne algoritmy učenia | 26 |
| 3.2.5.1 | TD učenie | 26 |
| 3.2.5.2 | Q-učenie | 26 |
| 3.2.5.3 | SARSA učenie | 27 |
| 3.2.5.4 | Aktér-Kritik učiace metódy | 28 |
| 3.2.5.5 | QV-učenie | 30 |
| 3.2.6 | Spojité algoritmy učenia | 30 |
| 3.2.6.1 | CACLA | 30 |
| 4 | Navrhovaný model | 33 |
| 4.1 | Model prostredia | 34 |
| 4.2 | Model agenta | 37 |
| 4.3 | Učenie | 39 |
| 4.4 | Zostavenie a vyhodnotenie | 41 |
| 5 | Záver | 44 |
| | Literatúra | 45 |
| | Prílohy | I |
| | Obsah CD | I |

Zoznam obrázkov

| | | |
|-----|--|----|
| 2.1 | Funkčné prvky pozornosti. (Knudsen, 2007). | 5 |
| 2.2 | Tri obdobia vývinu zdieľanej pozornosti. (Nagai a spol. 2003) . . . | 9 |
| 3.1 | Dopredná neurónová sieť. Informácia sa šíri zľava doprava. (Abazov a spol., 2009). | 16 |
| 3.2 | Príklad aktivačných funkcií.(de Smith, 2007) | 17 |
| 3.3 | Schéma interakcie agenta s prostredím (Sutton, Barto, 2004). | 23 |
| 3.4 | Typická aktér-kritik architektúra (Sutton, Barto, 2004). | 29 |
| 4.1 | Sférický súradnicový systém. (Wolfram MathWorld, [online]). . . . | 35 |
| 4.2 | Reprezentácia rodičovho pohľadu ako uhol predkladaný agentovi. . . | 36 |
| 4.3 | Zobrazenie agenta v prostredí. | 38 |
| 4.4 | Počet krokov potrebných ku nájdeniu objektov počas prvých 300 učiacich epizód. | 42 |
| 4.5 | Priemerný počet krokov pri testovaní. | 42 |
| 4.6 | Pohyb ku trom vybraným objektom. | 43 |

Zoznam tabuliek

| | | |
|-----|--|----|
| 4.1 | Tabuľka parametrov pre model. | 40 |
| 4.2 | Priemerná percentuálna úspešnosť vzniku zdieľanej vizuálnej pozornosti na nových tisíc testovaných objektoch po každom z 20 učení. . . | 41 |

Kapitola 1

Úvod

Už novorodenci sa na svet dostanú s určitými, no nie plne vyvinutými schopnosťami. Ich videnie ja rozmazané, pohyby sú nekontrolované a veľkosť pamäte je obmedzená. Počas ďalšieho vývinu sa ich schopnosti zlepšujú pomocou skúseností, ako sa hýbu, pozorujú a rozpoznávajú svet okolo seba, sa u nich rozvíjajú ich kognitívne schopnosti. Časom dokážu vyextrahovať základné charakteristiky z komplikovaného prostredia. Napríklad nájsť a vybrať určité jednoduché časti z komplexného jazykového aparátu dospelých a tak rozvíjať dôležité jazykové štruktúry. Podľa Baron-Cohena (1992) si deti ešte pred prvým rokom života začínajú uvedomovať význam pozornosti iných osôb. Pohľad na rodiča zameraného na nejaký objekt, si dieťa vie interpretovať ako podnet pre upriamenie svojej pozornosti ku danému objektu, tu sa začína prejavovať schopnosť zdieľať pozornosť. Zdieľanú pozornosť (ZP) môžeme definovať ako schopnosť jedinca upútať alebo byť upútaný pozornosťou iného alebo iných jedincov pohľadom, gestami alebo iným verbálnym či neverbálnym spôsobom ku určitému stimulu. Inak povedané zdieľaná pozornosť je súčasné zapojenie pozornosti viacerých jedincov na určitý vonkajší stimul. ZP tiež umožňuje deťom zisťovať čo je dôležité v ich prostredí na základe poznania dospelých. ZP spolu s rečou poskytujú možnosti vytvárať mentálne reprezentácie toho čo vnímajú, sú teda dôležité v pri akvizícii jazyka ako aj v komunikácií. Vo všeobecnosti môžeme rozdeliť ZP do dvoch častí a to na zahájenie a odpoveď na ZP.

Zdieľaná pozornosť má dôležitú úlohu vo vývine jazyka a je tiež dôležitá v spoločenských interakciách. Rozvinutie tejto schopnosti je dôležité pri vývine komunikačných schopností u detí (Tomasello, Farrar, 1986). Deficity v zdieľanej pozornosti

bývajú úzko spájané s autizmom (Baron-Cohen, Baldwin, Crowson, 1997).

Zároveň vznikajú snahy modelovať tieto schopnosti pre interaktívnych robotov a tiež súčasné výpočtové modely nám pomáhajú porozumieť týmto schopnostiam. Značné úsilie je vynakladané na vývoj robotov, ktorý by sa mohli učiť ako deti. Preto existuje množstvo modelov vývoja a simulácií ZP (Triesch J, Teuscher, Deák, 2006. Nagai, Hosoda, Morita, Asada, 2003. Yücel, Salah, 2009. Yucel, Salah, Mericli C., Mericli T., 2009. Doniec, Ganghua Sun, Scassellati 2006).

V tejto práci bol vytvorený nový model, ktorý simuluje vývin ZP pomocou neurónových sietí a učenia posilňovaním. Model, v ktorom je agent schopný sa naučiť otáčať za objektom a odpovedať na podnet ZP, tiež dokáže pracovať aj s veľkým množstvom objektov v okolí. Práca sa zaoberá iba vizuálnou formou zdieľanej pozornosti, teda zdieľanou vizuálnou pozornosťou.

Táto práca je rozdelená do troch častí, v nasledujúcej kapitole sa budeme zaoberať všeobecne vlastnosťami pozornosti a vývinom zdieľanej pozornosti a aké faktory a mechanizmy vplývajú na vývin ako pôsobí vývin ZP na rozvoj jazyka. V ďalšej kapitole sa budeme zaoberať výpočtovým modelovaním. Popíšeme si použité časti neurónových sietí, väčšiu pozornosť venujeme učeniu s posilňovaním. V poslednej kapitole si predstavíme navrhovaný model napodobňujúci funkciu zdieľanej vizuálnej pozornosti. Model otestujeme a vyhodnotíme jeho vlastnosti.

Kapitola 2

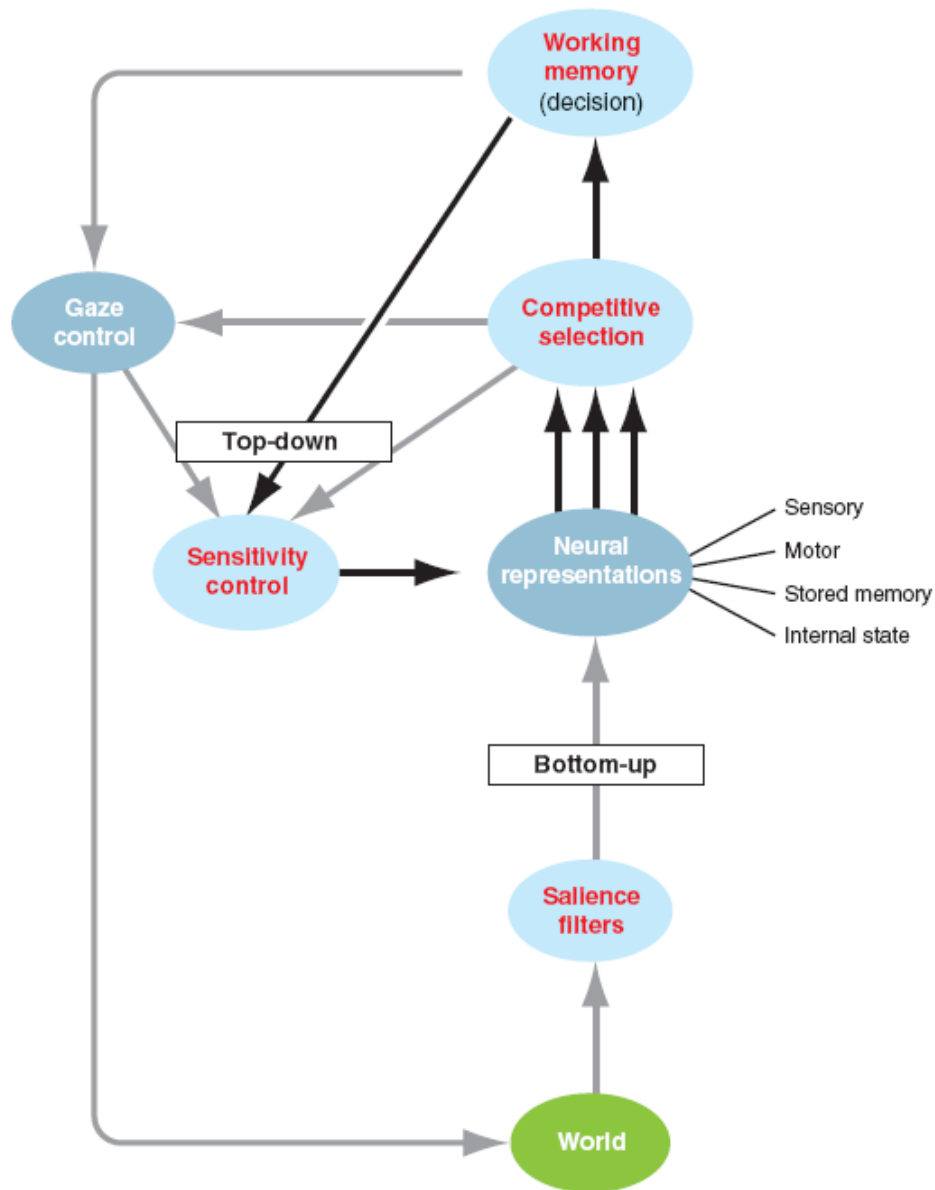
Zdieľaná pozornosť

2.1 Základné prvky pozornosti

Za pozornosť môžeme považovať stav, pri ktorom sa jedinec sústreďí na vybraný aspekt prostredia a viac menej ignoruje tie ostatné. Z pohľadu neurovied pozornosť nastáva pri zvýšenej aktivácii príslušných synchronne oscilujúcich neurónov v pozornosťných oblastiach frontálnej kôry mozgu, napríklad vplyvom zvýšenej aktivácie v sensorických oblastiach. Čiže pôsobením dostatočne silného podnetu je pozornosť k tomuto podnetu odvedená. Takáto aktivácia v sensorických oblastiach môže automaticky ovplyvňovať pozornosť, čo nazývame ako pôsobenie zdola nahor procesov. Naopak pozornosť môže byť tiež ovplyvnená plánovacími oblasťami vo frontálnej kôre, ktoré sa aktivujú pred tým, než je vykonaná vôľou ovplyvnená činnosť. Plánovanie môže ovplyvniť, ktoré sensorické neuróny sa synchronizujú, čo môžeme nazvať ako pôsobenie zhora-nadol (Rybár, Beňušková, Kvasnička, 2002).

Knudsen (2007) prezentuje svoju koncepciu kde navrhuje, že pozornosť závisí od kombinovaných príspevkov štyroch odlišných procesov ako: pracovná pamäť, kompetitívna selekcia, zhora nadol kontrola senzitivity a automatické filtrovanie pre výrazné stimuly. Pracovná pamäť je forma vysoko dynamickej krátkodobej pamäte, ktorá operuje s dočasne vybranými informáciami. Kompetitívna selekcia je proces, ktorý určuje aká informácia získa prístup do pracovnej pamäte. Zhora nadol kontrola senzitivity je proces, ktorý upravuje silu signálov rôznych informačných kanálov, ktoré súťažajú o prístup do pracovnej pamäte. Zdola nahor

pôsobiacie filtre pre výrazné stimuly automaticky zlepšujú reakcie na stimul, ktorý je zriedkavý. Súhra týchto procesov vedie k zvýšenej citlivosti neurónov a skraca je reakčné odozvy, potrebné pre vznik pozornosti. Pozornosť vyberá informáciu, ktorá získa prístup do pracovnej pamäte. Prístup do pracovnej pamäte je určený silami signálov súťažiacich reprezentácií informácie. Sila signálu je upravovaná automaticky zhora-nadol filtrami, zároveň je upravovaná zhora-nadol signálmi, ktoré kontroluje pracovná pamäť. Cielená kontrola pozornosti je sprostredkovaná rekurentnou slučkou zahŕňajúcou pracovnú pamäť, zhora-nadol kontrolu a kompetitívnu selekciu. Informácia je vyhodnotená a rozhodnutia zostávajú na pracovnej pamäti. Podľa tejto koncepcie, pozornosť neidentifikuje ciele, o to sa stará pracovná pamäť. Pozornosť do tohoto procesu ani veľmi nie je zapojená, je to len neustále pokračujúce súťaženie medzi hierarchiami spracovávajúce informácie, súperiace o prístup do pracovnej pamäti. To, čo je zapojené, sú zhora nadol signály o ktorých rozhoduje pracovná pamäť. Zhora nadol signály môžu voliteľne obohatiť reprezentácie určitej informácie tak, že informácia má stále vysokú pravdepodobnosť dostať sa do pracovnej pamäte. Pohyby očí aj s ostatnými orientačnými pohybmi sú tiež vedené rozhodnutiami v pracovnej pamäti a slúžia spolu s zhora nadol signálmi na zlepšenie rozlíšenia informácie poskytnutej pracovnej pamäti.



Obrázok 2.1: Funkčné prvky pozornosti. (Knudsen, 2007).

