



УНИВЕРЗИТЕТ У НОВОМ САДУ
ФАКУЛТЕТ ТЕХНИЧКИХ
НАУКА У НОВОМ САДУ



Срђан Савић

**Когнитивно инспирисани
рачунарски модел меморије са
применама у роботици**

ДОКТОРСКА ДИСЕРТАЦИЈА

Нови Сад, 2018.



УНИВЕРЗИТЕТ У НОВОМ САДУ • ФАКУЛТЕТ ТЕХНИЧКИХ НАУКА
21000 НОВИ САД, Трг Доситеја Обрадовића 6

КЉУЧНА ДОКУМЕНТАЦИЈСКА ИНФОРМАЦИЈА

Редни број, РБР:	
Идентификациони број, ИБР:	
Тип документације, ТД:	Монографска документација
Тип записа, ТЗ:	Текст, штампани материјал
Врста рада, ВР:	Докторска дисертација
Аутор, АУ:	Срђан Савић
Ментор, МН:	др Милан Ђатовић, доцент
Наслов рада, НР:	Когнитивно инспирисани рачунарски модел меморије са применама у роботизи
Језик публикације, ЈП:	српски
Језик извода, ЈИ:	српски/енглески
Земља публикавања, ЗП:	Република Србија
Уже географско подручје, УГП:	АП Војводина
Година, ГО:	2018.
Издавач, ИЗ:	Ауторски репринт
Место и адреса, МА:	Факултет техничких наука, 21000 Нови Сад, Трг Доситеја Обрадовића 6
Физички опис рада, ФО: <small>(поглавља/страна/ цитата/табела/слика/графика/прилога)</small>	6/118/108/5/22/0/0
Научна област, НО:	Мехатроника
Научна дисциплина, НД:	Мехатроника, роботика и аутоматизација
Предметна одредница/Кључне речи, ПО:	дуготрајна меморија, радна меморија, когнитивни систем, интеракција човека и робота, фокусно стабло
УДК	
Чува се, ЧУ:	Библиотека Факултета техничких наука, Универзитет у Новом Саду
Важна напомена, ВН:	



УНИВЕРЗИТЕТ У НОВОМ САДУ • ФАКУЛТЕТ ТЕХНИЧКИХ НАУКА
21000 НОВИ САД, Трг Доситеја Обрадовића 6

КЉУЧНА ДОКУМЕНТАЦИЈСКА ИНФОРМАЦИЈА

Извод, **ИЗ:**

У дисертацији је представљен нови рачунарски модел дуготрајне меморије, намењен за примене у конверзационим роботским агентима. Предложени модел је симболички, са методолошког аспекта, и инспирисан је изабраним когнитивним механизмима људског меморијског система, који укључују интеграцију менталних репрезентација, семантичку категоризацију, асоцијативно учење и контекстно зависно селектовање информација. У основи модела се налази симболички приступ за аутоматско моделовање домена интеракције између човека и робота. Релевантни функционални аспект предложеног модела односи се на проблеме адекватног активирања делова дуготрајне меморије, у складу са спољашњим стимулансима, историјом интеракције и тренутним контекстом интеракције. Ниво апстракције у спецификацији модела је довољан да омогући примену модела у широком спектру просторних, униформних домена који су карактеристични за интеракцију између човека и робота, а ниво детаља у спецификацији је довољан за рачунарску имплементацију модела.

Датум прихватања теме, **ДП:**

11.5.2017.

Датум одбране, **ДО:**

Чланови
комисије,
КО:

Председник:

др Бранислав Боровац,
редовни професор

Члан:

др Мирко Раковић, доцент

Члан:

др Милутин Николић, доцент

Члан:

др Коста Јовановић, доцент

Члан, ментор:

др Милан Ђњатовић, доцент

Потпис ментора

Образац **Q2.HA.06-05**- Издање 1



UNIVERSITY OF NOVI SAD ● FACULTY OF TECHNICAL
SCIENCES

21000 NOVI SAD, Trg Dositeja Obradovića 6

KEY WORDS DOCUMENTATION

Accession number, ANO :	
Identification number, INO :	
Document type, DT :	Monographic documentation
Type of record, TR :	Text, printed material
Contents code, CC :	PhD thesis
Author, AU :	Srđan Savić
Mentor, MN :	Milan Gnjatović, PhD, assistant professor
Title, TI :	Cognitively Inspired Computational Memory Model with Applications in Robotics
Language of text, LT :	Serbian
Language of abstract, LA :	Serbian/English
Country of publication, CP :	Republic of Serbia
Locality of publication, LP :	AP of Vojvodina
Publication year, PY :	2018.
Publisher, PB :	Author reprint
Publication place, PP :	Faculty of Technical Sciences, 21000 Novi Sad, Trg Dositeja Obradovića 6
Physical description, PD : (chapters/pages/ref./tables/pictures/graphs/app endixes)	6/118/108/5/22/0/0
Scientific field, SF :	Mechatronics
Scientific discipline, SD :	Mechatronics, Robotics and Automation
Subject/Key words, S/KW :	long-term memory, working memory, cognitive system, human-robot interaction, focus tree
UC	
Holding data, HD :	Library of the Faculty of Technical Sciences
Note, N :	



UNIVERSITY OF NOVI SAD ● FACULTY OF TECHNICAL
SCIENCES

21000 NOVI SAD, Trg Dositeja Obradovića 6

KEY WORDS DOCUMENTATION

Abstract, **AB**:

This dissertation proposes a novel computational model of long-term memory intended for applications in conversational robotic agents. The proposed model is symbolic, from the methodological point of view, and cognitively-inspired by selected cognitive mechanisms of the human memory system, including integration of mental representations, semantic categorization, associative learning, and context-dependent information selection. In the core of the model there is a symbolic approach to automatic modeling of domains of human-robot interaction. The relevant functional aspect of the proposed model concerns the problems of context-dependent retrieval from long-term memory, in accordance with external stimuli, the interaction history, and the current context of interaction. The level of abstraction in the model is sufficient to enable generalization of the model over a range of spatial, uniform domains that are characterical for human-robot interaction, while the level of detail contained in the specification of the model is sufficient for a computational implementation.

Accepted by the Scientific Board on,
ASB:

2017/05/11

Defended on, **DE**:

Defended Board,
DB:

President:

Branislav Borovac, PhD,
full professor

Member:

Mirko Raković, PhD,
assistant professor

Member:

Milutin Nikolić, PhD,
assistant professor

Member:

Kosta Jovanović, PhD,
assistant professor

Member,
Mentor:

Milan Gnjatović, PhD,
assistant professor

Menthor's sign

Сажетак

У дисертацији је представљен нови рачунарски модел дуготрајне меморије, намењен за примене у конверзационим роботским агентима. Предложени модел је инспирисан изабраним когнитивним механизмима људског меморијског система, који укључују интеграцију менталних репрезентација, семантичку категоризацију, асоцијативно учење и контекстно зависно селектовање информација.

У основи модела се налази симболички приступ за аутоматско моделовање домена интеракције између човека и робота. Кроз процесе семантичке категоризације и асоцијативног учења, систем на основу почетног, непотпуног скупа пропозиција, класификује пропозиције у односу на домене којима припадају и изводи недостајуће знање о доменима. Резултати експерименталне валидације прототипског система заснованог на предложеном моделу показују високу успешност предложеног механизма учења, као и његову скалабилност.

Адекватна пажња је посвећена и потребама практичних имплементација модела које укључују велике количине података у дуготрајној меморији. Релевантни функционални аспект предложеног модела односи се на проблеме адекватног активирања делова дуготрајне меморије, у складу са спољашњим стимулансима, историјом интеракције, и тренутним контекстом интеракције.

На методолошком нивоу, предложени модел је симболички, што га чини следљивим, а резултати извршавања предложених алгоритама могу се објаснити у оквиру претпостављених концепата на којима се заснива модел. Ниво апстракције у спецификацији модела довољан је да омогући примену модела у широком спектру просторних, униформних домена који су карактеристични за интеракцију између човека и робота, а ниво детаља у спецификацији је довољан за рачунарску имплементацију модела.

Abstract

This dissertation proposes a novel computational model of long-term memory intended for applications in conversational robotic agents. The proposed model is cognitively-inspired by selected cognitive mechanisms of the human memory system, including integration of mental representations, semantic categorization, associative learning, and context-dependent information selection.

In the core of the model there is a symbolic approach to automatic modeling of domains in human-robot interaction. Based on an incomplete set of propositions, the system classifies the propositions with respect to the interaction domain they belong, and derives the missing propositions. The experimental validation of the prototype system shows high efficacy and scalability of the proposed learning algorithm.

Appropriate attention is devoted to the requirements for implementations of the model that relate to big-data storage and processing in long-term memory. The relevant functional aspect of the proposed model concerns the problems of context-dependent retrieval from long-term memory, in accordance with external stimuli, the interaction history, and the current context of interaction.