Na obr. 2.1 sú procesy prispievajúce k pozornosti označené červenou farbou. Informácie o svete (zelená) sú spracované zdola nahor pôsobiacimi filtermi a odpovedajú na rozdielne, neobvyklé či dôležité stimuly. Neurálne reprezentácie v rôznych hierarchiách kódujú informáciu o svete, pohyboch, emóciách, spomienkach a ďalších. Kompetitívny proces vyberá reprezentáciu s najväčšou silou signálu pre vstup do okruhu podliehajúceho pracovnej pamäti. Pracovná pamäť môže usmerniť zhora

nadol signálmi senzitivitu reprezentácií, ktoré sú spracovávané pracovnou pamäťou. Výberový proces môže tiež ovplyvniť zhora nadol signály, čo môže ovplyvniť výsledky kompetitívnej selekcie. Pracovná pamäť a kompetitívna selekcia riadia pohyby očí a iné orientačné správanie, ktoré upravujú vplyv prostredia na nervový systém. Vôľou riadená pozornosť zahŕňa pracovnú pamäť, zhora nadol kontrolu a kompetitívnu selekciu operujúce v rekurentnej slučke (čierne šípky).

Správanie, ktoré nastáva pri pozornosti, a ako sa navonok prejavuje, je možné u ľudí či zvierat sledovať, čo je nutné pre vznik ZP. Deti túto schopnosť nadobúdajú postupne už takmer od narodenia a je veľmi dôležitá pri sociálnom vývine detí. V nasledujúcej sekcii nazrieme na vývin ZP a rôzne teórie, modely a prístupy, ktoré tento fenomén opisujú alebo sa ho snažia vysvetliť.

2.2 Vývin zdieľanej pozornosti

Zdieľaná pozornosť sa objavuje už niekedy medzi šiestym a dvanástym mesiacom života dieťaťa a plne sa rozvinie okolo osemnásteho mesiaca života (Butterworth, Jarrett, 1991). Niektoré pozorovania ukazujú, že niektoré prvotné schopnosti ako sledovanie stimulu sa objavujú už pred 6 mesiacom života dieťaťa. Toto sledovanie sa neskôr rozvinie do zdieľanej pozornosti. Spočiatku sa u dieťaťa vytvára schopnosť jednoduchým spôsobom ukazovať na objekty, ktorá je neskôr vo vyššom veku nahradená verbálnymi schopnosťami upútať ostatných. Jednou z prvých schopností u detí, potrebných pre ďalší rozvoj ZP, je schopnosť sledovania pohybu hlavy alebo smer pohľadu rodiča. Teda je to zameranie smeru pohľadu tak aby smeroval tam kam sa pozerá niekto iný. Schopnosť sa tiež nazýva vizuálna koordinácia a jej výskyt bol preukázaný u detí už vo veku od 2 mesiacov (Bruner, 1977) a v prvom roku života je už plne rozvinutá. Už D'Entremont, Hains, a Muir (1997) ukázali že 3 mesačné deti pohnú očami v smere v ktorom sledovaný dospelý natočí hlavu častejšie ako v opačnom smere. Toto je ale skôr iba prosté sledovanie pohyblivého podnetu, ktoré ale je spočiatku dôležité, neskôr asi v 9 mesiaci, už dieťa môže byť schopné podľa pozície hlavy určiť smer pohľadu.

Zdravé deti v prvých mesiacoch veku majú tendenciu reagovať na pohyb hlavy

matky, kdežto už v deviatom mesiaci začínajú ukazovať na blízke objekty (Schaffer, 1984). Sledovanie pohľadu, pohyby hlavy a ďalšie reakcie predznamenávajú prechod do fázy kedy dieťa začne rozpoznávať vzťahy medzi sebou, ostatnými a okolím, teda už chápe ak niekto iný niečo chce urobiť alebo že sa na neho pozerá.

ZP je veľmi úzko spojená s vývinom jazyka, môže byť označená ako počiatok vývinu jazyka (Tomasello, Farrar, 1986). Výskyt počiatočných pozorovacích schopností detí, ako je sledovanie pohľadu, má dôležitú funkciu v ranných fázach vývinu jazyka. Zdieľaná vizuálna pozornosť má základnú úlohu v spoločenských interakciách niektorých živočíšnych druhov ako sú napríklad gorily, ktoré ju využívajú tak aby ovplyvnili výhľad ostatných (Gomez, 1990).

Tomasello a Farrar (1986) tvrdia, že pre schopnosť dieťaťa naučiť sa pomenovať a pochopiť objekt v správnom kontexte je potrebné aby dieťa bolo v stave zdieľanej pozornosti s rodičom. Baldwin (1995) predpokladá že bez schopnosti zdieľať pozornosť, nie je možné rozpoznať, že ide o nejakú komunikáciu a nie je možné tak nadviazať kontakt.

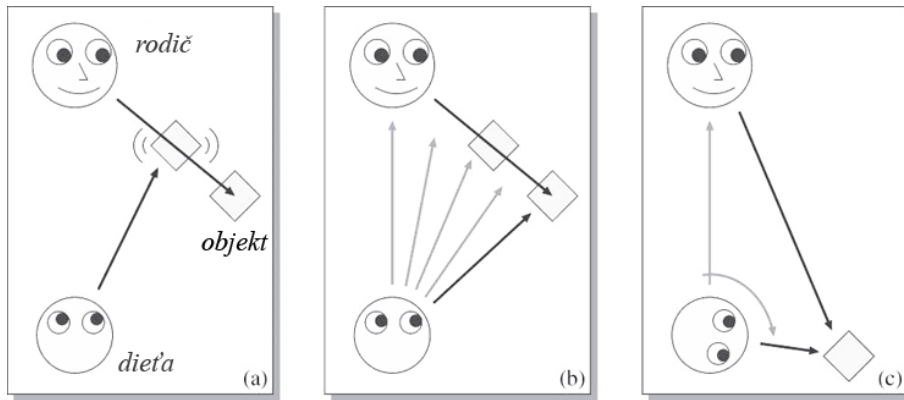
Pokusy detí usmerniť pozornosť niekoho ďalšieho na objekt sa zvyčajne prejavujú v dobe od 13 do 18 mesiaca, teda v čase rozvíjania prvých slov (Bates, Benigni, Bretherton, Camaioni, Volterra, 1979). Bates a spol. zistili že schopnosť nasledovať smer pohľadu, gestá, predvádzanie či ukazovanie, sú dôležitým ukazovateľom pre vývin reči. Ulvund a Smith (1996) dokázali že existuje silný vzťah medzi zahájením ZP, pochopením reči a tiež vyjadrovaním. Mundy a Gomes (1998) zistili že podnet pre ZP je spájaný so schopnosťami sa vyjadrovať, kdežto odpoveď na podnet ZP od druhých sa zase viaže s pochopením reči. Zdieľaná pozornosť je taktiež spájaná s rozvojom slovnej zásoby u detí (Kasari, Freeman, Paparella, 2001). Teda nadobudnutie ZP je kritické pre ďalší mentálny rozvoj človeka.

2.3 Butterworthov model vývinu zdieľanej vizuálnej pozornosti

Butterworth a spol.(1995) rozdelili pomocou rôznych experimentov mechanizmus vývinu zdieľanej vizuálnej pozornosti do troch období, spadajúcich do veku od šiestich do osemnástich mesiacov. V každom tomto období sa u dieťaťa objavujú nové schopnosti. Pri experimente zaradenom do prvého obdobia, deti vo veku šesť mesiacov sedeli oproti svojim matkám a už boli schopné z ich pohľadu určiť kde v miestnosti hľadať zaujímavé objekty, no už nevedeli presnejšie zistiť na ktorý objekt sa matka pozerá, tiež neboli schopné obzrieť sa za seba, ak sa matka pozerala na objekt za nimi. Deti pozorujúce svoje matky boli schopné otáčať sa v smere jej pohľadu, no zastavili sa na prvom objekte ktorý ich zaujal. Ak sa matka pozerala na objekt čo sa nachádzal za dieťaťom, dieťa sa pozrelo buď na objekt niekde pred sebou alebo nereagovalo. Butterworth toto nazval ekologickým mechanizmom, kde deti sú schopné využiť informáciu o niekoho pohľade a určiť tak miesto výnimočnej udalosti.

Podľa Butterwortha deti v dvanástich mesiacoch už vedia presnejšie lokalizovať pohľad matky. Vykazujú nové správanie, kedy uprene pozorujú matku počas jej pohybu, a ak sa matka niekam zahľadí, dieťa ju ešte krátko potom nehybne pozoruje a o chvíľu neskôr sa prudko obzrie v smere matkinho pohľadu na cieľ jej záujmu. Butterworth toto nazval geometrickým mechanizmom, predpokladal že tu deti uvažujú určitú neviditeľnú čiaru medzi matkou a cieľom, tak boli schopné prekonať iné prekážajúce objekty pri lokalizovaní správneho cieľa, ktorý sa môže nachádzať aj vo viac periférnej oblasti. No stále deti takto nedokážu nájsť objekty umiestnené za sebou.

Schopnosť obzrieť sa za objektom, v prípade že nie je žiadny iný vo výhľade, sa objavuje vo veku osemnástich mesiacov. Toto obdobie je označované ako reprezentatívne, pretože Butterworth predpokladal, že deti využívajú znalosť o priestore v ktorom sa nachádzajú a ktorý obsahuje objekty aj mimo zorného pola dieťaťa.



Obrázok 2.2: Tri obdobia vývinu zdieľanej pozornosti. (Nagai a spol. 2003)

Na obr. 2.2 môžeme vidieť jednotlivé obdobia, ktoré sa delia nasledovne:

- **Ekologické obdobie** - prvé obdobie vo veku okolo 6 až 9 mesiacov obr. 2.2(a). Dieťa má tendenciu sa pozeráť na blízke objekty ktoré sú výrazné alebo pozoruhodné ako aj na rodiča. (Prevažne sú to objekty ktoré sa pohnú alebo majú pre dieťa výrazné vizuálne charakteristiky ako lesk či farba.) Ak sa rodič pozerá na blízky objekt ktorý má dieťa v zornom poli, dieťa vie využiť pohľad rodiča na približné otočenie daným smerom. Ak je bližšie nejaký iný objekt dieťa sa zastaví pohľadom na tomto objekte. Ak sa rodič pozrie na objekt ktorý dieťa nevidí, buď sa pozrie na prvý objekt čo zbadá alebo nereaguje.
- **Geometrické obdobie** - ďalšie obdobie nastáva vo veku približne 12 mesiacov obr. 2.2(b). Dieťa v tomto období už vie sledovať smer pohľadu rodiča. Dieťa vie nájsť objekt ktorý je v smere pohľadu rodiča ak je objekt zároveň v zornom poli dieťaťa, zároveň je schopné už nájsť objekt v smere rodičovho pohľadu aj keď je pred ním iný objekt.
- **Reprezentačné obdobie** - obdobie okolo 18-tého mesiaca dieťaťa, dieťa už má plne vyvinutú schopnosť zdieľanej pozornosti obr. 2.2(c). Dokáže porozumieť informácií o rodičovom pohľade, vie sa otáčať v smere pohľadu rodiča a nájsť tak aj objekt ktorý je spočiatku mimo zorného poľa.

Táto teória je často kritizovaná pretože nevysvetľuje ako a kde vznikajú tieto tri mechanizmy, existujú dôkazy ktoré sa ukazujú nedostatky tejto teórie, ukázalo

sa že tieto mechanizmy sa objavujú už skôr. Deti v už 12. mesiaci, sú schopné nasledovať pohľad rodiča za seba a v 9. mesiaci sú schopné ignorovať objekt pred sebou a pozrieť sa na periférny objekt. Ukazuje sa že sú aj iné potrebné faktory ako habituácia a motivácia alebo spôsob podnetu, ktoré sú dôležité a nie sú sem nijako zahrnuté. (Deák, Triesch 2006)

Ukazuje sa že deti v rôznom veku určujú smer pohľadu rodiča podľa určitých charakteristík, napríklad deti mladšie ako 12 mesiacov určujú smer pohľadu podľa pozície hlavy rodiča, kdežto u starších detí medzi 12-14 mesiacmi sa rapídne zvyšuje citlivosť na pozíciu rodičovských očí. V dobe okolo osemnásť mesiacov na sledovanie pohľadu plne postačuje len pohyb očí. Otázkou je ako deti chápu takéto správanie rodiča. Woodward (2003) ukázal že deti nepotrebujú pochopiť vzťah medzi rodičom a objektom na ktorý sa rodič pozerá, na to aby sa naučili sledovať pohľad. Na úspešné sledovanie pohľadu nemusí byť vyžadovaná žiadna reprezentačná stratégia zahrňujúca identifikáciu rodičovských úmyslov. Zaiste sa takéto reprezentácie časom u starších detí objavia, ale nie sú potrebné pre vysvetlenie vzniku sledovania pohľadu. Sledovanie pohľadu je možné vnímať ako jednoduchú formu imitácie, ak sa niekto pozerá smerom ktorým sa pozerá niekto iný, imituje jeho jednanie, tak že sa pozerá rovnakým smerom. (Triesch a spol., 2006).

2.4 Baron-Cohenove moduly zdieľanej vizuálnej pozornosti

O vytvorenie biologicky a výpočtovo relevantného modelu vzniku zdieľanej vizuálnej pozornosti sa pokúsil Baron-Cohen (1995). Predstavil teóriu o existencii niekoľkých spolupracujúcich modulov v mozgu, zodpovedných za vznik ZP. Evolučne najstarším modulom u ľudí ako aj zvierat je tzv. detektor pohľadu očí (Eye Direction Detector, EDD). Zároveň majú ľudia vrodenný detektor intencionality (Intentionality Detector, ID). EED je vizuálny modul, ktorý rozpoznáva pohľad inej osoby a vytvára dyadické vzťahy pomocou presných geometrických výpočtov. ID je multimodálny modul, ktorý pridáva informáciu o vnútornej potrebe ku vnímanému ob-

jektu. EED a ID posielajú informáciu na vstup modulu zdieľanej pozornosti, ktorý z nich vytvára triadické reprezentácie. Nakoniec sú odoslané do modulu teórie mysle, schopného reprezentovať mentálne koncepty v plnom rozsahu.

Táto teória sa vyvoláva množstvo problémov, ako napríklad nedostatok relevantných neuropsychologických dôkazov o týchto moduloch. Taktiež sa nevie ako tieto moduly ovplyvňujú známe procesy ako sú habituácia, pozornostné procesy, vizuálne procesy, priestorová pamäť usudzovanie, atď. Zároveň existuje množstvo empirických dát, ktoré sú v neprospech tejto teórie (Deák, Triesch 2006).

2.5 Model zahrňajúci štruktúry a mechanizmy podieľajúce sa na vývine zdieľanej pozornosti

Triesch a spol. (2006) vychádzajúc z predpokladu, že sa dieťa naučí sledovať pohľad rodiča, pretože zisťuje, že smer rodičovho pohľadu umožňuje predvídať, kde sa bude nachádzať zaujímavý vizuálny stimul, navrhujú model, kde sa zdieľaná pozornosť objavuje na základe vzájomného pôsobenia rôznych štruktúr a mechanizmov, ako sú vnemové schopnosti a preferencie, habituácia, učenie posilňovaním a štruktúrované sociálne prostredie.

2.5.1 Vnemové schopnosti

Vnemové schopnosti a preferencie, ktoré sú už rozvinuté vo veku 3 mesiacov, môžu byť dôležité pre vývin ZP. Už novorodenci uprednostňujú viac ľudské stimuly ako tváre a hlasy. Deti vo veku 2 mesiacov začínajú reagovať viac sústredene na rodičov. Čo sa dá pripísať predpokladu, že ľudia sú pozoruhodnejší ako neživé veci (Bates a spol., 1979). V rovnakom čase sa u detí objavujú prvé sociálne prejavy ako úsmev a zvyšuje sa zapájanie pri interakciách s rodičmi (Cole a Cole, 1996). Tiež viac preferujú sledovanie očí ako napríklad ústa rodiča (Haith a spol. 1979). Pre je vývin ZP veľmi dôležité rozpoznávať tváre a ich orientáciu, ku úspešnému

sledovaniu pohľadu je tak potrebná schopnosť rozpoznať rodičov pohľad, tá musí byť namapovaná ku náležitým motorickým funkciám, k čomu je už potrebné zapojiť ďalšie učiace procesy.

2.5.2 Štrukturované sociálne prostredie

Triesch a spol.(2006) tvrdia, že najvhodnejšie situácie pre naučenie ZP, sú pri bežných úkonoch ako kŕmenie, kúpanie či hry s rodičom. Čo je ale dôležité pri týchto úkonoch je predvídateľná štruktúra nasledujúcich udalostí, ktorú je možné si osvojiť pomocou učenia posilňovaním a habituácie. Dieťa sa môže naučiť ako využiť situáciu pri týchto úkonoch. Dieťa začne predvídať, kde sa udejú zaujímavé udalosti na základe pohľadu rodiča. Pretože pohľad rodiča vedie väčšinou k ďalším osobám alebo udalostiam, ktoré môžu byť pozoruhodné a pre dieťa zaujímavé.

2.5.3 Učenie posilňovaním

Triesch a spol.(2006) zastávajú názor že učenie posilňovaním je dôležité pri vývine zdieľanej pozornosti. Učenie posilňovaním nastáva už v prvých mesiacoch života dieťaťa. Už dvojmesačné sa dieťa vie naučiť počas niekoľkých minút predvídať následnú akciu pri opakovaných úkonoch. Triesch a spol.(2006) navrhujú model učiacich mechanizmov získavania zdieľanej pozornosti je neurálne plauzibilný proces učenia posilňovaním zvaný TD učenie (temporal difference) (Sutton & Barto, 1998). Excitácia určitých neurónov v častiach bazálnych ganglií bola asociovaná ku TD signálu, podľa ktorého TD učenie prevzalo názov. Je možné že pomocou TD metód by bolo možné pochopiť vývin správania, ktoré zahŕňa určitú formu odmeny, keď sú deti odmenené stimulom pri nasledovaní ZP. Avšak dodávajú, že učenie posilňovaním nemusí byť jediný učiaci mechanizmus, čo sa podieľa na vývine zdieľanej pozornosti.

2.5.4 Habituácia

Habituácia je jeden z dôležitých faktorov pri učení spolu s TD učením, kde proces habituácie mení vnútorné ohodnotenie pozorovaného objektu a tým ovplyvňuje správanie dieťaťa. Napríklad ak sa dieťa pozrie na tvár rodiča alebo na určitý objekt, nastáva habituácia ktorá môže byť modelovaná ako postupné znižovanie odmeny pozorovaného stimulu. Táto napomáha presunutiu pozornosti na nový zaujímavejší objekt. Naopak dishabituácia čiastočne obnovuje hodnotu už habituovaného objektu, v prípade že dieťa pozoruje iný objekt. Tieto procesy spolu s TD učením, vytvárajú opakované presúvanie pozornosti z tváre rodiča na iné objekty a zase späť na tvár rodiča. Takéto presuny pozornosti sú dôležité pri získavaní schopnosti ZP v závislosti od správania rodiča (Triesch a spol., 2006).

No aj keď v súčasnosti existujú rôzne teórie o vzniku ZP, zatiaľ nie je žiadna, ktorá by dostatočne uspokojivo zahrnula všetky aspekty vývinu ZP počas prvých osemnástich mesiacov.

Kapitola 3

Výpočtové modelovanie

V tejto kapitole sa budeme venovať teórií niektorých výpočtových modelov, ktoré následne využijeme pri vytváraní zjednodušeného modelu ZP. Nazrieme do problematiky neurónových sietí a tiež učenia posilňovaním.

3.1 Umelé neurónové siete

Umelé neurónové siete (UNS) ako výpočtová paradigma, inšpirovaná biologickými neurónovými sieťami. Neurónové siete sú zložené z veľkého množstva vzájomne poprepájaných výpočtových prvkov (neurónov) spolupracujúcich na riešení špecifických problémov. UNS sa učia z predkladaných príkladov. S ich pomocou je možné riešiť rôzne problémy ako rozpoznávanie vzorov, klasifikácia dát či iné spracovanie a analyzovanie dát. Učenie v UNS je rovnako ako v biologických sieťach postavené na zmenách synaptických spojení ktoré spájajú neuróny medzi sebou.

Jedným zo základných typov sietí sú dopredné neurónové siete, ale aj existuje množstvo ďalších typov sietí ako sú rekurentné siete, samoorganizačné mapy či impulzné siete a množstvo ďalších.

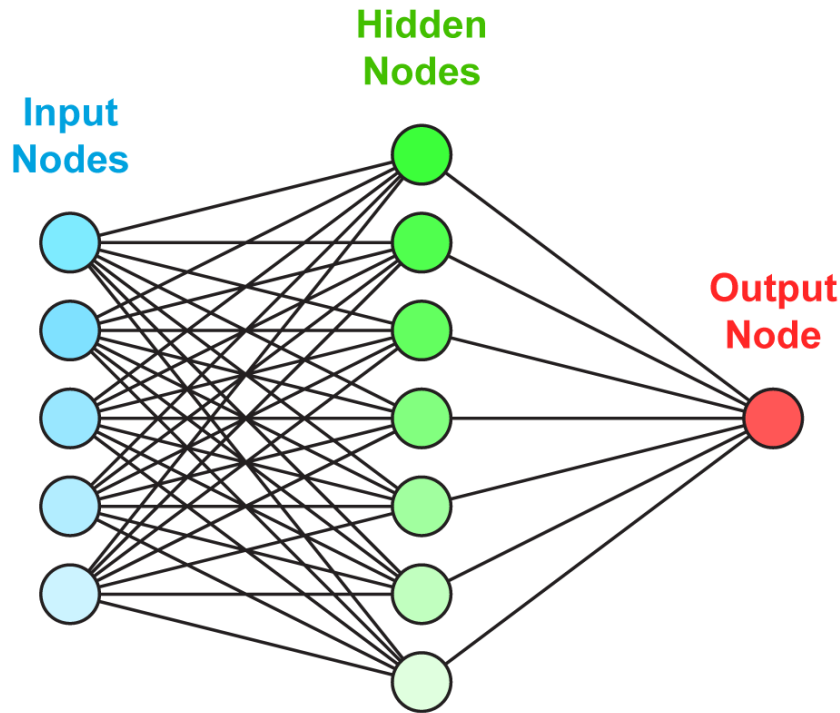
Učenie v neurónových sieťach môže byť rozdelené do troch hlavných kategórií a to učenie s učiteľom, učenie bez učiteľa a učenie posilňovaním. Učenie s učiteľom vyžaduje vplyv nejakej externej zložky, ktorá vie určiť pre dané vstupy siete správne hodnoty na výstupoch. Tieto hodnoty sú potom využité pri učení na úpravu váh si-

ete aby bola chyba medzi výstupnými a správnymi hodnotami čo najmenšia. Úlohou je nájsť také váhy pre ktoré je chyba čo najmenšia. Sem spadajú typy učenia ako napríklad učenie so spätným šírením chyby.

Učenie bez učiteľa nemá žiadnu externú zložku, ktorá by poznala správne hodnoty, preto závisí čisto na lokálnej informácií v sieti. Často sa takéto učenie odkazuje na samoorganizáciu, v zmysle že si sieť sama vie zorganizovať dáta ktoré sú sieti prezentované a rozpozná nejaké ich spoločné vlastnosti, korelácie, príznaky či drobné štatistické pravidelnosti len na základe vstupov. V takýchto typoch sietí je učenie reprezentované pomocou hebbovského učenia ako aj jeho rôznych derivátov. Poslednej kategórií učenia, teda učeniu posilňovaním, sa budeme zaoberať omnoho podrobnejšie neskôr.

3.1.1 Viacvrstvový perceptrón

V typickej doprednej sieti sa nachádzajú 3 alebo viac vrstiev neurónov. Rozdelené na vstupnú vrstvu, jednu či viac skrytých vrstiev a výstupnú vrstvu. Informácia sa šíri iba jedným smerom, zo vstupnej do výstupnej vrstvy. Synaptické spojenia alebo váhy označujú prepojenia medzi jednotlivými neurónmi medzi vrstvami. Doprednú sieť, kde sú neuróny v jednej vrstve pospájané so všetkými neurónmi z ďalšej vrstvy doprednými spojeniami, môžeme nazvať viacvrstvom perceptrónom.



Obrázok 3.1: Dopredná neurónová sieť. Informácia sa šíri zľava doprava.
(Abazov a spol., 2009).

Ak vstupné dáta do siete označíme ako vektor X a výstupné ako vektor Y , tak potom funkcia UNS môže byť opísaná ako operácia ktorá mapuje X do Y , teda $f : X \rightarrow Y$.

Váhy medzi vstupnou a skrytou vrstvou môžeme označiť ako maticu váh W , kde w_{ij} označuje váhu medzi i -tým vstupným a j -tým skrytým neurónom. Obdobne máme tiež maticu váh Z medzi skrytou a výstupnou vrstvou. Váhy určujú ako silno ovplyvňuje jeden neurón nasledujúci.

Vstupné dáta X do neurónu sú upravené podľa korešpondujúcich váh a následne spočítané. A následne upravené aktivačnou funkciou g . Pre j -ty neurón skrytej vrstvy h_j bude aktivácia na výstupe:

$$h_j = g\left(\sum_i x_i w_{ij}\right). \quad (3.1)$$

Pre výpočet aktivácie výstupného neurónu y_k využijeme aktivácie z predošlej

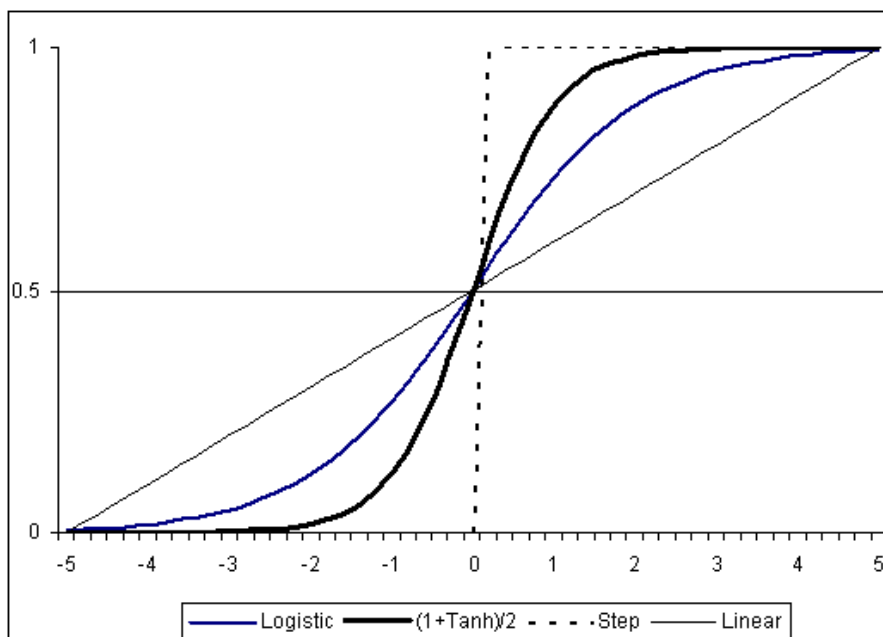
vrstvy, z_{jk} označuje váhy medzi skrytou a výstupnou vrstvou:

$$y_k = g \left[\sum_j z_{jk} g \left(\sum_i x_i w_{ij} \right) \right]. \quad (3.2)$$

Aktivačná funkcia perceptrónu býva typicky v tvare lineárnej či sigmoidálnej (logistickej) funkcie. A napríklad pri RBF sietiach býva v podobe gaussovskej funkcie. Najbežnejšie používanou aktivačnou funkciou vo viacvrstvom preceptróne býva sigmoidálna funkcia, ktorú môžeme definovať ako:

$$g(net) = \frac{1}{1 + e^{-\beta net}}, \quad (3.3)$$

kde zisk β určuje sklon funkcie. Na obr. 3.2 sú zobrazené ďalšie možné aktivačné funkcie.



Obrázok 3.2: Príklad aktivačných funkcií. (de Smith, 2007)

3.1.2 Učenie so spätným šírením chyby

(Kvasnička a i. 1997) Na to aby dopredná UNS dokázala riešiť nejaký konkrétny problém je nutné ju natrénovať, tak že jej predkladáme vstupné dáta a sledujeme

aké vráti hodnoty na výstupoch, ak tieto hodnoty nezodpovedajú požadovaným hodnotám, snažíme sa upraviť váhy aby rozdiel výstupnej hodnoty a požadovanej hodnoty bol čo najmenší. K tomuto potrebujeme mať vhodne zvolené tréningové a testovacie dáta. Tréningové dáta použijeme na tréning, testovacie dáta poslúžia na zistenie schopnosti siete generalizovať. Jedným zo spôsobov učiť doprednú sieť je učenie so spätným šírením chyby. Učenie prebieha tak, že zistíme aktivácie na výstupoch siete y_j dopredným šírením z vybraných tréningových dát. Použijeme sigmoidálnu aktivačnú funkciu:

$$y_i = \frac{1}{1 + e^{-x_i}},$$

y_i je aktivácia na i -tom neuróne, x_i je jeho váhovaná suma vstupov.