At the methodological level, the proposed model is symbolic, which makes it computationally and analytically tractable, with explanatory power. The level of abstraction in the model is sufficient to enable generalization of the model over a range of spatial, uniform domains that are characterical for human-robot interaction, while the level of details contained in the specification of the model is sufficient for a computational implementation.

Захвалница

Желео бих овом приликом да захвалим свом ментору др Милану Гњайновићу на указаном поверењу, подстицају, корисним саветима и несебичном дељењу знања током израде ове дисертације. Захвалан сам што ме је приближио свету научне и критичке мисли.

Посебну захвалност дујем својим родитељима који су увек били уз мене и који су ми пружили моћност и привилегију да бирам свој животињи позив и пут. Хвала им што су испрајно били уз мене на том путу, свих ових година и помогли ми да не скренем са њега.

Највећу захвалност дујем својој сурузи Соњи на бескрајном стрпљењу, безрезервној подршци, разумевању и љубави, што ми је уливало снаге и подстицало ме да испрајем.

Посвећено мојим родитељима

Садржај

Сажетак	I
Abstract	II
Захвалница	III
Садржај	V
Списак слика	VII
Списак табела	X
Увод	1
1.1 Потреба за истраживањем.....	2
1.2 Предмет истраживања.....	4
1.3 Научни допринос тезе	5
1.4 Структура тезе.....	6
Преглед стања у области	7
2.1 Неурокогнитивни увиди у меморију.....	7
2.2 Методолошки приступи развоју рачунарских модела људске меморије	16
2.2.1 Симболички модели меморије	16
2.2.2 Статистички модели меморије.....	25
Аутоматско моделовање домена интеракције	31
3.1 Основна нотација.....	32
3.2 Представљање знања о домену интеракције.....	36
3.3 Моделовање домена интеракције – механизам учења	39
3.3.1 Категоризација семантичких ентитета	39

3.3.2 Асоцијативно учење менталних репрезентација	43
3.4 Кодовање и складиштење информација	48
3.4.1 Усмерене семантичке мреже	48
3.4.2 Аутоматско генерисање семантичких мрежа	50
3.4.3 Аутоматско генерисање фокусног стабла	54
Валидација модела	61
4.1 Валидација модела на корпусу	61
4.2. Валидација у оквиру интегрисаног конверзационог агента	63
Моделовање односа између дуготрајне и радне меморије	70
5.1 Класификација стимуланса	72
5.2 Асоцијативно активирање меморије	75
5.3 Приоритизација активiranог садржаја	79
5.3.1 Мера семантичке сличности менталних репрезентација	79
5.3.2 Активациони потенцијал семантичке мреже	83
5.3.2.1 Релевантност мреже током историје интеракције	83
5.3.2.2 Релевантност мреже за тренутни контекст интеракције	85
5.3.2.3 Адаптивно одређивање вредности прага активације	87
5.3.3 Процена промене контекста интеракције	88
5.4 Концептуализација радне меморије	89
Закључак	90
Библиографија	92

Списак слика


Сл. 2.1	Класификација модула дуготрајне меморије и одговарајући неурални корелати. Илустрација је преузета из [12].	9
Сл. 2.2	Хијерархијска организација појмова и одлика у моделу хијерархијских мрежа. Илустровани пример је модификација примера преузетог из [14].	17
Сл. 2.3.	Топографија појмова и одлика у оквиру модела ширеће активације. Илустровани пример је модификација примера преузетог из [14].	18
Сл. 3.1.	Блок-дијаграм који описује ток извршавања алгорита за аутоматско моделовање домена интеракције.	32
Сл. 3.2	Верзија проблема ханојских кула са три диска (лево) и пример слагалице Танграм (десно). Сlike су преузете из [99].	37
Сл. 3.3	Алгоритам семантичке категоризације. Релација пропозиционе повезаности (\sim) дефинисана је једначином (9). Процедура интеграције менталних репрезентација дефинисана је једначином (10).	42
Сл. 3.4	Алгоритам асоцијативног учења.	45
Сл. 3.5	Семантичка мрежа која одговара домену интеракције везаном за редуковану верзију слагалице Танграм са означеним основним елементима мреже.	49
Сл. 3.6	Аутоматско генерисање семантичких мрежа. Процедура укључује одређивање одговарајућих хијерархијских нивоа у мрежи за семантичке конституенте из менталне репрезентације m_{LI} и успостављање веза између ентитета на суседним нивоима.	52

Сл. 3.7	Скуп семантичких мрежа генерисаних из скупа научених менталних репрезентација M_L [в. јед. (20)]. У оквиру приказаних семантичких мрежа назначена је графичка нотација дефинисаних семантичких јединица. а) Семантички ентитети су представљени чворовима, б) пропозиције су представљене усмереним путањама од почетног до одредишног чвора, в) семантички конституенти су представљени као групе чворова на истом хијерархијском нивоу мреже.	53
Сл. 3.8	Пример фокусног стабла које одговара семантичкој мрежи SM_2 и описује домен везан за редуковану слагалицу Танграм.	55
Сл. 3.9	Фокусно стабло које одговара семантичкој мрежи SM_1 (в. сл. 3.7). Чвор „квадрат“ је произвољно изабран као тренутни фокус пажње.	56
Сл. 3.10	Превођење семантичке мреже SM_2 у фокусно стабло.	57
Сл. 3.11	Фокусно стабло генерисано на основу скупа семантичких мрежа датих на слици 3.7.	60
Сл. 4.1	Индустријски робот АBB IRB140 са две мрежне камере за стереовизију AXIS 211 IP. На столу се налази скуп дрвених објеката који представљају просторни контекст који робот дели са корисником приликом интеракције. Слика преузета из [102].	64
Сл. 4.2	Хуманоидни робот МАРКО. Слика преузета из [104].	64
Сл. 4.3	Средња вредност и стандардна девијација одзива.	67
Сл. 4.4	Средња вредност и стандардна девијација прецизности.	68
Сл. 4.5	Средња вредност и стандардна девијација Ф-мере.	68
Сл. 5.1	Блок-дијаграм који илуструје функционални аспект модела меморије. (1) Механизам асоцијативног активирања меморије, који издваја подскуп садржаја дуготрајне меморије, (2) Механи-	

	зам приоритизације активираниог садржаја, који издваја најрелевантнији подскуп активираниог садржаја (3) Аутоматско генерисање фокусног стабла и контекстно зависно интерпретирање стимуланса.	71
Сл. 5.2	Графичка репрезентација различитих типова стимуланса. Чворови мреже, који су уоквирени задебљаном линијом, представљају експлицитни садржај стимуланса, док су задебљаним стрелицама означене менталне репрезентације на које је могуће прсликати стимуланс. Пуне стрелице означавају једнозначно прсликавање, а испрекидане стрелице представљају потенцијалне путање за прсликавање некомплетних или вишесмислених стимуланса. Пример комплетног стимуланса је приказан у левом делу слике, некомплетног стимуланса у средини, а вишесмисленог стимуланса у десном делу слике.	74
Сл. 5.3	Илустрација ажурирања бафера приликом активирања семантичке мреже. Идентификатор на последњем месту бафера односи се на најскорије активiranу мрежу.....	85

Списак табела

- Табела 3.1 Комплетни скуп пропозиција који представља селектоване домене интеракције повезане са Танграм слагалицом (p_1-p_{12}) и проблемом ханојских кула ($p_{13}-p_{21}$). У претпоследњој колони је означен подскуп пропозиција, који је коришћен као улаз за когнитивног агента.38
- Табела 4.1 Резултати експеримента: средње вредности и стандардне девијације прецизности, одзива и Φ -мере. Вредности су израчунате за 9 различитих величина улазних скупова на узорку од по 100 вредности.67
- Табела 5.1 Структура инвертованог индекса генерисана за скуп семантичких мрежа приказан на слици 3.7. Семантички ентитети у првој колони су уређени према растућим вредностима. Сваки ентитет има придружену уређену листу једнозначних идентификатора семантичких мрежа у којима се јавља. Инвертоване листе су сортиране у растућем низу.76
- Табела 5.2 Векторска репрезентација семантичких мрежа представљених на слици 3.7. Компоненте вектора одговарају релативним учесталостима појављивања ентитета у оквиру семантичких мрежа.81
- Табела 5.3 Мера семантичке сличности израчуната за сваки пар семантичких мрежа представљених на слици 3.7.82



Увод

Science, my boy, is made up of mistakes, but they are mistakes which it is useful to make, because they lead little by little to the truth.

- Jules Verne

Shall I refuse dinner because I do not fully understand the process of digestion?

- Oliver Heaviside

The saddest aspect of life right now is that science gathers knowledge faster than society gathers wisdom.

- Isaac Asimov

The future ain't what it used to be.