Ak máme všetky aktivity na výstupných neurónoch, zistíme ako sa hodnoty na výstupoch siete líšia od žiadaných hodnôt, určíme si preto účelovú funkciu:

$$E = \frac{1}{2} \sum_j (y_j - d_j)^2,$$

kde d_j je požadovaná aktivácia na y_j výstupnom neuróne.

Následne zistíme ako sa mení chyba v závislosti od zmeny aktivity výstupného neurónu

$$EA_j = \frac{\partial E}{\partial y_j} = y_j - d_j,$$

vypočítame ako rýchlo sa mení chyba keď sa zmení súčet vstupov vo výstupnom neuróne

$$EI_j = \frac{\partial E}{\partial x_j} = \frac{\partial E}{\partial y_j} \frac{\partial y_j}{\partial x_j} = EA_j y_j (1 - y_j)$$

a určíme aká bude zmena chyby pri zmene váhy na výstupnom neuróne

$$Ew_{ij} = \frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial x_j} \frac{\partial x_j}{\partial w_{ij}} = EI_j y_i$$

Výpočet zmeny chyby od zmeny aktivity neurónu z predchádzajúcej vrstvy:

$$EA_i = \frac{\partial E}{\partial y_i} = \sum_j \frac{\partial E}{\partial x_j} \frac{\partial x_j}{\partial y_i} = \sum_j EI_j w_{ij}$$

takto zistíme chyby aktivácií EA pre všetky neuróny po vrstvách. Potom už len stačí vypočítať zmenu váh Ew pre konkrétne váhy a upraviť ich napríklad pomocou metódy najprudšieho spádu, tak aby sme minimalizovali účelovú funkciu opakovaním nasledujúceho vzťahu:

$$w_{ij}^{k+1} = w_{ij}^k - \eta E w_{ij}^k + \mu \Delta w_{ij}^k,$$

kde w_{ij}^{k+1} je nová hodnota váhy w_{ij}^k , η je učiaci parameter ktorý zabezpečuje vhodnú konvergenciu, typicky $0,01 \leq \eta \leq 0,1$. $\mu \Delta w_{ij}^k$ je momentový člen, ktorý pomáha vyhnúť sa lokálnym minimám, ktorý je určený pomocou rozdielu koeficientov z posledných dvoch iterácií $\Delta w_{ij}^k = w_{ij}^k - w_{ij}^{k-1}$ a μ je parameter $0,5 \leq \mu \leq 0,7$.

3.2 Učenie posilňovaním

Učenie posilňovaním, ďalej len RL, ako jeden z výpočtových prístupov sa snaží podchytiť podstatu cieľovo orientovaného učenia a rozhodovania relevantného z biologického pohľadu. Od iných prístupov sa RL odlišuje hlavne na základe poňatia problému cez priamu interakciu s prostredím a bez akejkoľvek priamej kontroly učenia, ako v prípade učenia s učiteľom (supervised learning), zároveň dokáže pracovať v prostredí, ktoré nemusí byť je plne pozorovateľné. RL využíva systém interakcií medzi učiacim agentom a prostredím popísaný pomocou stavov, akcií a odmien (Sutton, Barto, 2004).

RL sa snaží vyberať akcie tak aby dokázal maximalizovať číselnú odmenu. Učiaci agent musí sám objaviť, ktoré akcie su najlepšie pre dosiahnutie najväčšej odmeny pomocou použitím pokus-omyl metódy učenia, teda sa uči pomocou svojich nadobudnutých skúseností.

RL je definovaný charakterizovaním učiaceho problému, ktorý je definovaný cez markovovský rozhodovací proces. Podstatou je zachytiť najdôležitejšie aspekty

problému, ktorému sa musí učiaci agent postaviť aby dosiahol svoj cieľ. Agent musí byť schopný určiť stav svojho prostredia a musí vedieť vykonávať akcie v tomto stave, ktoré pomôžu k tomu aby sa agent dostal do cieľa, ktorý je taktiež definovaný ako určitý stav prostredia.

Jeden z dôležitých aspektov RL je nájdenie kompromisu medzi skúmaním nových možností a využitím nadobudnutých skúseností, pre nadobudnutie čo najlepšej odmeny si RL agent musí vybrať či zvolí akciu o ktorej vie že je výhodná vo význame najväčšej získanej odmeny alebo skúsi zvoliť novú akciu pre neho doteraz s neznámou odmenou, tá sa môže ukázať ako prospešnejšia. Agent musí voliť akcie tak aby boli čo najefektívnejšie. Napríklad pri stochastických úlohách je potrebné urobiť veľké množstvo akcií aby sme získali dostatočne dobrý odhad odmeny.

3.2.1 Prvky učenia posilňovaním

V RL systéme môžeme odhaliť tiež štyri dôležité prvky: *stratégiu*, *úžitkovú funkciu*, *ohodnocovaciu funkciu* a *model prostredia*, ktoré si následne v krátkosti popíšeme.

Stratégia určuje ako sa učiaci agent správa v danom čase. Čiže stratégia je mapovanie vnímaných stavov prostredia do akcií, ktoré budú vykonané ak sa agent nachádza v týchto stavoch. Stratégia môže byť napríklad rozhodovacia tabuľka, funkcia alebo neurónová sieť, ale taktiež môže byť reprezentovaná formou rôznych iných výpočtovo náročných procesov.

Úžitková funkcia v podstate definuje cieľ v RL problému. Funkcia priradí ku každému vnímanému stavu prípadne danému páru stav-akcia prostredia určitú numerickú hodnotu - *odmenu*, ktorá vyjadruje ako veľmi je žiadúce sa v tomto stave nachádzať. Hlavnou úlohou RL učiaceho agenta je získať čo najvyšší súčet odmien z dlhodobého hľadiska. Úžitková funkcia teda stanovuje ako sú situácie pre agenta dobré či zlé. Dôležité je aby úžitková funkcia bola úplne nezávislá od agenta a teda ju nemohol ovplyvniť. No taktiež slúži ako významný údaj pre zmenu stratégie, teda ak súčasná stratégia vyberie akciu ktorá vedie ku odmene s malou hodnotou, potom by sa mala stratégia, tak aby vybrala v budúcnosti v tejto situácii vybrala inú akciu.

Ohodnocovacia funkcia určuje *odhadovanú hodnotu* stavu, ktorá má význam z dlhodobého hľadiska. Odhadovaná hodnota stavu je daná súčtom odmien, ktoré agent bude môcť zozbierať ak bude pokračovať z tohoto stavu ďalej. Kde odmeny rozhodujú o okamžitej vhodnosti stavu, odhadované hodnoty vyjadrujú vhodnosť stavu z dlhodobého hľadiska, odhadované hodnoty berú do úvahy odmeny v stavoch ktoré budú nasledovať. Napríklad v stave kde je odmena malá, stále môže byť odhadovaná hodnota, ktorá je veľká pretože nasledujúce stavy dávajú veľké odmeny. Alebo naopak. Teda odmeny a odhadované hodnoty sú vzájomne závislé. Bez odmien tu nie sú odhadované hodnoty a zároveň účelom hodnôt je získať čo najväčšie odmeny. Pri rozhodovaní o vhodnosti akcií sú pre nás dôležitejšie odhadované hodnoty. Akcie ktoré hľadáme majú mať čo najvyššiu odhadovanú hodnotu, aktuálna odmena nieje až tak dôležitá, tieto akcie s vysokou odhadované hodnotou nám dopomáhajú získať najpriaznivejší súčet odmien z dlhodobého hľadiska. Nanešťastie určiť odhadované hodnoty je náročnejšie ako odmeny. Odmeny nám v podstate dáva priamo prostredie, zato odhadované hodnoty sú vypočítavané neustále podľa pozorovaní až do konca učenia. Dá sa tvrdiť že nájdenie najefektívnejšej ohodnocovacej funkcie sa stáva najdôležitejšou zložkou v RL algoritmoch. Avšak riešenie RL problémov existujú ďalšie spôsoby ako napríklad použitie genetických algoritmov či simulovaného žihania, ktoré k úspešnému nájdeniu riešenia nepotrebujú žiadnu ohodnocovaciu funkciu.

Posledným prvkom niektorých RL systémov je *model prostredia*, ktorý napodobuje správanie prostredia. Agent podľa stavu a akcie vie predikovať nasledovný stav a odmenu v tomto stave, tým vytvorí model použiteľný pre plánovanie, teda spôsob jednania ovplyvňujú aj situácie v budúcnosti ktoré zatiaľ neboli spoznané. Teda RL systémy sú schopné učiť sa v zmysle pokus-omyl učenia a zároveň sa naučiť model prostredia následne ho využiť ako model pre plánovanie. V súčasnosti existuje množstvo RL od jednoduchého pokus-omyl učenia až po vysokoúrovňové plánovanie.

3.2.2 Markovovský rozhodovací proces

RL ktorý spĺňa Markovovskú vlastnosť sa nazýva markovovský rozhodovací proces (MDP). Markovovská vlastnosť je určená ako vlastnosť systému, kde budúce stavy

sú závislé iba od súčasného stavu a nezávislé od stavov minulých.

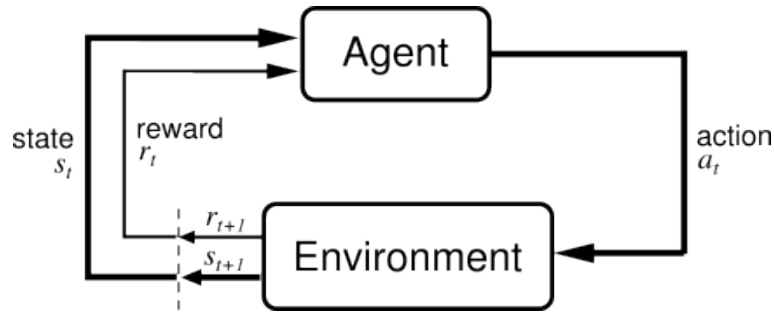
Problémy ktoré môžu byť riešené pomocou RL môžu byť modelované ako MDP. Konečný MDP je (S, A, R, P) , kde $s_t \in S$ označuje stav agenta, v ktorom sa nachádza v čase t . r_t označuje možnú odmenu v čase t a $R_{ss'}^a$ označuje očakávanú odmenu pri prechode z s do s' po vykonaní akcie a . P je prechodová funkcia, $P_{ss'}^a$, ktorá určuje pravdepodobnosť, že po vykonaní akcie a v stave s sa ocitneme v stave s' .

$$P_{ss'}^a = Pr \{s_{t+1} = s' | s_t = s, a_t = a\} \quad (3.4)$$

$$R_{ss'}^a = E \{r_{t+1} | s_t = s, a_t = a, s_{t+1} = s'\} \quad (3.5)$$

3.2.3 Definícia problému učenia posilňovaním

RL problém môžeme zjednodušiť na problém učenia pomocou interakcie agenta s prostredím na dosiahnutie určitého cieľa. Entita ktorá sa učí a rozhoduje sa sa nazýva *agent*, ostatné čo je mimo agenta a s čím je v interakcii je *prostredie*. Vzájomne na seba pôsobiac, agent vykonáva akcie a od prostredia dostáva odmeny, zároveň sa ocitá v nových situáciách. Agentovou úlohou je maximalizovať odmeny ktoré získa. Agent a prostredie sú spolu v interakcii v určitých časových diskretných krokoch $t = 0, 1, 2, 3, \dots, n$. V každom kroku t agent získa špecificky definovanú reprezentáciu stavu tohoto prostredia, $s_t \in S$, kde S je množina všetkých možných stavov. V každom stave s_t môže agent zvoliť akciu $a_t \in A(s_t)$, kde $A(s_t)$ je množina akcií dostupných v tomto stave. V ďalšom kroku následkom svojej akcie získa agent numerickú *odmenu*, $r_{t+1} \in R$ a ocitne sa v novom stave s_{t+1} .



Obrázok 3.3: Schéma interakcie agenta s prostredím (Sutton, Barto, 2004).

V každom kroku agent mapuje pravdepodobnosti výberu z akcií. Toto mapovanie sa nazýva stratégia π_t , kde $\pi_t(s, a)$ je pravdepodobnosť že $a_t = a$ ak $s_t = s$. Agent časom mení svoju stratégiu na základe získaných skúseností, ako ju mení závisí od RL metódy. V RL modeloch stavy a akcie môžu byť takmer ľubovoľne definované, či už ako jednoduché signály alebo komplexné rozhodovacie procesy. Veľmi dôležité je oddelenie agenta prostredia, tak aby agent nemohol ľubovoľne ovplyvňovať prostredie.

Úlohou agenta je získať čo najväčší súčet odmien z dlhodobého hľadiska. V skutočnosti sa agent snaží maximalizovať očakávaný zisk, kde R_t je funkcia zahŕňajúca sled budúcich odmien $r_{t+1}, r_{t+2}, r_{t+3}, \dots, r_T$ po kroku t . T je posledný krok. Takto môžeme rozdeliť interakcie agenta s prostredím do epizód, napríklad na hry, epochy alebo akékoľvek opakované úkony. Každá epizóda končí v terminálnom stave, nasledovaná obnovením východzieho stavu. Takéto epizódy sa nazývajú epizodickými úlohami naproti tomu opačné typy úloh sa nazývajú trvalé (continuing) úlohy. Tu v prípade že $T = \infty$ sa môže stať že odmena bude ∞

Preto si zavedieme koncept diskontu. Agent sa teda bude snažiť vyberať akcie tak aby maximalizoval sumu budúcich diskontovaných odmien. Pre a_t teda dostane:

$$R_t = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \dots = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1}, \quad (3.6)$$

kde γ parameter $0 \leq \gamma \leq 1$ je diskontný faktor (discount rate). Pre k krokov v budúcnosti je odmena iba γ^{k-1} -krát menšia. Ak $\gamma = 0$ agent v a_t maximalizuje iba okamžitú odmenu r_{t+1} .