- Yogi Berra

Ова дисертација представља резултат ауторовог научноистраживачког рада у области хуманоидне роботике и когнитивних техничких система. У њој је презентован резултат мултидисциплинарног приступа истраживачком проблему развоја когнитивно инспирисаног рачунарског модела људске меморије, погодног за примене у техничким когнитивним агентима. Појам когнитивног агента се овде користи у широком значењу, и односи се на техничке системе чија функционалност симулира биолошки иманентне когнитивне способности људи, укључујући различите типове конверзационих агената, виртуалних анимираних ликова (аватара) и сервисних и хуманоидних робота, који на функционалном нивоу испољавају одређене аспекте интелигентног

понашања. На основу анализе доступне литературе и детаљног прегледа стања у области, према најбољем сазнању аутора, резултати ове дисертације у потпуности представљају оригинални научни допринос.

У оквиру уводног поглавља, након кратког осврта на потребу за спровођењем овог истраживања, биће представљен предмет истраживања и научни допринос тезе. У циљу јасног и концизног излагања, а због непостојања ширег концензуса у истраживачкој заједници у вези са дефиницијама одређених фундаменталних концепата из ове области, у прва два поглавља биће дефинисани основни појмови релевантни за даље излагање.

1.1 Потреба за истраживањем

Развој рачунарских модела људске меморије, применљивих у хуманоидној роботичкој и код других типова техничких когнитивних агената, представља активну област истраживања. Потреба за оваквим истраживањем произлази из актуелних смерова истраживања у роботичкој и вештачкој интелигенцији, који су одређени циљем да се у будућности значајно повећа ниво животне и радне коегзистенције људи и робота. За успешну интеграцију робота у неструктурирано и динамично људско окружење, неопходно је развити роботе који су по својој функционалности упоредиви са људима. Задаци који се постављају пред роботе углавном подразумевају манипулацију објектима из непосредног окружења и кретање у неструктурираној околини која укључује препреке, што захтева, између осталог, извесне способности планирања, учења и одлучивања. За разлику од примена у индустрији где су роботи програмирани да извршавају унапред дефинисане задатке у контролисаном окружењима, роботи који су предвиђени да функционишу у непосредном људском окружењу мораће да испољавају адаптивно понашање у односу на задатак који извршавају, тренутни ситуациони контекст и претходну интеракцију са окружењем.

Још један битан услов за прихватање робота од стране крајњих корисника јесте његово друштвено прихватљиво понашање, тј. да је понашање усклађено са прихваћеним нормама друштвене комуникације. Пошто је језик, и у говорној и писаној форми, основни и најзначајнији

вид изражавања и комуницирања између људи, један од фундаменталних захтева за друштвено прихватљиво понашање робота је да могу да учествују у говорној интеракцији са људима на природном језику. Поред очигледних функционалности аутоматског препознавања и синтезе говора, природна интеракција између људи и робота суштински је заснована на функционалности адаптивног управљања дијалогом, која укључује моделовање знања и контекста, контекстно зависно интерпретирање дијалогских чинова и одлучивање, учење из претходне интеракције, итд. Ове функционалности су у великој мери засниване на меморијском систему. Иако је, до сад, предложен велики број рачунарских модела меморије, они не пружају у довољној мери основу за развој адаптивног управљања интеракцијом. Због доминантног тренда примене статистичких приступа у развоју система вештачке интелигенције, методолошки недостатак већине постојећих модела огледа се у недовољној или никаквој могућности за моделовањем ширег контекста интеракције, што је од кључног значаја за природну и дугорочну интеракцију човека и робота.

У овој тези је предложен рачунарски модел дуготрајне меморије који је проистекао из вишегодишњег континуитета истраживања у области интеракције између људи и машина и представља проширење претходно развијеног когнитивно инспирисаног модела људске радне меморије, названог модел фокусног стабла [1], [2], [3]. Модел фокусног стабла укључује низ правила за контекстно зависно интерпретирање дијалогских чинова и примену адаптивних дијалогских стратегија, што је демонстрирано кроз низ дијалогских система заснованих на овом моделу. С друге стране, модел фокусног стабла претпоставља да је хијерархијска структура, која представља конкретни домен интеракције, унапред дата. У ранијим применама овог модела, релације између семантичких ентитета у фокусном стаблу, као и топологија самог стабла, ручно су кодирани [2], или научени на основу унапред припремљеног корпуса [4]. Модел дуготрајне меморије, предложен у овој тези, омогућава аутоматизовано генерисање симболичке структуре која моделује домен интеракције. Ова структура, између осталог, омогућава и аутоматско генерисање фокусног стабла, што предложени модел меморије чини погодним за интегрисање у нову модуларну архитектуру за конверзационе роботске агенте [5].

1.2 Предмет истраживања

Предмет истраживања ове докторске дисертације је развој когнитивно инспирисаног рачунарског модела дуготрајне људске меморије. Одредница „когнитивно инспирисани“ означава да предложени модел не претендује да објасни когнитивне механизме људског меморијског система, већ само користи одређене увиде из когнитивних неуронаука као инспирацију за развој рачунарског модела за примену у техничким когнитивним системима. Другим речима, рачунарски модел опонаша, на функционалном, не и на анатомском, нивоу изабране људске когнитивне механизме.

Три основна механизма људског меморијског система, која су интегрисана у предложени модел, су: складиштење (кодовање) информација, контекстно зависно селектовање информација и консолидовање информација. Складиштење информација подразумева дефинисање адекватне форме за репрезентацију знања у меморијском систему. Селектовање информација подразумева контекстно зависно изабирање информација из меморије. Контекст је од кључног значаја за правилно интерпретирање стимуланса, па је посебна пажња посвећена истраживачком проблему моделовања контекста интеракције између човека и машине. Консолидовање знања се односи на промену и реструктурирање садржаја меморије и представља основни аспект учења из искуства.

Једно од полазних становишта овог истраживања, утемељено на скорашњим резултатима когнитивних неуронаука, јесте да се радна меморија конципира као функционално стање (тј. активирани део) дуготрајне меморије, а не као засебни анатомски ентитет. Због тога је механизам пажње, који врши селекцију и приоритизацију релевантних информација у складу са тренутним контекстом интеракције, од кључног значаја за предложени модел и посебно је разрађен у оквиру дисертације.

1.3 Научни допринос тезе

Са методолошке тачке гледишта, модел меморије развијен у оквиру ове дисертације припада групи симболичких модела. Овај приступ је изабран, јер омогућава адекватније моделовање контекста интеракције од тренутно доминантно заступљених статистичких приступа.

Главни допринос тезе се може сумирати у следећим важним особинама и функционалностима предложеног меморијског модела. Фундаменталне особине овог модела су:

- *Амодалност* – Семантички ентитети су представљени на униформни начин, независно од сензорског модалитета коришћеног за аквизицију екстерних стимуланса.
- *Рачунарска и аналитичка следљивост* – Извршавање предложених алгоритама је следљиво кроз све кораке.
- *Експланаторна моћ* – Резултати извршавања предложених алгоритама могу се објаснити у оквиру претпостављених концепата на којима се заснива модел.
- *Скалабилност* – Модел прихвата улазни скуп података произвољног обима (за разлику од статистичких модела, који обично захтевају велики корпус за обуку).
- *Независност од домена интеракције* – Модел се може применити на различите домене интеракције између људи и робота.
- *Независност од граматике* – Током интеракције са машином, кориснику нису наметнута никаква синтаксна ограничења у вези са структурирањем дијалошких чинова.
- *Независност од хардверске структуре робота* – Модел меморије се може применити на роботе различитих структура и намена, као и на виртуелне когнитивне агенте.

Најважније функционалности предложеног рачунарског модела меморије укључују:

- функционалност селектовања и приоритизације релевантних информација у складу са тренутним контекстом интеракције,

- функционалност контекстно зависног интерпретирања екстерних стимуланса,
- функционалност категоризације појмова и учење релација између њих.

Валидација предложеног модела је извршена на наменски развијеном симулационом систему. Све улазне вредности у симулациони систем су сматране идеално познатим, јер процес њиховог мерења и аквизиције са сензорских система не представља тему дисертације. Конкретни домени интеракције на којима је илустрована функционалност система изабрани су у складу са спецификацијом постојећег хуманоидног роботског система развијеног на Катедри за мехатронику, роботiku и аутоматизацију Факултета техничких наука у Новом Саду.

1.4 Структура тезе

Текст дисертације је организован у шест поглавља. У 2. поглављу је пружен детаљни приказ тренутног стања истраживања у релевантним областима. У 3. поглављу је презентован модел дуготрајне меморије, при чему је прво дефинисана основна нотација, а затим је разматран складишни аспект меморије, укључујући репрезентацију знања и механизме складиштења информација. Након тога је описан механизам учења за аутоматско моделовање домена интеракције између човека и робота. У поглављу 4. описана је валидација модела на корпусу снимака интеракције између људи и машине и презентован је прототипски систем интегрисан у два конверзациона робота. У 5. поглављу су разматрани функционални аспекти дуготрајне меморије, укључујући механизме за активацију меморије и приоритизацију активираниог садржаја, који су од значаја за практичне примене које укључују велике количине података. Последње поглавље садржи закључак и опис потенцијалних смерова будућег истраживања.