Skoro všetky RL algoritmy majú ohodnocovaciu funkciu. Funkcie stavov alebo stav-akcia párov odhadujú ako dobré je pre agenta byť v tomto stave alebo ako je dobré vybrať akciu v tomto stave, vo význame očakávania budúcich odmien teda očakávaný návrat. Ohodnocovacie funkcie sú definované vo vzťahu ku stratégií.

Stratégia π je zobrazenie dvojíc (s, a) do pravdepodobnosti $\pi(s, a)$ vybratia akcie a v stave s . odhadovaná hodnota stavu s v stratégií π označíme $V^\pi(s)$, čo je očakávaná hodnota v stave s nasledovaním stratégie π . $V^\pi(s)$ sa nazýva stavová-hodnota funkcie pre stratégiu π nasledujúcu zo stavu s . Pre MDP môžeme definovať:

$$V^\pi(s) = E_\pi \{R_t | s_t = s\} = E_\pi \left\{ \sum_{k=0}^{\infty} \gamma^k r_{t+k} + k + 1 | s_t = s \right\}, \quad (3.7)$$

E_π určuje očakávanú hodnotu keď agent nasleduje stratégiu π v kroku t .

Podobne definujeme odhadovanú hodnotu vybratia akcie a v stave s v π stratégií, $Q^\pi(s, a)$ ako očakávaný návrat zo stavu s výberom akcie a nasledovaním stratégie π . Q^π sa nazýva akciová-hodnota pre funkciu v stratégií π :

$$Q^\pi(s, a) = E_\pi \{R_t | s_t = s, a_t = a\} = E_\pi \left\{ \sum_{k=0}^{\infty} \gamma^k r_{t+k} + k + 1 | s_t = s, a_t = a \right\} \quad (3.8)$$

3.2.4 Explorácia a exploatácia

Vlastnosť ktorá odlišuje RL od iných typov učenia je že RL hodnotí akcie ktoré boli zvolené miesto toho aby sa učil z predom daných správnych akcií. Preto je nutné aby agent vedel prehľadávať priestor a tu vzniká potreba explorácie, keďže sa agent učí spôsobom pokus-omyl.

Ak uvažujeme nejaký RL problém, tak pre každú akciu máme nejakú očakávanú odmenu ktorú môžeme získať ak túto akciu vyberieme. Nazvime ju teda *odhadovaná hodnota* akcie. V prípade že poznáme hodnoty pre všetky akcie, tak vieme nájsť optimálne riešenie problému, pretože vyberieme vždy akcie s najvyššou hodnotu. V skutočnosti hodnoty ale nemusíme poznať no môžeme ich odhadovať. Keď v danom čase existuje aspoň jedna akcia ktorej odhadovaná hodnota je väčšia než ostatné, označujeme ju ako „hladnú“ (greedy) akciu. Teda ak je takáto akcia zvolená, využívame teda svoju znalosť hodnôt akcií (exploatácia). Ak zvolíme inú

akciu (nongreedy), teda hľadáme nové možnosti (explorácia), a môžeme tak upraviť hodnotu tejto (nongreedy) akcie a zistiť či je perspektívnejšia. Napríklad ak vieme že hodnota hladnej akcie je už známa ale stále existujú podobne vhodné akcie, ktorých hodnoty sú ešte stále čiastočne nejasné môže to znamenať že niektorá z akcií bude pravdepodobne lepšia ako hladná akcia. Preto môže byť z dlhodobého hľadiska výhodné overiť tieto akcie. V rôznych prípadoch voľba závisí na presnosti odhadu, nejasnosti hodnôt a zvyšnom počte krokov do konca učenia. Existujú rôzne sofistikované matematické metódy ako presne vyvážiť exploráciu a exploataciu ale pre použitie v RL sú nepoužiteľné. No existujú jednoduchšie metódy ktoré sú dostatočné pre rôzne RL problémy.

V tejto časti sa pozrieme na niektoré spôsoby hľadania správneho vyváženia explorácie a exploatacie.

3.2.4.1 ϵ -greedy

ϵ -greedy je jedným zo základných a najjednoduchších spôsobov explorácie. ϵ -greedy pravidlo zvolí náhodnú akciu určitou pravdepodobnosťou ϵ . Inak je zvolená hladná odhadovaná akcia s pravdepodobnosťou $(1 - \epsilon)$.

3.2.4.2 Softmax výberové pravidlo

Nevýhodou ϵ -greedy je že pri explorácii volí akciu rovnomerne spomedzi všetkých akcií, teda môže viesť k výberu najhoršej možnej akcie, čo je v niektorých prípadoch veľmi nežiadúce. Softmax pravidlo odstupňuje akcie podľa odhadovaných hodnôt. Greedy akcia má stále najväčšiu pravdepodobnosť byť zvolená, ostatné akcie môžu byť vybrané podľa ich odhadovanej hodnoty. Najčastejšie používajú Gibbsovu alebo Boltzmanovu distribúciu pre výber akcie s pravdepodobnosťou:

$$\frac{e^{Q_t(a)}/\tau}{\sum_{b=1}^n e^{Q_t(b)}/\tau} \quad (3.9)$$

kde τ parameter sa označuje ako teplota. Vysoký parameter teploty spôsobuje že akcie budú mať takmer podobnú pravdepodobnosť výberu, ak $\tau \rightarrow 0$ tak výber akcie sa stáva greedy.

3.2.4.3 Gaussovská explorácia

Pre niektoré algoritmy ako CACLA je vhodná veľmi častá explorácia a zároveň keď sú vyberané akcie blízko súčasnej optimálnej odhadovanej akcií (van Hasselt, H, Wiering, M. A. 2009). Preto je vhodné využiť náhodný výber z gaussovskej pravdepodobnostnej funkcie $G(x, \mu, \sigma)$. μ označuje aktuálnu optimálnu akciu, σ bude reprezentovať exploračný parameter.

$$\pi_t(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (3.10)$$

S najväčšou pravdepodobnosťou bude zvolená akcia ktorá bude najbližšie k výslednej hodnote.

3.2.5 Diskrétne algoritmy učenia

3.2.5.1 TD učenie

TD učenie je spôsob predikcie využívajúci následných odhadovaných hodnôt. Tento princíp umožňuje vyhodnotenie odhadovanej hodnoty $V(s_t)$ v čase t bez znalosti finálnej odmeny R_t . Na úpravu odhadovanej hodnoty $V(s_t)$ v najjednoduchšej forme TD(0), stačí poznať odmenu r_{t+1} , nasledujúcu odhadovanú hodnotu $V(s_{t+1})$ a vypočítať rozdiel. Formálne výpočet môžeme zapísať ako:

$$V(s_t) \leftarrow V(s_t) + \alpha [r_{t+1} + \gamma V(s_{t+1}) - V(s_t)], \quad (3.11)$$

α je krokový parameter, ktorý ovplyvňuje veľkosť zmeny $0 \leq \alpha \leq 1$. Podobne môžeme zapísať aj TD pre akciovú-hodnotu $Q(s_t, a_t)$.

3.2.5.2 Q-učenie

Q-učenie je jeden z najznámejších a najpoužívanejších učiacich algoritmov RL. Algoritmus používa jednokrokový TD(0) na aktualizovanie aproximácie odhadovanej hodnoty $Q(s_t, a_t)$. Q-učenie priamo odhaduje optimálnu Q^* , to aj v prípade explorácie. $\max_a Q(s_{t+1}, a)$ vyberie najväčšiu odmenu r_{t+1} z nasledujúceho stavu s_{t+1} , aktualizácia Q-učenia bude :

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left[r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right].$$

Algoritmus Q-učenia (van Hasselt, [online]):

Inicializujeme $Q_0(s, a)$, pre všetky s, a

Vyberieme s_0

V každom kroku $t = 0, 1, 2, \dots$:

Odvodíme stratégiu π_t z Q_t (aj s exploraáciou)

Vyberieme a_t podľa π_t (pravdepodobnosť výberu a je $\pi_t(s_t, a)$)

Vykonáme a_t , zistíme r_t, s_{t+1}

Ak s_{t+1} je terminálny stav:

$$Q_{t+1}(s_t, a_t) \leftarrow Q_t(s_t, a_t) + \alpha_t(r_t - Q_t(s_t, a_t))$$

vyberieme nový stav s_{t+1} (ako začiatok pre novú epizódu)

inak

$$Q_{t+1}(s_t, a_t) \leftarrow Q_t(s_t, a_t) + \alpha_t[r_t + \gamma \max_a Q_t(s_{t+1}, a) - Q_t(s_t, a_t)]$$

3.2.5.3 SARSA učenie

SARSA oporiť Q-učenie nepotrebuje na aktualizáciu Q hodnôt najväčšiu hodnotu, ale zvolí v stave s_{t+1} ďalšiu akciu a_{t+1} podľa danej stratégie, a využije $Q_t(s_{t+1}, a_{t+1})$ na úpravu :

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r_{t+1} + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)].$$

Algoritmus pre SARSA učenie:

Inicializujeme $Q(s, a)$, pre všetky s, a

Vyberieme s_0

Odvodíme stratégiu π_0 z Q (aj s exploraáciou)

Vyberieme a_0 podľa π_0 (pravdepodobnosť výberu a je $\pi_0(s_0, a)$)

V každom kroku $t = 0, 1, 2, \dots$:

Vykonáme a_t , zistíme r_t , s_{t+1}

Ak s_{t+1} je terminálny stav:

vyberieme nový stav s_{t+1} (ako začiatok pre novú epizódu)

Odvodíme stratégiu π_t z Q (aj s exploraáciou)

Vyberieme a_{t+1} podľa π_t (pravdepodobnosť výberu a je $\pi_t(s_t + 1, a)$)

Ak s_{t+1} je terminálny stav:

$$Q_{t+1}(s_t, a_t) \leftarrow Q_t(s_t, a_t) + \alpha_t(r_t - Q_t(s_t, a_t))$$

inak

$$Q_{t+1}(s_t, a_t) \leftarrow Q_t(s_t, a_t) + \alpha_t[r_t + \gamma Q_t(s_{t+1}, a_{t+1}) - Q_t(s_t, a_t)]$$

3.2.5.4 Aktér-Kritik učiace metódy

Metódy typu Aktér-Kritik sú TD metódy, ktoré majú oddelenú štruktúru pre stratégiu a ohodnocovaciu funkciu. Štruktúra stratégie sa nazýva aktér, pretože vyberá akcie. Ohodnocovacia funkcia nazýva kritik, pretože kritizuje akcie zvolené aktérom. Kritika je forma TD chyby. Čo je skalárna hodnota z výstupu kritika a riadi všetko učenie v aktérovi aj kritikovi.

Typicky kritik je funkcia pre stavovú-hodnotu. Po každej zvolenej akcii, kritik zhodnotí nový stav a určí ako vhodná bola zvolená akcia. Pre výpočet TD chyby:

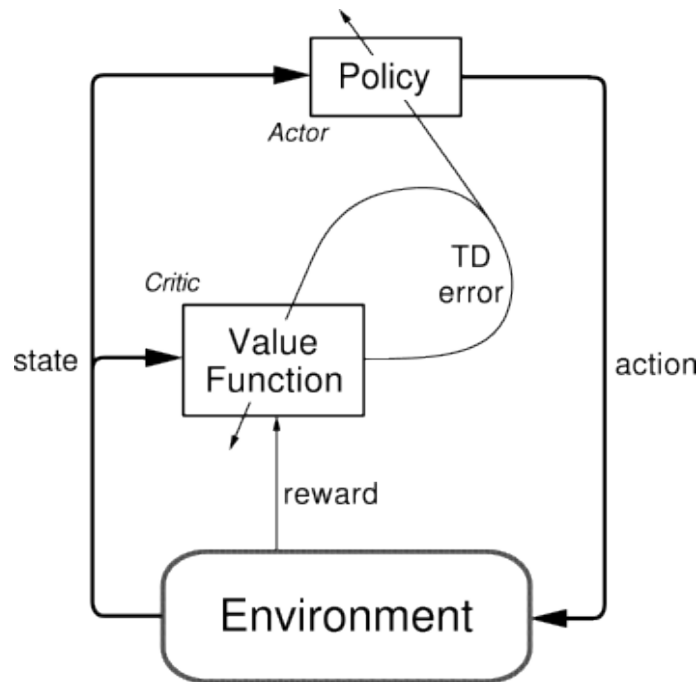
$$\delta_t = r_{t+1} + \gamma V(s_{t+1}) - V(s_t) \quad (3.12)$$

kde V je ohodnocovacia funkcia, čiže kritik. TD chyba je použitá pre zhodnotenie akcie a_t zvolenej v stave s_t . Ak je chyba kladná, preferencia výberu a_t by mala byť zvýšená, keď je TD chyba záporná mala by sa znížiť.

Preferenčná funkcia P , vráti preferenciu pre každú akciu:

$$P(s_t, a_t) = P_t(s_t, a_t) + \alpha \delta_t, \quad (3.13)$$

α je učiaci parameter.



Obrázok 3.4: Typická aktér-kritik architektúra (Sutton, Barto, 2004).

Inicializujeme $P_0(s, a)$, pre všetky s, a

Inicializujeme $V_0(s)$, pre všetky s

Vyberieme s_0

V každom kroku $t = 0, 1, 2, \dots$:

Odvodíme stratégiu π_t z P_t (aj s exploraáciou)

Vyberieme a_t podľa π_t (pravdepodobnosť výberu a je $\pi_t(s_t, a)$)

Vykonáme a_t , zistíme r_t, s_{t+1}

Ak s_{t+1} je terminálny stav:

$$P_{t+1}(s_t, a_t) \leftarrow P_t(s_t, a_t) + \alpha_t[r_t - V_t(s_t)]$$

$$V_{t+1}(s_t) \leftarrow V_t(s_t) + \beta_t[r_t - V_t(s_t)]$$

vyberieme nový stav s_{t+1} (ako začiatok pre novú epizódu)

inak

$$P_{t+1}(s_t, a_t) \leftarrow P_t(s_t, a_t) + \alpha_t[r_t + \gamma V_t(s_{t+1}) - V_t(s_t)]$$

$$V_{t+1}(s_t) \leftarrow V_t(s_t) + \beta_t[r_t + \gamma V_t(s_{t+1}) - V_t(s_t)]$$

3.2.5.5 QV-učenie

Tento typ učenia je možné uvažovať ako kombináciu predošlých algoritmov. Používa stavové-hodnoty vďaka čomu sa učí rýchlejšie, no odhaduje Q hodnoty a nie preferenčné hodnoty ako u aktér-kritik metódy.

$$Q_{t+1}(s_t, a_t) \leftarrow Q_t(s_t, a_t) + \alpha[r_t + \gamma V_t(s_t + 1) - Q_t(s_t, a_t)]$$

Algoritmus pre QV-učenie:

Inicializujeme $Q(s, a)$, pre všetky s, a

Inicializujeme $V(s)$, pre všetky s

Vyberieme s_0

V každom kroku $t = 0, 1, 2, \dots$:

Odvodíme stratégiu π_t z Q_t (aj s exploraáciou)

Vyberieme a_t podľa π_t (pravdepodobnosť výberu a je $\pi_t(s_t, a)$)

Vykonáme a_t , zistíme r_t, s_{t+1}

Ak s_{t+1} je terminálny stav:

$$Q_{t+1}(s_t, a_t) \leftarrow Q_t(s_t, a_t) + \alpha_t(r_t - Q_t(s_t, a_t))$$

$$V_{t+1}(s_t) \leftarrow V_t(s_t) + \beta_t(r_t - V_t(s_t))$$

vyberieme nový stav s_{t+1} (ako začiatok pre novú epizódu)

inak

$$Q_{t+1}(s_t, a_t) \leftarrow Q_t(s_t, a_t) + \alpha_t[r_t + \gamma V_t(s_{t+1}) - Q_t(s_t, a_t)]$$

$$V_{t+1}(s_t) \leftarrow V_t(s_t) + \beta_t(r_t + \gamma V_t(s_{t+1}) - V_t(s_t))$$

3.2.6 Spojité algoritmy učenia

3.2.6.1 CACLA

V klasických RL problémoch, sa využívajú MDP s konečnou množinou akcií. No v mnohých prípadoch riešenia s rozdelením priestoru do diskretných akcií nie sú vhodné. Napríklad v niektoré časti priestoru akcií môžu byť dôležitejšie ako iné,

vyžadujúce veľmi jemné rozdelenie priestoru pre dostatočne vhodné výsledky a spočiatku nemusí byť zrejmé, kde sú tieto oblasti. Taktiež rozdelenie na veľké množstvo akcií spôsobuje spomaľovanie učenia. Preto je potreba algoritmu schopného plne pracovať v spojitom priestore akcií. Takýmto algoritmom je Continuous Actor Critic Learning Automaton (CACLA)(van Hasselt, Wiering, 2007; van Hasselt, Wiering, 2009).

CACLA je postavená na princípe Aktér-Kritik metód. Pre problém spojitých priestorov sú využívané funkčné aproximátory (FA). Kritik ktorý ukladá očakávanú sumu diskontovaných odmien pre stavy je formou ohodnocovacej funkcie. Potom TD chyba δ_t je použitá na určenie vhodnosti a_t akcie. FA Ac ako aktér ktorý odhaduje funkciu $Ac^* : S \rightarrow A$, kde $Ac^*(s)$ označuje optimálnu akciu pre stav s . Aktér je upravený ak bola TD chyba pozitívna, teda vykonaná akcia bola lepšia ako očakávaná a mala by byť posilnená. Aktér je upravený smerom k tejto akcii, s tým že jeho výstup bude viac podobný naposledy vykonanej akcii v tomto stave.

$$Ac_{t+1}(s_t) \leftarrow Ac_t(s_t) + \alpha_t(a_t - Ac_t(s_t)), \text{ ak } \delta_t > 0, \quad (3.14)$$

táto zmena aktéra nastane iba vtedy keď a_t je rozdielne od $Ac_t(s_t)$, preto sa CACLA algoritmus môže naučiť nové akcie len pri explorácií. Exploatácia zvyšuje presnosť ohodnocovacej funkcie. CACLA nepoužíva hodnotu získanú z TD chyby, ale iba je použité znamienko z výslednej chyby. Ak je chyba záporná, stratégia nieje upravená.

Ak chceme pracovať v spojitom priestore stavov, na uloženie odhadovaných hodnôt stavov budeme potrebovať funkčné aproximátory. V tomto prípade nám poslúžia neurónové siete. Váhy siete sú teda parametre FA a θ^V nám označuje vektor parametrov FA. Na zmenu použijeme teda upravené pravidlo pre TD:

$$\theta_{i,t+1}^V = \theta_{i,t}^V + \alpha \delta_t \frac{\partial V_t(s_t)}{\partial \theta_{i,t}^V} \quad (3.15)$$

$\theta_{i,t}^V$ je i -ta zložka vektoru θ^V v čase t a $V_t(s)$ je výstup FA v čase t a vstupom s .

Ak máme výstup aktéra siete $Ac_t(s_t)$ v čase t a vektor parametrov váh θ^{Ac} , upravíme parametre nasledovne ak $\delta_t > 0$:

$$\theta_{i,t+1}^{Ac} = \theta_{i,t}^{Ac} + \alpha(a_t - Ac_t(s_t)) \frac{\partial Ac_t(s_t)}{\partial \theta_{i,t}^{Ac}} \quad (3.16)$$

Algoritmus pre CACLA:

Inicializujeme A_0 , pre všetky s

Inicializujeme V_0 , pre všetky s Vyberieme s_0

V každom kroku $t = 1, 2, 3, \dots$:

Vyberieme a_t s exploraáciou okolo $A_t(s_t)$

Vykonáme a_t , zistíme r_t, s_{t+1}

Ak je s_{t+1} terminálny

$$V_{t+1}(s_t) \leftarrow V_t(s_t) + \beta_t[r_t - V_t(s_t)]$$

vyberieme nový stav s_{t+1} (ako začiatok pre novú epizódu)

inak

$$V_{t+1}(s_t) \leftarrow V_t(s_t) + \beta_t[r_t + \gamma V_t(s_{t+1}) - V_t(s_t)]$$

Ak je $V_{t+1}(s_t) > V_t(s_t)$

$$A_{t+1}(s_t) \leftarrow A_t(s_t) + \alpha + t[(a_t) - A_t]$$

Kapitola 4

Navrhovaný model

Navrhovaný model vychádza z niektorých myšlienok z Nagai a spol. (2003). Tu navrhujú konštruktívny model kde robot nadobudne schopnosť zdieľanej pozornosti založený na vlastných vnútorných mechanizmoch vizuálnej pozornosti a učením so samovyhodnocovaním. Zo začiatku je dôležité aby sa robot pozrel na objekt vo svojom výhľade, neskôr sa naučil senzomotorickú koordináciu keď nastane vizuálna pozornosť. Keďže výsledok vizuálnej pozornosti sa nie vždy zhodoval s výsledkom zdieľanej pozornosti, robot mal správne a nesprávne informácie pre učenie zdieľanej pozornosti. Avšak robot postupne prestal uvažovať nesprávne informácie, pretože tieto informácie vzájomne nekorelovali so senzomotorickú koordináciou, kým správne informácie korelovali. Robot následne nadobudol schopnosť zdieľať pozornosť, nájdením korelácie v senzomotorickej koordinácii aj v prípade väčšieho množstva objektov náhodne rozmiestnených v okolitom prostredí. Podľa zistených výsledkov robot napodoboval vývinový proces zdieľanej pozornosti u dieťaťa.

V tejto práci prezentovaný model simuluje schopnosť zdieľať pozornosť použitím CACLA algoritmu. Model je značne zjednodušený a predpokladá už zahájenie ZP. Naučený model má schopnosť odpovedať na ZP v plne spojitom trojrozmernom prostredí. Model má rovnako ako ľudia obmedzený výhľad, teda môže vnímať iba iba časť z celého prostredia. Model agenta pracuje s predspracovanými abstrahovanými informáciami z vizuálneho vstupu, predloženými ako hodnoty stavov agenta. Priame vizuálne informácie ako napríklad snímky z kamery alebo obrázky tváre rodiča, nie sú v tomto modeli zahrnuté, no je ho možné o niektoré tieto časti jednoducho rozšíriť. Model narába priamo so súradnicami objektov v priestore a

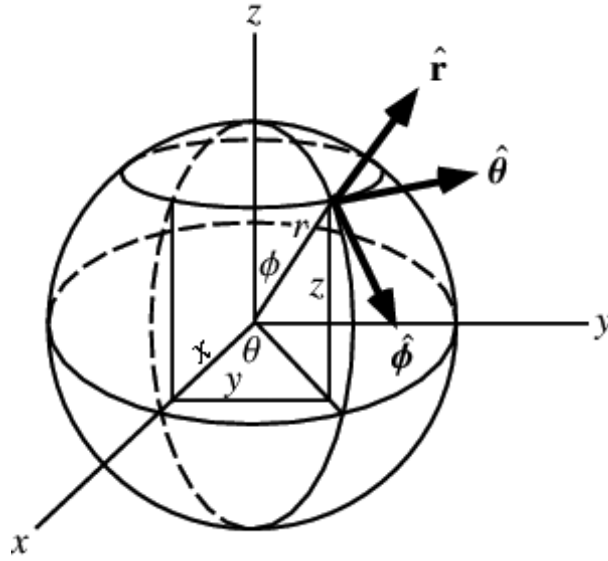
uhlami pohľadu agenta dieťaťa a rodiča vo sférickom súradnicovom systéme.

Smer pohľadu agenta v tomto modeli je určený dvoma uhlami, θ_p a θ_t , kde θ_p reprezentuje uhol horizontálneho natočenia agentovho pohľadu a θ_t vertikálny uhol natočenia, respektíve ich reprezentácie vo sférickom súradnicovom systéme. Teda pohľad agenta môžeme reprezentovať ako $\theta = [\theta_p, \theta_t]$. Rovnako je reprezentovaný tiež objekt $\mathbf{O} = [O_p, O_t]$. Hlavnou úlohou agenta je minimalizovať rozdiel medzi korešpondujúcimi uhlami pohľadu θ_p a uhlami určeného objektu \mathbf{O} , teda v ideálnom prípade $\theta_p - O_p = 0$ a $\theta_t - O_t = 0$.

Pohľad rodiča reprezentujú uhly rovnako ako v predošlom prípade. Teda pohľad rodiča je reprezentovaný uhlami, ktoré korešpondujú s jedným vybraným objektom na ktorý sa pozerá.

4.1 Model prostredia

Model pozostáva z agenta a objektov v priestore okolo neho v simulovanom prostredí. Súradnice agenta aj objektov sú reprezentované vektorom $\mathbf{s} = (p, t) \in \mathfrak{R}^2$ kde p, t sú uhly vo sférickom súradnicovom systéme (SSS). V klasickom karteziánskom systéme bod v priestore vieme reprezentovať vektorom $\mathbf{v} = (x, y, z) \in \mathfrak{R}^3$. V SSS môžeme akýkoľvek bod reprezentovať dvoma uhlami a polomerom (θ, φ, r) . Kde v našom prípade platí $-\pi < \varphi \leq \pi$, $0 \leq \theta \leq \pi$ a $r \geq 0$.



Obrázok 4.1: Sférický súradnicový systém. (Wolfram MathWorld, [online]).

Na prevody medzi súradnicovými sústavami sú použité nasledovné vzorce:

$$\begin{aligned}
 r &= \sqrt{x^2 + y^2 + z^2} \\
 \theta &= \sin^{-1}\left(\frac{z}{r}\right) \\
 \varphi &= \text{atan2}(y, x),
 \end{aligned}
 \tag{4.1}$$

Na prevod bodu $(x, y, z) \rightarrow (\theta, \varphi, r)$ z karteziánskej sústavy do sférickej a

$$\begin{aligned}
 x &= r \cos(\theta) \cos(\varphi) \\
 y &= r \cos(\theta) \sin(\varphi) \\
 z &= r \sin(\theta),
 \end{aligned}
 \tag{4.2}$$

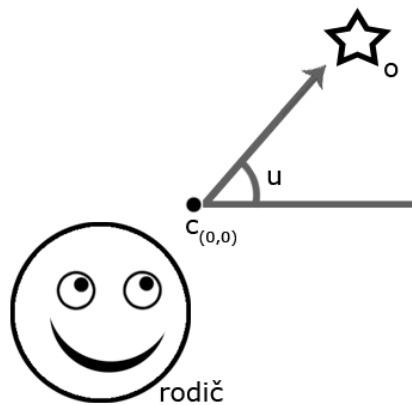
$(\theta, \varphi, r) \rightarrow (x, y, z)$ naopak. $\text{atan2}(x, y)$ je variácia arctan funkcie ktorá vracia kvadrant podľa x, y hodnôt.

V našom prípade objekt $\mathbf{O} = (O_p, O_t)$ môžeme povedať $p \leftarrow \theta, t \leftarrow \varphi$ a r je vzdialenosť od stredu SSS. Pozícia objektu \mathbf{O} v priestore bude vyjadrená pomocou súradníc $\mathbf{O} = (O_p, O_t, O_r)$, vzdialenosť r skúsime pri učení zanedbať a uvidíme ako bude model pracovať len s p, t súradnicami.

Kedže p, t sú v stupňoch, pre vhodnú implementáciu v CACLA, sú zobrazené do intervalu $-\frac{1}{2} \leq p, t \leq \frac{1}{2}$. Vzdialenosť 0,1 medzi takto danými dvoma uhlami potom odpovedá otočeniu o 36°

Model prostredia pozostáva z množiny M objektov \mathbf{O}_m , $m \in M$, rozmiestnených v okolí agenta. Agentovou úlohou je nájsť a natočiť sa na vybraný objekt, v tomto zjednodušenom prípade agentov vektor musí korešpondovať s predpočítaným vektorom ku objektu.