2

Преглед стања у области

A year spent in artificial intelligence is enough to make one believe in God.

- Alan Perlis

Intelligence is the wife, imagination is the mistress, memory is the servant.

- Victor Hugo

У овом поглављу је дат преглед стања у релевантним областима истраживања и моделовања људске меморије. У првом делу поглавља су презентовани изабрани увиди у људски меморијски систем из когнитивних неуронаука и когнитивне психологије, који су коришћени као инспирација у поступку развоја когнитивно инспирисаног рачунарског модела меморије. Након тога је дат преглед постојећих меморијских модела, организован према критеријуму коришћених методолошких приступа који омогућује сагледавање методолошких и имплементационих доприноса предложене дисертације у ширем контексту.

2.1 Неурокогнитивни увиди у меморију

Иако меморија представља један од фундаменталних концепата у когнитивним наукама, и даље не постоји општеприхваћена дефиниција овог појма. Речник Америчког психолошког удружења (енгл. АРА) дефинише меморију као менталну способност кодовања, складиштења и проналажења информација [6]. Стенфордова енциклопедија филозофије

дефинише меморију као скуп различитих когнитивних способности помоћу којих се памте информације и реконструишу претходна искуства, зарад тренутних циљева [7]. Појам меморије се уско везује за појам когниције, који Кембрицов речник психологије [8] дефинише као општи термин за све форме менталних процеса, који укључују свесне процесе попут перцепције, мишљења и меморије, и несвесне процесе као што су превођење стимуланса у перцепте, управљање физиолошким процесима на неуралном нивоу, итд. Неоспорно је да је меморија механизам људског когнитивног система који је од суштинског значаја за комплексне когнитивне задатке, као што су одлучивање, учење, итд. Поред функције складиштења информација, меморија је у значајној мери укључена у когнитивну обраду и манипулацију ускладиштених информација и тесно повезана са механизмом пажње.

Различити когнитивни модели, развијани од средине XX века до данас, подразумевају више функционално разграничених типова меморије, при чему основна подела препознаје сензорску (чулну), оперативну (радну) и дуготрајну меморију. Основна разлика између ових типова меморије односи се на период задржавања информација у меморији. У сензорској меморији се информације најкраће задржавају, док у дуготрајној меморији остају практично трајно. Потребно је напоменути да је понекад тешко одредити јасну границу између различитих типова меморије и њихових механизма функционисања, као и то да постоје различите дефиниције дуготрајне меморије и извесна неслагања међу истраживачима у вези са њеном структуром.

Ендел Талвинг¹ је предложио класификацију дуготрајне меморије засновану на разлици између знања које се односи на лично искуство, са једне стране, и знања о појмовима и процедурама, са друге стране. Према овој класификацији, знање које се односи на наше лично искуство чини део садржаја епизодне меморије, док знање о појмовима и процедурама чини садржај семантичке меморије. Епизодну меморију чини свесно знање о временски и просторно лоцираним, лично проживљеним догађајима и епизодама, док семантичку меморију чини знање о свету, знање о речима и појмовима, њиховим карактеристикама и односима [9].

¹ Endel Tulving (1927-) – експериментални психолог и неурокогнитивни научник.

Алтернативна класификација дуготрајне меморије претпоставља декларативну (експлицитну) и недеklarативну (имплицитну) меморију [10] и [11]. Према овој класификацији, декларативну меморију сачињавају садржаји којима се може свесно приступити, тј. садржаји који се могу декларисати. У ову групу спадају, већ поменуто, епизодна меморија и семантичка меморија. Насупрот томе, недеklarативна меморија обухвата садржаје дуготрајне меморије којима се не може свесно приступити, а који се одражавају као несвесна промена у понашању. Пример недеklarативне меморије је процедурално знање које се односи на вештине које се аутоматски обављају, попут ходања, скијања, итд. Класификација меморије дата у [12], са одговарајућим неуралним корелатима за појединачне меморијске модуле, илустрована је на слици 2.1.



Сл. 2.1 Класификација модула дуготрајне меморије и одговарајући неурални корелати. Илустрација је преузета из [12].

Садржај епизодне меморије примарно потиче из сензорских информација и кодован је у оквиру различитих чулних модалитета, у виду визуелних, аудитивних, тактилних, густативних и олфакторних представа. Такве чулне представе, визуелне перцепте које је око регистровало, у одређеној мери репрезентују стварне особине објекта или процеса који представљају. Из тог разлога се овакве представе називају аналогним. Насупрот томе, семантичка меморија је симболичког карактера и доминантно је представљена у вербалном коду. Истраживања у области неурокогнитивних наука још нису објаснила како се одређене информације из епизодне меморије осмишљавају и структурирају у знање које постаје инвентар семантичке меморије. Осим тога, још увек се води полемика да ли су семантичка и епизодна меморија два анатомски засебна меморијска модула са одговарајућим неуралним корелатима, или само представљају два функционална аспекта једног меморијског система дуготрајне меморије.

Приликом разматрања структуралне организације меморије, поставља се питање форме представљања и складиштења знања у семантичкој меморији. У предложеним теоријама из области когнитивне психологије, као основне јединице за представљање људске семантичке меморије, издвајају се: појмови (концепти), пропозиције и шеме. Појам или концепт, као основна меморијска јединица семантичке меморије, одређен је преко скупа својих дистинктивних одлика и релација са другим појмовима.

Пропозиција је сложенија меморијска јединица којом се могу представити комплекснији односи између појмова. Почетком седамдесетих година XX века развијена је пропозициона теорија према којој менталне репрезентације нису ускладиштене у форми слика или речи, него у апстрактној пропозиционој форми [11]. У том периоду, многи когнитивни психолози прихватили су пропозициону форму као прикладни начин за представљање знања у дуготрајној меморији. Међу утемељивачима пропозиционе теорије били су Зенон Пилишин², Џон Андерсон³ и Гордон Бауер⁴. Пилишин је у оквиру своје теорије

² Zenon Pylishin (1937-) – когнитивни научник и филозоф.

³ John Robert Anderson (1947-) – когнитивни психолог.

⁴ Gordon Bower (1932-) – когнитивни психолог.

јединственог кода [13] изнео схватање да визуелне представе нису реплике објеката, већ само садрже њихове кључне карактеристике. То би значило да се визуелне менталне представе могу превести у скуп пропозиција. Ова теорија постулира да се информације смештене у вербалном и визуелном коду могу изразити у јединственом пропозиционом коду, који је основа организације садржаја у семантичкој меморији [14].

Џон Андерсон и Гордон Бауер су у оквиру теорије НАМ (енгл. *human associative memory*) детаљно разрадили пропозиционе структуре као форму репрезентовања знања која обухвата асоцијативне везе између појмова [15]. Андерсон је касније дефинисао пропозицију као минималну јединицу знања која може да стоји као самостални исказ [16]. Пропозициона форма садржи значење исказа и не зависи од његове граматичке структуре. То значи да би различите варијације исте реченице, настале нпр. променом њене синтаксичке структуре, имале исту пропозициону форму.

Изабрани пропозициони оквир за репрезентацију знања је такође конзистентан са лингвистичком теоријом трансформационо-генеративне граматике, предложене од стране Ноама Чомског⁵ [17]. Језгро ове теорије чини дистинкција између два нивоа репрезентовања реченице, тзв. дубинске структуре и површинске структуре реченице. Дубинска структура представља суштинске семантичке релације у реченици у форми пропозиције, тј. представља значење саме реченице. Свака пропозиција се може преликати на више различитих површинских структура, пратећи правила трансформационе граматике.

Током времена су предлагани различити начини за формално представљање пропозиција, укључујући следеће: логичка форма (пропозициона логика или предикатски рачун), графови и семантичке мреже, фразно-структурална репрезентација коју је предложио Чомски, и шаблони (енгл. *frames*) које је дефинисао Марвин Мински⁶.

⁵ Noam Chomsky (1928-) – лингвиста и филозоф.

⁶ Marvin Minsky (1927-2016) – математичар и истакнути научник у области вештачке интелигенције.

Когнитивна релевантност пропозиција је потврђена и експериментално. Жаклина Сакс⁷ је у експерименту доказала да пропозициона форма није осетљива на синтаксне карактеристике исказа. То значи да временом људи заборављају површинску структуру исказа, тј. синтаксичку форму реченице, али памте значење пропозиције [18]. Волтер Кинч⁸ је у свом експерименту показао да је људски когнитивни систем осетљив на комплексност исказа и да се она може изразити бројем пропозиција које садржи [19]. У експерименту који су спровели Ретклиф⁹ и Мек-Кун¹⁰ испитивана је мера у којој пропозициони оквир може да утиче на обраду речи. Испитиван је утицај контекста на обраду стимуланса, тзв. ефекат примовања. У експерименту је изолован ефекат примовања који потиче од саме пропозиционе структуре исказа, при чему је елиминисан ефекат семантичког или асоцијативног примовања између приказаних парова појмова. Ретклиф и Мек-Кун су дали свој допринос утврђивању когнитивне релевантности пропозиција утврдивши да ширина пропозиционог оквира утиче на време обраде речи [20].

Осим појмова и пропозиција, као јединица организованог знања могу се издвојити и шеме. Шеме су шире целине у којима је знање организовано и оне, поред централног појма, садрже и велики број других појмова, ситуација, догађаја и акција повезаних са централним појмом. Шема може да садржи подшеме које су хијерархијски организоване. Редослед догађаја и операција у оквиру једне шеме назива се сценарио. Као пример за шему може се узети шема супермаркета. Она обухвата низ појмова који се везују за менталну представу супермаркета, утемељену у нашем претходном искуству. Такви појмови су, нпр. робни артикли, касе, колица, корпе, фрижидери, храна, пиће, итд. Поред скупа појмова везаних за контекст супермаркета, шема садржи и низ акција и догађаја, као и њихов редослед који описују сценарио одласка у куповину. Под сценаријом куповине подразумева се да купац улази у супермаркет, узима корпу или колица, бира артикле поређане на полицама или у фрижидерима, плаћа робу на каси, пакује робу и излази. Ако би приликом уласка у радњу продавац затражио од купца да плати, пре него што је

⁷ Jacqueline Sachs – психолог.

⁸ Walter Kintsch (1932) – психолог.

⁹ Roger Ratcliff – когнитивни психолог.

¹⁰ Gail McKoon – когнитивни психолог.

уопште нешто и изабрао за куповину, то би одударало од менталне шеме и сценарија везаног за куповину у супермаркету. У светлу пропозиционе теорије, шема се може посматрати као скуп пропозиција везаних за одређени домен интеракције (знање о одређеној теми).

У разматрањима везаним за садржај и структуру семантичке меморије битно је направити разлику између менталне представе појма, одређеног дистинктивним одликама, и менталне представе речи која означава тај појам. Менталне репрезентације речи, поред значења, укључују и низ лингвистичких карактеристика (нпр. графемске, фонолошке, морфолошке и синтаксне карактеристике) и чине садржај засебног домена дуготрајне меморије, тзв. менталног лексикона [14].

Садржај семантичке меморије организован је на основу асоцијативне и категоријалне повезаности појмова. То значи да су меморијски ентитети у семантичком простору груписани на основу сличности њиховог значења и контекстуалног преклапања. Груписање (кластероване) менталних представа о појмовима има експерименталну потврду у неуронаукама. Техником функциналне магнетне резонанце добијена је тродимензионална реконструкција семантичке мапе кортекса људског мозга, где се види како су менталне репрезентације појмова груписане по површини церебралног кортекса [21].

Категоризација менталних представа о појмовима подразумева њихову хијерархијску организацију, укључујући релације између општијих и специфичнијих категорија. Сваки појам је одређен скупом одлика, при чему се одлике општијег појма пропагирају ка нижим (специфичнијим) нивоима хијерархијске структуре, тј. важе и за све изведене појмове, док обрнуто не важи. То значи да су појмови на вишим хијерархијским нивоима апстрактнији, док су појмови на нижим нивоима одређени већим бројем дистинктивних одлика.

Још један значајан механизам меморијског система је тзв. семантичко примовање које се односи на утицај контекста на когнитивну обраду стимуланса. Код испитаника којима се као стимуланс прикаже графемски запис неке речи на екрану у одређеној мери се активирају менталне представе у мозгу које су семантички повезане са стимулансом. Ако им се као следећи екстерни стимуланс прикаже запис који

представља неку од активираних менталних представа, време когнитивне обраде новог стимуланса је краће него што би било да стимуланс није семантички повезан са тренутно активираним менталним представама. Ефекат примовања може да се јави у виду фацитације (појачања) или инхибиције (слабљења), а такав механизам, који одређује које ће се менталне представе и у коликој мери побудити, назива се механизмом ширеће активације. У вези са тим, заборављање такође представља један од когнитивних феномена везаних за меморију. У контексту дуготрајне меморије, заборављање може да се манифестује као потпуно брисање трага информације у меморији. Друге форме заборављања су парцијална деградација меморијског записа или стапање више меморијских записа у један [14]. Временска крива заборављања варира између различитих меморијских домена. Сматра се да заборављање није узроковано само протоком времена, него и интерференцијом информација до које долази уписивањем нових меморијских садржаја.

Још један аспект људског когнитивног система који је од значаја за ову дисертацију тиче се радне меморије и механизма пажње. Мишљења око статуса радне меморије су подељена. Одређени број истраживача сматра да је радна меморија физички и функционално одвојена меморијска јединица са одговарајућим неуралним корелатима [22], [23], док поједини истраживачи сматрају да је радна меморија само активирани део дуготрајне меморије. За друго становиште, које је заступљено у овој тези, карактеристично је што радну меморију концептуализује као функционално стање дуготрајне меморије [1], [24], [25]. Развијено је много различитих теорија и модела радне меморије који наглашавају различите особине овог меморијског домена. Детаљнији преглед когнитивних и неуралних модела радне меморије је дат у [26], [27].

Људска радна меморија има одређене специфичности које су предмет истраживања когнитивних неуронаука, а од интереса су за развој рачунарских модела меморије. Једна од веома важних особина радне меморије је њен ограничен капацитет који према мишљењу неких аутора износи 7 ± 2 меморијских ентитета [28]. Сматра се да људска радна меморија тежи да групише елементе у смислене целине (енгл. *chunks*), које представљају независне смисаоне јединице информације. Оно што је

заправо ограничено капацитетом меморије је број тих целина, а не број појединачних елемената, што значајно увећава количину података које је могуће привремено сместити у радну меморију. У научној заједници и даље не постоји консензус у вези са идентификовањем механизма који ограничавају капацитет радне меморије. Разматрани су многи механизми као потенцијални узроци ограниченог капацитета радне меморије, а најновија истраживања наводе механизам пажње [29], [30] и способност довлачења садржаја из дуготрајне меморије [31], као два главна узрока. Предложен је и тзв. двокомпонентни модел радне меморије (енгл. *dual-component model*) [32], [33], који оба наведена механизма сматра подједнако важним за ограничење капацитета.

Када механизам пажње побуди одређене делове дуготрајне меморије, део активираних ентитета постаје садржај радне меморије и предмет директне обраде у когнитивним процесима. Задржавање података у радној меморији је релативно краткотрајно, а једна од теорија претпоставља спонтано, нелинарно губљење трага информације у периоду од 3 до 18 секунди [34], [35]. Постоје и алтернативна тумачења која нелинеарно губљење трага информације објашњавају механизмима интерференције и проактивне инхибиције.

Код радне меморије је изражен ефекат позиције у низу [36], [37]. Експериментални резултати показују да се у задацима секвенцијалне репродукције листе задатих стимуланса најуспешније памте елементи на почетку листе. Разлог је тај што се елементи са почетка најдуже понављају (енгл. *rehearsal*). За разлику од тога, када је потребно репродуковати елементе из листе у произвољном редоследу, приликом слободне репродукције, најуспешније се памте елементи са краја листе, који су најактуелнији у радној меморији.

2.2 Методолошки приступи развоју рачунарских модела људске меморије

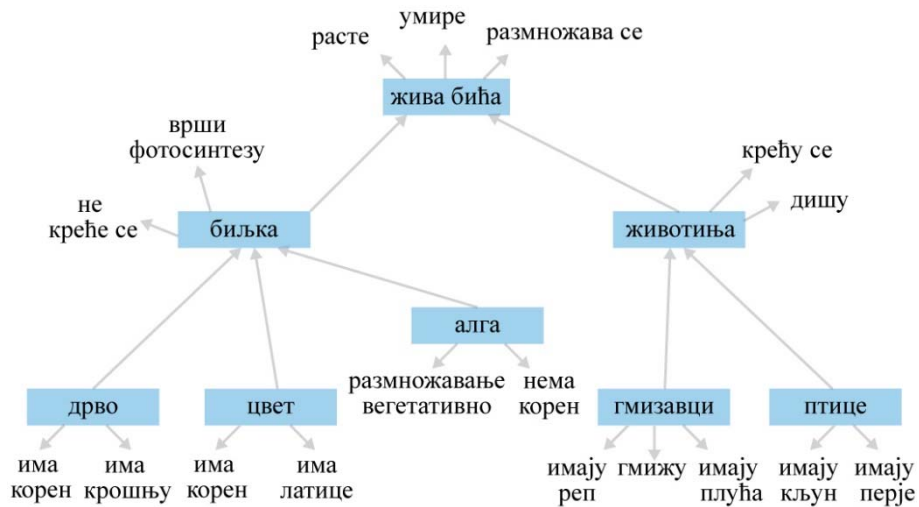
Постојећи модели меморије истичу различите аспекте људског меморијског система. Модели који су проистекли из когнитивне психологије имају за циљ да обухвате што већи број експерименталних резултата везаних за способности и ограничења људске меморије. Овакви модели теже да објасне људске когнитивне механизме на функционалном нивоу. Модели проистекли из неуронаука су фокусирани на одговарајуће неуралне корелате појединих меморијских модула и механизма. Са развојем техника неуроосликавања и проучавањем можданих лезија у клиничкој пракси, предлагана су различита пресликавања когнитивних механизма на регионе мозга. Поједини модели из когнитивне психологије и неуронаука имали су и своје рачунарске имплементације. У наставку поглавља је дат преглед стања у области когнитивно инспирисаног моделовања меморије, организован на основу методолошке поделе на симболичке и статистичке моделе.