Pohľad rodiča reprezentuje taktiež vektor uhlov a to $\mathbf{y} = (y_p, y_t)$, no prezentovaný agentovi je ako jeden uhol, čo je jedna skalárna hodnota určujúca smer pohľadu rodiča. Tento uhol korešponduje s niektorým vybraným objektom. Rodič a agent sú umiestnení oproti sebe, kde dieťa agent je na počiatkovej pozícii $0, 0, 0 \in \mathbb{R}^3$ v oboch súradnicových systémoch, rodič je posunutý na x zložke v karteziánskom priestore, potom stred súradnicového systému je presunutý do tohoto bodu a podľa toho sú prevedené uhly rodiča v SSS. Uhol, ktorý reprezentuje informáciu o natočení rodiča, je jednoducho vypočítaný medzi objektom, horizontálnou osou v a stredom pohľadu agenta pomocou atan2 funkcie.



Obrázok 4.2: Reprezentácia rodičovho pohľadu ako uhol predkladaný agentovi.

obr. 4.2 Informácia rodiča zobrazená v pohľade agenta. $c_{(0,0)}$ určuje stred agentovho zorného poľa, o je objekt a uhol u , ktorý je medzi horizontálnou osou a vektorom bodu $c_{(0,0)}$, o . Tento uhol je prezentovaný ako vstup pre agenta.

Odmena pre agenta je vypočítavaná na základe euklidovskej vzdialenosti stre-

du pohľadu agenta \mathbf{x} (vektor agenta) a konkrétneho objektu \mathbf{O}_m (vektor daného objektu). Odmena je vypočítaná podľa nasledovného vzorca:

$$r = \frac{1}{\sqrt{(x_p - O_{mp}) + (x_t - O_{mt})}}.$$

4.2 Model agenta

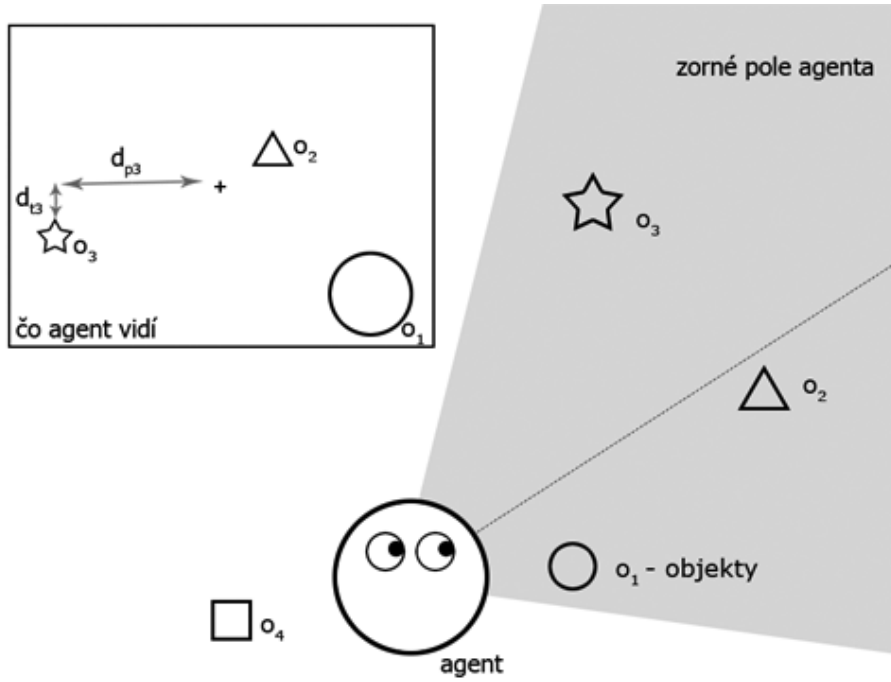
Model agenta (MA) môžeme opísať nasledovne: na začiatku je smer pohľadu agenta náhodne natočený, agent sa pozerá do priestoru s určitým vymedzeným výsekom ako zorným poľom. V celom priestore sú náhodne rozložené objekty. Agentovi je vybraný jeden objekt, ktorý má v zornom poli. Následne sa sa musia pomocou RL algoritmu vybrať vhodné akcie, pomocou ktorých agent natočí svoj pohľad ku danému objektu.

Model agenta pracuje so spojitými akciami a stavmi, je teda reprezentovaný pomocou algoritmu CACLA. Algoritmus je reprezentovaný dvoma trojvrstvovými doprednými neurónovými sieťami s hyperbolickou aktivačnou funkciou \tanh , pre aktéra aj kritika. Sieť pre aktéra disponuje len jedným vstupom-stavom, ktorý je ovplyvnený $x = 2$ súradnicami reprezentujúce vektor uhlov \mathbf{x} aktuálneho natočenia pohľadu agenta. Tento vstup $y = 1$ vstup nahrádza informáciu pohľadu rodiča, ako natočenie hlavy alebo jeho pozícia očí, abstrahovanú na uhol y pohľadu rodiča. Vstup y je stav RL v rozsahu i , pre i platí $-\frac{1}{2} \leq i \leq \frac{1}{2}$. Teda stav v prostredí s , $s \in \mathfrak{R}$ je daný ako $s = y$

Sieť aktéra má určené 2 výstupy na ktorých sieť vracia informáciu o zvolených akciách. Tieto akcie určujú zmenu uhlov natočenia agenta. Akcie vybrané aktérom sú vždy vyhodnotené kritikom, ak sú vyhodnotené ako lepšie, stavy sú patrične upravené. Akcie sú definované ako vektor uhlov $\mathbf{A} \in \mathfrak{R}^2$. Pomocou \mathbf{A} sú modifikované vektory \mathbf{x} v čase t :

$$\mathbf{x}_{t+1} = \mathbf{x}_t + \mathbf{A}_t,$$

kde \mathbf{x} je vektor súradníc aktuálnej pozície agenta.



Obrázok 4.3: Zobrazenie agenta v prostredí.

Na obr. 4.3 je reprezentácia agenta v prostredí, na ľavej strane je zobrazený výhľad agenta. Ak objekt o_3 je hľadaný objekt, potom vzdialenosti d_{p3} a d_{t3} sú použité na výpočet odmeny.

Model vyváženého exploračného-exploatačného je riešený pomocou Gaussovej exploračnej v okolí aktuálnej aproximácie optimálnej akcie.

Ak $A_t(s_t)$ označuje akciu na výstupe aktéra v čase t , pravdepodobnosť výberu akcie a bude:

$$\pi_t(s_t, a) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(a - A_t(s_t))^2 / (2\sigma^2)} \quad (4.3)$$

Pretože agent vykonáva v jednom kroku dve akcie zároveň, definované ako dvo-prvkový vektor, je každá akcia vyberaná zvlášť.

4.3 Učenie

Učenie je proces natrénovania agentových CACLA sietí aktéra a kritika, tak aby boli schopné nájsť a natočiť sa na ľubovoľný objekt v priestore podľa pohľadu rodiča. Pre tento model máme zostavenie, kde agent a rodič sú umiestnený oproti sebe v rovnakej výške a nehýbu sa. Na začiatku učenia aj testovania sú vytvorené objekty, tie sú rozložené v rovine okolo agenta, zároveň sú predpočítané uhly medzi objektom a agentom aj objektom a rodičom aby nemuseli byť neustále prepočítavané počas simulácií. Uhly medzi objektom a agentom poslúžia na výpočet odmeny pri učení. Trénovanie agenta bolo rozdelené na 1000 epizód. Na začiatku bolo v okolitom priestore rozmiestnených náhodne vygenerovaných 100 objektov. Objekty boli reprezentované ako body v priestore a počas učenia sa nachádzali stále na rovnakej pozícií. Na začiatku každej epizódy bol pohľad agenta nastavený niektorým náhodným smerom, zároveň bol náhodne vybraný jeden z množiny agentom aktuálne viditeľných objektov, ktorý bol prezentovaný ako cieľový. Agentov vstup bol nastavený podľa aktuálneho pohľadu a pozície objektu v prostredí. Siete boli vyresetované a úlohou bolo naučiť tieto CACLA siete natáčať sa ku objektom, teda vyberať akcie tak, aby viedli ku natočeniu k objektu. Agent sa učil z náhodnej pozície natočiť sa na správny objekt. Vždy keď sa mu to podarilo, bol odmenený zvýšenou odmenou a epizóda bola ukončená. Nasledovala nová epizóda s náhodne nastaveným pohľadom a vybraným objektom a proces sa opakoval. Takto bol trévaný CACLA modul až do poslednej epizódy.

Takto sa sieť učí vo všetkých epizódach. V takto určených podmienkach, bolo agentovou úlohou nájsť a vytvoriť si vhodnú reprezentáciu sietí, tak aby ním zvolené akcie viedli z náhodného stavu smerom ku vybranému objektu, minimalizovaním vzdialenosti medzi vektorom objektu a vlastným vektorom natočenia (pohľadu). Pre učenie bola požadovaná euklidovská vzdialenosť agenta a objektu nastavená na hodnotu 0,0005, ak sa agent dostal do vzdialenosti rovnej či menšej, dostal väčšiu odmenu a epizóda učenia bola ukončená. Následne sa začala nová epizóda s novým objektom a učenie pokračovalo až pokiaľ neskončili všetky epizódy. Po dotrénovaní bol agent testovaný na nových objektoch neobsiahnutých v trérovacej množine. Výsledky je možné nájsť v sekcii vyhodnotenie.

Ako prebieha učenie CACLA modulu je možné opísať nasledovne: ak sa agent

nachádza v stave s_t , na vstupoch má informácie o svojom natočení, objektoch a pohľade rodiča. Aktér v tomto stave vyberie najvhodnejšiu akciu spolu s gaussovskou exploračiou a_t , teda podľa zvolenej stratégie, čo znamená že dopredným šírením informácie zo vstupov daných stavom s_t , získa na výstupe vhodnú akciu. Agent tieto akcie vykoná a od prostredia dostane odmenu r_t . Agent sa teraz nachádza v stave s_{t+1} . Kritik zo svojej siete zistí odhadovanú hodnotu v s_{t+1} a podľa odmeny r_t upraví svoje parametre váh spätným šírením chyby pre stav s_t . Ak je nová odhadovaná hodnota s_t lepšia ako predošlá, sieť aktéra je upravená smerom ku akciám a_t . Následne agent v stave s_{t+1} zvolí akciu, vykoná exploračiu a dostane akciu a_{t+1} , po jej vykonaní sa ocitne v stave s_{t+2} a siete sú zase náležite upravené. Ak sa agent dostane do terminálneho stavu s_T , dostane odmenu r_T , plus ďalšiu malú odmenu, za to že sa dostal do cieľa a epizóda je ukončená. Učenie sa zopakuje pre nasledujúce objekty a pokračuje až do konca všetkých epizód. Pri testovaní sa postupuje rovnako ale bez úprav sietí aktéra a kritika, taktiež je vynechaný krok exploračie, keďže je nepotrebný.

V nasledujúcej tabuľke (tab. 4.1) sú zhrnuté spoločné parametre pre učenie a testy modelu.

Tabuľka 4.1: Tabuľka parametrov pre model.

| Názov | Označenie | Hodnota |
|----------------------|-----------|---------|
| počet objektov | M | 100 |
| aktér sieť | A_n | 3 |
| kritik sieť | C_n | 4 |
| aktér rýchly uč. | α | 0,07 |
| kritik rýchly uč. | β | 0,09 |
| diskont | γ | 0,8 |
| exploračný parameter | σ | 0,01 |
| zorné pole agenta | ρ | 0,2 |
| akcie-výstupy | A | 2 |
| stavy-vstupy | y | 1 |

Vysvetlenie ku všetkým parametrom z tab. 4.1. M určuje počet všetkých objektov

rozmiestnených v prostredí použitých pri učení. Množstvo neurónov v skrytej vrstve neurónovej siete aktéra a kritika označujeme ako A_n , respektíve C_n . Učiace alebo krokové parametre pre siete aktéra a kritika sú α pre sieť aktéra a β pre sieť kritika. γ určuje diskontný faktor, ktorý upravuje odhadovanú hodnotu budúcich odmien. Exploračný parameter σ je štandardná odchýlka určujúca pravdepodobnosť výberu novej akcie pri gaussovskej explorácii. Zorné pole alebo výhľad agenta je polomer kruhového výseku od agentovho vektoru ρ z celkového priestoru, v ktorom agent môžeme vnímať objekty. V prípade že $\rho > 0,5$, agent vidí všetky objekty v prostredí, hodnota $\rho = 0,1$ odpovedá zornému polu v rozsahu 72° . Počet akcií A ktoré agent vykoná v každom kroku, teda aj počet výstupov zo siete aktéra. Vstup y určený uhlom pohľadu rodiča.

4.4 Zostavenie a vyhodnotenie

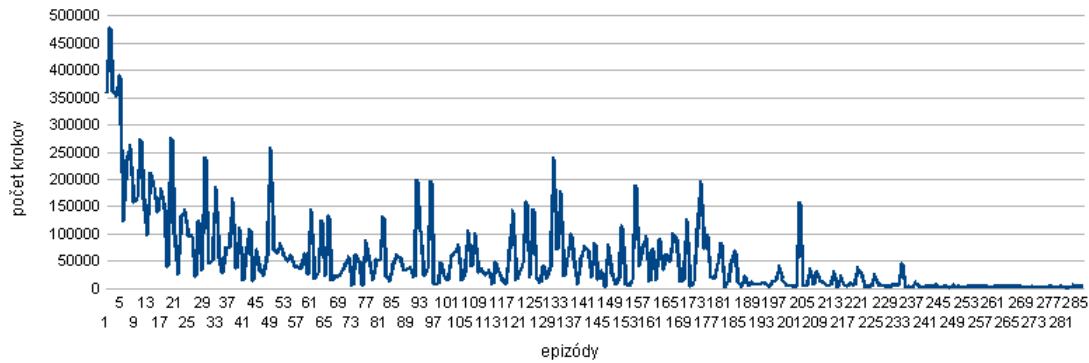
Samotný model je implementovaný v jazyku c++, pomocou ktorého bola vytvorená aplikácia, ktorá zahrňovala všetky časti modelu. S použitím tejto aplikácie boli vyhotovené aj všetky tu uvedené testy.