2.2.1 Симболички модели меморије

Са методолошког аспекта, прва група модела се односи на симболичке, тј. репрезентационе, моделе меморије. Крајем шездесетих година XX века постаје актуелно питање организације појмова у семантичкој меморији и предложени су први модели. Један од пионирских радова у овој области је модел хијерархијских мрежа Роса Килијана¹¹ [38], [39], који претпостављену хијерархијску организацију појмова пресликава у семантичку мрежу међусобно повезаних појмова. Чворови мреже представљају категорије, а везе између појмова и њихових одлика су представљене усмереним гранама које повезују чворове. Хијерархија у мрежи је заснована на односу између општијих и специфичнијих категорија, тако што су гране усмерене од општијих ка специфичнијим појмовима. Веза између два појма на различитим хијерархијским нивоима има статус копуле (помоћни глагол „је“), а веза између појма и његове одлике има статус предиката. Значајна карактеристика модела је

¹¹ Ross Quillian (1931-) – когнитивни научник.

да се одлике појма на вишем хијерархијском нивоу пропaгирају на све његове потомке. На слици 2.2 је приказан пример једне хијерархијске мреже. Ограничење овог модела је то што је прикладан само за обраду релативно једноставних исказа.



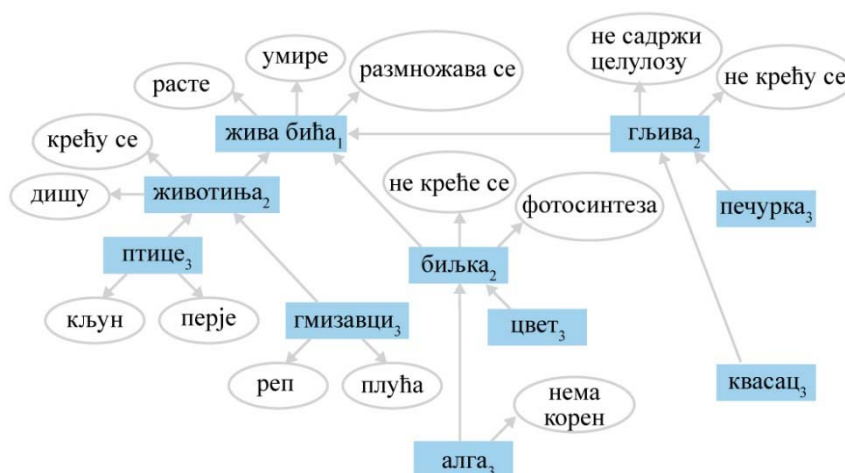
Сл. 2.2 Хијерархијска организација појмова и одлика у моделу хијерархијских мрежа. Илустрирани пример је модификација примера преузетог из [14].

Разлика у временима потребним за верификацију исказа, у зависности од броја хијерархијских нивоа између субјеката и предиката, у то време је објашњавана увођењем претпоставке о ширећој активацији. Механизам ширеће активације је био заснован на идеји о прагу побуђивања. Сматрано је да свака ментална представа има специфични праг побуђивања, при чему је нижем прагу одговарало краће време верификације појма. Побуђивање једног појма у меморији изазива побуђивање других појмова који су у семантичкој, граматичкој или фонолошкој вези са њим, тј. активација појмова зависи од њихове релативне позиције у мрежи. Појмови и одлике, активирани су у мањој мери, ако су удаљенији од иницијалног појма.

Каснијим експериментима је установљено да постоји низ особина људског меморијског система који се нису могли објаснити моделом

хијерархијских мрежа, попут утицаја типичности појма на време верификације [40], зависности времена категоризације појма од учесталости одлике [41] и зависности брзине верификације појмова на истом хијерархијском нивоу од надређених категорија [42]. Ови увиди су подстакли развој неколико нових когнитивних модела семантичке меморије, као што су модел ширеће активације, модел прототипа и модел преклапања одлика.

Модел ширеће активације [43] заменио је хијерархијску организацију појмова специфичном топографијом, изведеном из семантичких веза. Јачина семантичких веза између појмова је одређена њиховом релативном удаљеношћу унутар топографије, при чему је релативна позиција појма одређена степеном преклапања особина датог појма са особинама њему надређене категорије. Један пример топографије у оквиру модела ширеће активације, који приказује део садржаја семантичке меморије и илуструје њену структурну организацију, приказан је на слици 2.3.



Сл. 2.3. Топографија појмова и одлика у оквиру модела ширеће активације. Илустрирани пример је модификација примера, преузетио из [14].

Квантификација јачине семантичких веза је одређена хеуристички, на основу процене испитаника. Смер веза у топографији није произвољан

и одређује однос између појмова. Појмови су представљени четвороугаоним правоугаоницима, при чему индекси, наведени уз појмове, служе да означе појмове који припадају истом нивоу хијерархије. Одлике појмова су представљене овалним чворовима.

Идеја о типичности представника категорије довела је до развоја још једног когнитивног модела семантичке меморије, названог моделом прототипа који је предложила Еленор Рош¹² [44], [45]. У оквиру овог модела, прототип се дефинише као идеални представник категорије који представља просек особина свих представника. Прототип је апстрактна инстанца категорије на чије одређивање утиче учесталост јављања одређених инстанци у категорији. Прототип је доминантно одређен особинама оних инстанци категорије које се најчешће јављају. Категорије у овом моделу нису међусобно искључиве, што значи да један појам може да припада једној категорији или више категорија, са различитим степенима припадности. Инстанце које су блиске прототипу категорије деле већи број одлика са осталим члановима категорије и код тих инстанци је ефекат примовања израженији.

Упоредо са моделом прототипа је развијен још један когнитивни модел семантичке меморије, назван моделом преклапајућих одлика који практикује суштински другачији приступ [46]. У овом моделу су одбачени претпоставка о просторној организацији појмова и идеја о ширећој активацији, а ментална репрезентација појма је представљена скупом семантичких одлика. Одлике појма су хијерархијски организоване по значају и могу се поделити на одређујуће и карактеристичне. Одређујуће одлике се налазе највише у хијерархији и омогућавају основну дистинкцију између појмова. Карактеристичне одлике обухватају својства која су специфична за дати појам. Задаци категоризације и верификације се своде на поређење одлика два појма.

Рани, тзв. класични модели семантичке меморије, као што су хијерархијски мрежни модел [38], [39], модел ширеће активације [43] и модел поређења одлика [46], нису одредили како су менталне репрезентације научене [47]. Две заједничке карактеристике раних симболичких модела су њихова уобичајена зависност од екстерне

¹² Eleanor Rosch (1938-) – когнитивни психолог.

онтологије и потреба за унапред задатим релацијама између семантичких ентитета (нпр. израженим кроз тополошку структуру мреже). Ове релације су најчешће ручно кодоване или хеуристички процењиване (нпр. попуњавањем упитника од стране испитаника).

Сви поменути модели су ограничени на релативно једноставни тип везе између појмова и одлика, најчешће у форми копуле „је“ или предиката. Пропозиције су предложене да би се концептуализовало знање о комплекснијим односима између појмова или о везама између појмова и акција у оквиру неког контекста. Оне могу бити изражене у форми реченице или у форми релације између конституената пропозиције и нису ограничене само на семантичке и асоцијативне везе међу појмовима, него могу да представе и садржаје који укључују субјекат и предикат. Оваква пропозициона структура била је у основи групе когнитивних модела познатих под називом АСТ модели (енгл. *adaptive control of thought*), који су развијани од средине седамдесетих до средине деведесетих година XX века. На основу модела АСТ* из ове групе [48] имплементиран је рачунарски модел који је прерастао у когнитивну архитектуру опште намене са циљем симулирања функционисања целокупног когнитивног система човека.

У скороје време често се рачунарски модели дуготрајне меморије имплементирају у оквиру когнитивних архитектура опште намене, уместо да буду дискретни когнитивни модули. Детаљнији преглед когнитивних архитектура и дискусија везана за истраживачке изазове и проблеме су доступни у [49], [50]. а овде ће бити представљене когнитивне архитектуре релевантне за даље излагање.

Једна од најзначајнијих хибридних рачунарских когнитивних архитектура је АСТ-R (енгл. *adaptive control of thought - rational*), која представља наставак развоја поменуте групе АСТ модела [51], [52]. Меморијски систем у овој архитектури је подељен на процедурални модул и декларативни модул. Основне меморијске јединице у овим модулима су симболичке структуре које се односе на процедурално и декларативно знање (енгл. *productions* и *chunks*). АСТ-R интегрише ове симболичке структуре са подсимболичким скупом процеса који служе за селектовање најадекватнијих процедуралних структура (тј. оних са

највишим степеном очекиване корисности у датом контексту) и најадекватнијих декларативних структура (тј. оних са највишим степеном активације). Ови подсимболички процеси, описани скупом једначина, представљају језгро декларативног модула и моделују утицај тренутног контекста и претходних пруступања меморији на вероватноћу будућег активирања декларативних структура [52]. Ова архитектура је примењивана у моделовању феномена повезаним са многим аспектима људског понашања, укључујући: семантичку категоризацију [53], асоцијативно учење [54], и интеракцију [55]. У поређењу са моделом АСТ-R, у овој тези се предлаже примарно симболички и рачунски следљив приступ моделовању меморије, задржавајући могућност уопштења на различите домене интеракције између људи и робота.

Меморијски систем когнитивне архитектуре ICARUS [56], [57] поседује два различита типа меморијске репрезентације којима се прави разлика између концепата и вештина. Оба меморијска домена поседују краткотрајни и дуготрајни модул. Концептуална меморија садржи знање о класама објеката и релацијама између објеката. Дуготрајна меморија ове архитектуре има изражену хијерархијску организацију, при чему се примитивни концепти налазе на дну хијерархије, а комплексни концепти ближе врху. Закључивање у концептуалној меморији користи приступ од дна ка врху (енгл. *bottom-up*), а избор акција у меморији вештина обавља се приступом од врха ка дну (енгл. *top-down*) [49].

Концептуални дизајн контекстно зависне архитектуре дуготрајне меморије презентован је у [58]. Архитектура подразумева модел меморије који разликује радну и дуготрајну меморију. Рачунарски модел радне меморије је заснован на когнитивном моделу Алана Бедлија¹ [59]. Модел дуготрајне меморије обухвата три типа меморије: епизодну, семантичку и процедуралну. Централна управљачка једница у предложеној архитектури управља кодовањем и декодовањем информација коришћењем контекста. Главни допринос овог модела је увођење контекстне информације као средства за повезивање различитих

¹ Alan Baddeley (1934-) – психолог, познат по својим радовима везаним за радну меморију. меморију.

типова меморије и анализирање начина на који увођење контекста утиче на процес проналажења информација у меморији [58].

Меморијски модел $CRAM_m$ (енгл. *cognitive robot abstract machine*) омогућава роботима да памте претходно стечено искуство у форми која им омогућава апостериорно доношење закључака, анализу и реконструкцију формиране представе о околини, у различитим временским тренуцима [60]. Овај модел користи логику првог реда за извођење предиката који описују концептуализације намера робота, његовог физичког стања, перцепције и предвиђених последица акција. Главни доприноси овог модела су: (i) развој меморијских репрезентација које комбинују симболичко планирање догађаја са подсимболичким сензорским подацима, (ii) развој метода за временско, просторно, дијагностичко и каузално резонување које се примењују на симболичке и подсимболичке меморијске структуре и (iii) развој ефикасних механизма за вођење хронолошке евиденције података о функционисању робота без деградирања перформанси робота. Једна компонента модела $CRAM_m$ је тзв. база знања *KnowRob*, која садржи структуре за репрезентацију задатака, као и њиховог просторног и временског контекста који укључује догађаје, објекте, просторне мапе околине и компоненте робота [61]. Ова база је имплементирана у програмском језику Пролог (енгл. *Prolog*), док је за репрезентацију знања коришћен језик OWL (енгл. *web ontology language*).

У оквиру когнитивне архитектуре ISAC (енгл. *intelligent soft arm control*) развијен је когнитивно инспирисани модел адаптивне радне меморије за примену на хуманоидном роботу [62]. Софтвер је заснован на моделима из области неуронаука који истичу улогу интеракције између префронталног кортекса и мезолимбичког допаминског система у функционисању радне меморије [63], [64]. У рачунарском моделу, учење је реализовано алгоритмом временске разлике (енгл. *temporal difference learning*), који је имплементиран у форми неуронске мреже. Модел је верификован у неколико симулација које су укључивале обучавање робота за задатке одложених сакада¹ [62] и перцептивног учења повезаног са навигацијом и препознавањем оријентира у простору

¹ Сакаде су брзи покрети очима.

[65]. Поменути алгоритам за учење спада у класу алгоритама обучавања са подстицајем (енгл. *reinforcement learning*), инспирисаних улогом допаминских ћелија у кодовању промена очекивања будуће награде [66]. Алгоритам временске разлике се користи за селекцију одговарајуће акције и доношење одлуке о задржавању информација у меморији.

Значајно је поменути и когнитивну архитектуру OpenCogPrime, која представља комбинацију различитих парадигми вештачке интелигенције и поседује комплексни меморијски систем са различитим модулима [67], [68]. Поред декларативне, сензорске, процедуралне и епизодне меморије, ова архитектура садржи и два специфична меморијска модула: меморију пажње¹ и меморију намере².

Дејвид Денкс³ [69] је предложио нови правац рачунарског моделирања когнитивних функција, заснован на графовима, који делимично објашњава два главна аспекта људског когнитивног система: интегрисану природу различитих когнитивних механизма и способност фокусирања пажње на факторе релевантне за тренутни ситуациони контекст. Денкс заступа став да се структура многих когнитивних репрезентација може разумети и интерпретирати у контексту графова. Приступ изложен у овој дисертацији је на линији ове идеје – предложени рачунарски модел такође користи графове за репрезентацију различитих когнитивних структура и механизма.

Ако разматрање фокусирамо на развој модела радне меморије човека, један од раних концепата, познат под називом модални модел (енгл. *modal model*) [22] Аткинсона⁴ и Шифрина⁵, постулирао је постојање три одвојене компоненте меморијског система: сензорског регистра, краткотрајног регистра и дуготрајног регистра. Краткотрајни регистар је био дефинисан као привремено складиште у ком се информације из сензорског система задржавају приближно од 15 до 30 секунди. Након тог периода, меморијски траг информације почиње

¹ Attentional memory

² Intentional memory

³ David Danks - филозоф

⁴ Richard Chatham Atkinson (1929-) – психолог.

⁵ Richard Shiffrin (1942-) – психолог и когнитивни научник.

постепено да деградира, осим уколико се поступком понављања не преведе у садржај дуготрајног регистра.

Бедли је предложио радикално нову интерпретацију радне меморије и модел који је произашао из ње [23]. Бедлијев модел поставља раније виђење краткотрајне меморије у шири оквир система за обраду информација, тј. радну меморију. Овај модел разликује три компоненте краткотрајне меморије: централну извршну јединицу, фонолошку петљу и визуелно-просторну матрицу. Централна извршна јединица је надзорна компонента која управља током информација. Преостале две компоненте су краткотрајни бафери, специфични по садржају, за вербалне и визуелно-просторне информације. Модел је касније проширен и четвртм компонентом, епизодним бафером. Ова компонента представља складиште које омогућава да се информације, кодоване у различитим модалитетима, повежу и интегришу како би се креирала епизода као смислена целина [59].

Ова два традиционална модела меморије конципирала су механизам пажње као систем независан од радне меморије, за разлику од неких каснијих модела који га интегришу. Нелсон Кован¹ је предложио концепт радне меморије, назван моделом интегрисаног процеса (енгл. *embedded process model*), који дефинише радну меморију не као засебну компоненту него као функционално стање дуготрајне меморије [24], [70]. Кованов модел садржи четири компоненте: централну извршну јединицу, дуготрајну меморију, активирани део меморије, и фокус пажње. Клаус Оберауер² је предложио концентрични модел меморије (енгл. *concentric model*) [25], у ком је извлачење информација из радне меморије еквивалентно постављању информације у фокус пажње. Овај модел дефинише следећа меморијска стања: активирани део дуготрајне меморије, област директног приступа меморији и фокус пажње. У складу са овим приступом, у [1] предложен је модел управљања пажњом у интеракцији између људи и машина, који ће у овој дисертацији бити интегрисан са предложеним симболичким моделом дуготрајне меморије.

¹ Nelson Cowan (1951-) – психолог.

² Klaus Oberauer – когнитивни психолог.