Pre vyhodnotenie modelu sa testovalo na 1000 náhodne rozmiestnených objektoch v priestore, vždy po 1000 epizódach učenia. Za úspešné vytvorenie ZP sa pokladalo ak bola euklidovská vzdialenosť uhlov agenta od objektu menšia ako 0,02, čo zodpovedalo ich rozdielu o $7,2^\circ$. Podľa takto definovaných nastavení boli dosiahnuté veľmi dobré výsledky v úspešnosti nachádzania objektov v trojrozmernom spojitom prostredí, zobrazené v tab. 4.2.

Tabuľka 4.2: Priemerná percentuálna úspešnosť vzniku zdieľanej vizuálnej pozornosti na nových tisíc testovaných objektoch po každom z 20 učení.

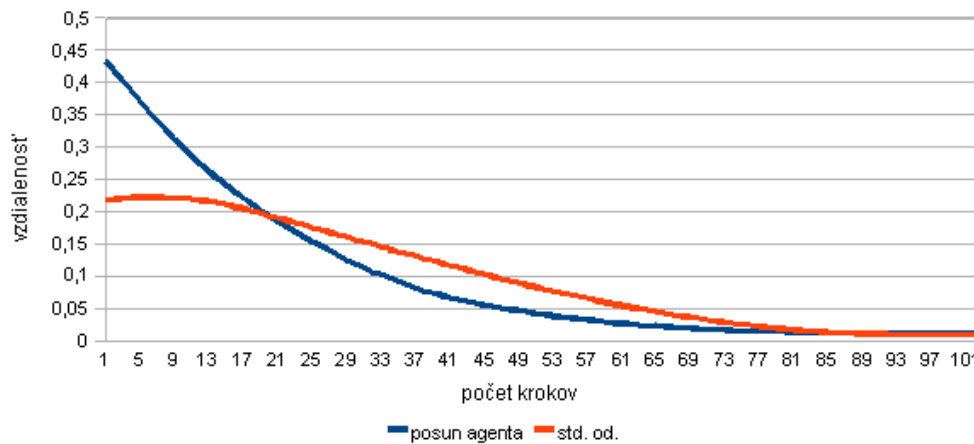
| | v zor. poli | mimo, zor poľa |
|-----------------|-------------|----------------|
| % úspešnosti ZP | 86,13 | 86,01 |
| std. od. | 10,09 | 8,99 |

tab. 4.2 tiež ukazuje že agent sa naučil rovnako dobre nachádzať aj objekty ktoré sú spočiatku mimo jeho zorného poľa.



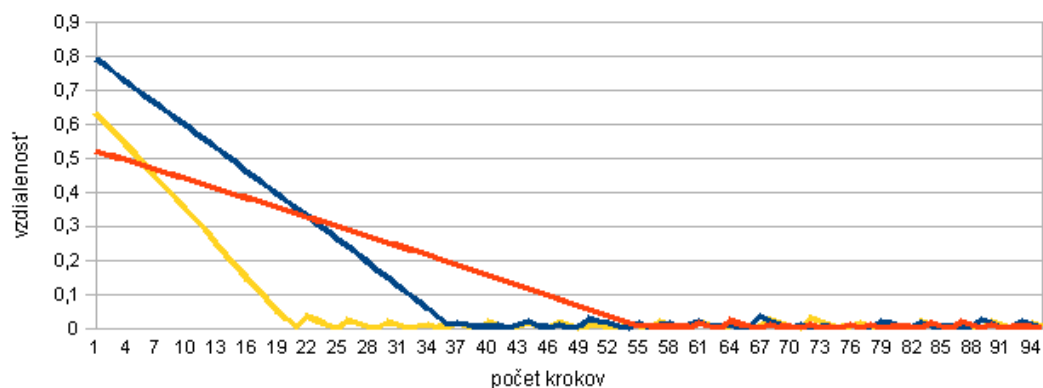
Obrázok 4.4: Počet krokov potrebných ku nájdeniu objektov počas prvých 300 učiacich epizód.

obr. 4.4 ukazuje priemerný počet krokov ktoré agent musel vykonať keď sa učil otáčať ku objektom. Graf znázorňuje ako sa agent učil otáčať ku daným sto objektom. Je možné vidieť, že po nájdení každého objektu dvakrát, už bol schopný bez problémov sa naň natočiť z hociakej pozície.



Obrázok 4.5: Priemerný počet krokov pri testovaní.

obr. 4.5 znázorňuje priemerný počet krokov do nájdenia cieľa pri testovaní na náhodných objektoch. Hodnoty sú priemerom z 200 testovaní na rôznych objektoch.



Obrázok 4.6: Pohyb ku trom vybraným objektom.

obr. 4.6 zobrazuje zmenu vzdialenosti pri testovaní konkrétnych troch objektov. Ukazuje s, a že ku niektorým objektom sa agent vie natočiť rýchlejšie.

Z daných výsledkov je možné tvrdiť, že agent sa dokázal naučiť plynulo otáčať za rôznymi objektami, ktoré môžu byť umiestnené kdekoľvek v jeho okolí, len na základe jednej hodnoty zodpovedajúcej uhlu natočenia v 2D reprezentácií. Čo by sme mohli prirovnať ku abstrahovanej reprezentácií procesu zdieľanej pozornosti.

Kapitola 5

Záver

Existuje množstvo experimentov a pozorovaní vývinu zdieľanej pozornosti, no stále chýbajú teoretické modely, ktoré by boli schopné hodnoverne vysvetliť procesy počas vývinu. Nové modely s použitím neurovedeckých poznatkov v spojení s výpočtovými prostriedkami v súčasnosti umožňujú odhaľovať vodítka k poznaniu ľudských schopností. Zároveň existujú snahy o tvorbu interaktívnych robotov, ktorí by boli schopní vnímať svet a učiť sa z prostredia podobne ako ľudia. Preto sa aj táto práca zaoberala problémom zdieľanej pozornosti od jej vývoja z psychologického hľadiska až po navrhnutie a vytvorenie nového výpočtového modelu, ktorý umožňuje modelovať zjednodušený problém zdieľanej vizuálnej pozornosti v spojitom priestore, ktorý sa učil na základe biologicky relevantného spôsobu učenia, bez vplyvu akéhokoľvek externého učiteľa, teda len na základe interakcií s prostredím. Tento model dosahuje dobré výsledky, no je však značne zjednodušený v mnohých ohľadoch, avšak v budúcnosti by mohol byť vylepšený a vďaka jeho schopnosti pracovať v spojitom priestore môže byť efektívne využitý v reálnych podmienkach.

Literatúra

- [1] Baldwin, D. A. (1995). Understanding the link between joint attention and language. In C. Moore, P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 131-158). Hillsdale, NJ: Lawrence Erlbaum.
- [2] Baron-Cohen (1992). Precursors to a Theory of Mind: Understanding Attention in Others. In: A. Whiten (Ed.): *Natural Theories of Mind: Evolution, Development and simulation of Everyday Mindreading*. Oxford, Basil Blackwell.
- [3] Baron-Cohen, S. (1995). The eye direction detector (EDD) and the shared attention mechanism (SAM): Two cases for evolutionary psychology. In C. Moore, P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 41-59). Hillsdale, NJ: Lawrence Erlbaum.
- [4] Baron-Cohen, S., Baldwin, D. A., Crowson, M. (1997). Do children with autism use speaker's direction of gaze strategy to crack the code of language? *Child Development*, 68, 58-57.
- [5] Bates, E., Benigni, L., Bretherton, I., Camaioni, L., Volterra, V. (1979). Cognition and communication from 9-13 months: Correlational findings. In E. Bates (Ed.), *The emergence of symbols: Cognition and communication in infancy* (pp. 69-140). New York: Academic Press.
- [6] Bruner, J. S. (1977). Early social interaction and language acquisition. In H. R. Schaffer (Ed.), *Studies in mother-infant interaction* (pp. 271-289). New York: Academic Press.

- [7] Butterworth, G., Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, 9, 55-72.
- [8] Butterworth, G. (1995). Origins of mind in perception and action. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development*. Hillsdale, NJ: Erlbaum.
- [9] Cole, M., Cole, S. (1996). *The development of children* (3rd edn.). New York: Freeman.
- [10] Deák, G. & Triesch, J. (2006). Origins of shared attention in human infants. Fujita, S. Itakura (Ed.). *Diversity of cognition*. University of Kyoto Press.
- [11] D'Entremont, B., Hains, S., Muir, D. (1997). A demonstration of gaze following in 3- to 6-month-olds. *Infant Behavior and Development*, 20 (4), 569–572.
- [12] Doniec, M.W. Ganghua Sun, Scassellati, B. (2006) *Active Learning of Joint Attention*. Dept. of Comput. Sci., Yale Univ., New Haven, CT;
- [13] Gomez, F. (1990). The emergence of intentional communication as a problem solving strategy in the gorilla. In S. T. Parker & K. R. Gibson (Ed.), *“Language” and intelligence in monkeys and apes: Comparative developmental perspectives* (pp. 333-335). New York: Cambridge University Press.
- [14] Haith, M.M., Bergman, T., Moore, M.J. (1979). Eye contact and face scanning in early infancy. *Science*, 198, 853–855.
- [15] Kasari, C., Freeman, S. F. N., Paparella, T. (2001). Early intervention in autism: Joint attention and symbolic play. *International Review of Research In Mental Retardation*, 23, 203-237.
- [16] Knudsen, E. I. (2007) Fundamental Components of Attention. *Annual Review of Neuroscience* Vol. 30: 57-78 (Volume publication date July 2007).
- [17] Kvasnička V., Beňušková., Pospíchal J., Farkaš I., Tiňo P. a Kráľ A. (1997). *Úvod do teórie neurónových sietí.*; Iris: Bratislava.

- [18] Mundy, P., Gomes, A. (1998). Individual differences in joint attention skill development in the second year. *Infant Behavior and Development*, 21, 469-482.
- [19] Nagai, Y., Hosoda K., Morita A., Asada, M. (2003) A constructive model for the development of joint attention. *Connection Science*, vol. 15, no. 4, pp. 211-229, December 2003.
- [20] Rybár, J., Beňušková, L. Kvasnička, M. (2002). *Kognitívne vedy*. Bratislava, Kalligram.
- [21] Schaffer, H. R. (1984). *The child's entry into a social world*. New York: Academic Press.
- [22] Sutton R.S., Barto, A.G. (2004). *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 1998 A Bradford Book.
- [23] Tomasello, M., Farrar, M. J. (1986). Joint attention and early language. *Child Development*, 57, 1454-1463.
- [24] Triesch, J., Teuscher, Ch., Deák, G., Carlson, E. (2006) Gaze following: why (not) learn it?; *Developmental Science* 9:2 (2006), pp. 125–157
- [25] Ulvund, S. E., Smith, L. (1996). The predictive validity of nonverbal communicative skills in infants with prenatal hazards. *Infant Behavior and Development*, 19, 441-449.
- [26] van Hasselt. H, Wiering, M. A. (2009) Using continuous action spaces to solve discrete problems. *Proceedings of the 2009 international joint conference on Neural Networks*; Atlanta, Georgia, USA; Pages: 1144-1151; Year of Publication: 2009; 1098-7576.
- [27] van Hasselt. H, Wiering, M. A. (2007). Reinforcement learning in continuous action spaces. *In Proceedings of the IEEE International Symposium on Adaptive Dynamic Programming and Reinforcement Learning*, p. 272–279, .
- [28] Woodward, A.L. (2003). Infants' developing understanding of the link between looker and object. *Developmental Science*, 6 (3), 297–311.

- [29] Yücel, Z., Salah, A. A. (2009) Head Pose and Neural Network Based Gaze Direction Estimation for Joint Attention Modeling in Embodied Agents. *The Annual Meeting of the Cognitive Science Society, CogSci 2009*.
- [30] Yucel, Z., Salah, A. A., Mericli, C., Mericli, T. (2009) Joint Visual Attention Modeling for Naturally Interacting Robotic Agents; *International Symposium on Computer and Information Sciences*;
- [31] van Hasselt, H., A Short Introduction To Some Reinforcement Learning Algorithms [online]. [cit. 2010-04-07]. Dostupné na WWW: <http://people.cs.uu.nl/hado/rl_algs/rl_algs.html>
- [32] de Smith M., J., Goodchild M., F., Longley P., A., 2007 *Geospatial Analysis : A Comprehensive Guide to Principles, Techniques and Software Tools*. 3rd edition. Dostupné na WWW: <<http://www.spatialanalysisonline.com/output/>>
- [33] Abazov, V. M. and Abbott, B. and Abolins, M. and Acharya, B. S. and Adams, M. and Adams, T. and Aguilo, E. and Ahsan, M. and Alexeev, G. D. and Alkhazov, G. and Alton, A. and Alverson, G. and Alves, G. A. and Ancu, L. S. and Andeen, T. and Anzelc, M. S. and Aoki, M. and Arnoud, Y. and Arov, M. and Arthaud, M. and Askew, A. and Åsman, B. and Atramentov, O. and Avila, C. and BackusMayes, J. and Badaud, F. and Bagby, L. (2009) PhysRevLett.103.092001. *Observation of Single Top-Quark Production* [online], Phys. Rev. Lett., 103,9, 092001, American Physical Society. Upravené September 30, 2009 [cit. 2010-05-02]. Dostupné na WWW: <http://www-d0.fnal.gov/Run2Physics/top/singletop_observation>
- [34] Christos Stergiou & Dimitrios Siganos. NEURAL NETWORKS [online]. [cit. 2010-04-25]. Dostupné na WWW: <http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html>
- [35] Wolfram MathWorld. Spherical Coordinates [online]. [cit. 2010-04-25] Dostupné na WWW: <<http://mathworld.wolfram.com/SphericalCoordinates.html>>

Prílohy

Obsah CD

K tejto práci je priložené CD, na ktorom sú uložené zdrojové kódy a aplikácia umožňujúca otestovať model. Na CD je textový súbor `readme.txt`, ktorý obsahuje informácie pre použitie aplikácie. Štruktúra adresárov na CD je nasledujúca:

Adresár [Bin] obsahuje spustiteľné súbory.

Adresár [Src] obsahuje zdrojové kódy aplikácie.

Adresár [Text] obsahuje elektronickú verziu tejto práce.