2.2.2 Статистички модели меморије

Поткласа статистичких модела меморије од посебног значаја за ову дисертацију односи се на дистрибуционе моделе, повезане са тзв. статистичком семантиком. Дистрибуциони модели семантичке меморије спецификују когнитивне механизме формирања и учења семантичких репрезентација, најчешће заснованих на анализи језичких корпуса. Настанак дистрибуционих модела има своје корене у филозофији Лудвига Витгенштајна¹ [71] и радова енглеског лингвисте Џона Руперта Фирта² [72]. Фирт је заступао идеју да је право значење речи увек контекстно зависно и познат је по својој изреци да се реч може познати по друштву у ком се налази. Ова идеја је касније разрађена у дистрибуциону хипотезу контекстуалног преклапања [73]. Ова хипотеза тврди да речи, које се јављају у сличним лингвистичким контекстима, теже да имају слично значење. У овом прегледу стања навешћемо неколико дистрибуционих модела релевантних за ову тезу.

Модел LSA (енгл. *latent semantic analysis*) [74] је високодимензиони линеарни асоцијативни модел општег механизма учења који анализира обрасце (енгл. *pattern*) симултаног појављивања одређених група речи унутар докумената у текстуалним корпусима, како би генерисао семантичке репрезентације. Претпоставка овог модела је да се речи сличног значења јављају заједно у тексту. Основна структура података на којој се заснива идеја алгорита је матрица инциденције термина и докумената, у којој врсте представљају термине (тј. основне појавне облике речи присутних у колекцији докумената), а колоне - документе (тј. неструктуриране текстове). Елементи ове матрице су тежине, дефинисане као функције учесталости појављивања посматраног термина у посматраном документу, и представљају мере релевантности термина. Пошто је матрица инциденције великих димензија и по правилу ретка, у пракси се користи редукована матрица ниског ранга добијена применом сингуларне декомпозиције (енгл. *singular value decomposition*). Након декомпозиције се узимају прве N сингуларне вредности, тј. издваја се N латентних семантичких обележја која немају обавезно аналогију у

¹ Ludwig Wittgenstein (1889-1951) – филозоф, присталица логичког позитивизма.

² John Rupert Firth (1890-1960) – лингвиста.

неким разумљивим физичким концептима, али анализа редуковане матрице може да покаже семантичку и контекстуалну сличност између два термина, чак и ако се термини у оквиру језичког корпуса за обуку нису јављали у истим документима. Семантичка репрезентација речи у редукованој матрици може се посматрати као образац који садржи N латентних семантичких обележја, тј. сваки термин је представљен тачком у векторском семантичком простору. Латентна семантичка анализа је примењивана за моделовање низа феномена везаних за семантичку меморију, укључујући категоризацију, обраду синонима и вишезначних термина, итд.

У моделу „Topic“ су разрађене поставке модела LSA у бајесовском¹ теоријском оквиру, где је предложен когнитивни механизам за одређивање значења речи [75]. Полазећи од претпоставки да одређени скупови термина теже да се јављају заједно и да се шаблон њиховог груписаног појављивања понавља кроз различите контексте, модел открива латентне теме. Значење речи се концептуализује као расподела вероватноћа над скупом могућих тема, а значење теме као расподела вероватноћа над скупом речи. Важна особина овог модела је таква да је он заснован на генеративном приступу, што омогућава да се документ посматра као мешавина више семантичких променљивих [76].

Другу групу дистрибуционих модела чине модели који примењују блоковску обраду текста, тзв. прозорирање (енгл. *moving window models*). Један од представника ове групе је модел HAL (енгл. *hyperspace analogue to language*) [77]. HAL се може сврстати и у групу модела пасивног копојављивања (енгл. *passive co-occurrence models*) који примењују једноставан механизам акумулације, заснован на Хебовом правилу како би изградили семантичке репрезентације речи. Разлика ових модела видљива је у односу са моделима латентне анализе тако што не анализирају читав документ као дискретну јединицу у оквиру матрице инциденције термина и докумената него континуирано анализирају текстуални корпус применом блоковске обраде [47]. У моделу HAL се примењује прозор који укључује 10 речи наведених испред циљне речи и 10 речи иза ње. У свакој итерацији прозор се помера за по једну реч

¹ Thomas Bayes (1702-1761) – математичар и свештеник

унапред. Алгоритам формира глобалну матрицу истовременог појављивања парова термина из доступног речника, која се прерачунава приликом сваког померања прозора. Вредности елемената ове матрице укључују тежинске коефицијенте који су инверзно пропорционални дистанци од циљне речи. Семантичка репрезентација речи настаје спајањем вектора врсте и вектора колоне из матрице истовременог појављивања термина [78]. Примењујући релативно једноставни алгоритам учења заснован на пребројавању речи које се јављају, модел HAL је примењиван за моделовање когнитивних феномена везаних за семантичке задатке, као што су семантичко примовање и решавање проблема. Неколико новијих дистрибуционих модела је развијено проширивањем модела HAL. Модел „Hidex“, нпр. узима у обзир чињеницу да високофреквентне речи, као што су неки предлози или речца „је“, носе мању информациону добит, због чега се уводи посебни тежински коефицијент који је инверзно пропорционалан учесталости посматране речи [77].

Још једну значајну подгрупу дистрибуционих модела чине случајни векторски модели (енгл. *random vector models*), чији је репрезентативни представник модел BEAGLE (енгл. *bound encoding of the aggregate language environment model*) [79]. Овај модел почиње од произвољне репрезентације речи и постепено развија семантичку структуру кроз анализу епизода из текстуалног корпуса, узимајући у обзир контекстуално преклапање речи (као модел LSA) и позицију у односу на друге речи (као модел HAL). Свакој речи, приликом њеног првог појављивања, придружује се тзв. вектор окружења који се третира као непроменљива одлика саме речи. Осим тога, свакој речи се придружује и меморијски вектор, чија вредност се ажурира сваки пут када се посматрана реч појави у тексту. Меморијски вектор представља суму вектора окружења оних речи које деле заједнички контекст са посматраном речју. Ако се две речи често јављају у истом контексту, са истом групом других речи, њихови меморијски вектори ће имати сличне вредности. Другим речима, две речи имају сличније значење уколико је мања дистанца између њих у векторском простору.

Пошто значење речи често зависи од контекста употребе, у неким статистичким приступима се посматра значење речи у променљивим

контекстима. Временски контекст је први пут употребљен у моделу TCM (енгл. *temporal context model*) [80], да би се узели у обзир ефекат повезан са временом последњег приступа [81] и ефекат суседства у задацима присећања, везаним за епизодну меморију. У каснијем раду исти концепт временског контекста употребљен је за генерисање семантичких репрезентација [82].

Дистрибуциони модели се углавном односе на семантички аспект дуготрајне меморије и користе статистичку редундансу језика како би извели семантичке репрезентације. Ови модели су посебно успешни у задацима селектовања информација из меморије, јер су робусни у случајевима делимично познатих стимуланса [83]. Ипак, ови модели се обично не односе на друге важне аспекте меморијског система повезане са хијерархијском и асоцијативном природом меморије, механизмом пажње, итд.

Другу значајну класу статистичких модела чине неуронски мрежни модели. Неуронска парадигма је утемељена у раду Румелхарта¹ и МекКлиланда² [84]. Она се односи на класу модела који користе неуролошки инспирисани приступ моделовању когнитивних и бихејвиористичких процеса као појавних феномена. Овај приступ се назива још и конекционистички. Архитектура ових модела је сачињена од једноставних неурона међусобно повезаних у мрежу, при чему се параметри неурона одређују и адаптирају кроз неки од алгоритама учења. Когнитивни процес који се моделује неуронском мрежом није кодован неким скупом правила или алгоритмом, него се испољава као последица понашања мреже као целине, тако да концепти представљају функционално стање система, тј. образац активације мреже. Мрежни модели су коришћени за симулирање динамике семантичке меморије и интеракције семантичке меморије са другим когнитивним процесима [78]. Детаљни преглед неуронских модела је дат у [85], а овде ћемо навести поједине моделе, чије су примене од интереса за даље излагање.

Румелхарт и МекКлиланд су истраживали проблем рачунарске симулације перцепције и дефинисали су један од првих модела неуралне

¹ David Rumelhart (1942-2011) – психолог.

² James McClelland (1948-) – когнитивни психолог.

обrade [84]. Румелхартов неуронски модел семантичке меморије [86] је имао улазни слој неурона, излазни и два скривена слоја неурона. Улазни слој је садржао скуп неурона који су представљали концепте (нпр. врабац, канаринац, итд.), док је други слој садржао неуроне који су представљали релације (нпр. јесте, може, има, и др.). Применом алгоритма за надгледано машинско учење, мрежа је обучена да успостави везу између концепта и релације са улазне стране, са семантичком одликом на излазу (нпр. летети, певати, итд.). Унутрашња репрезентација одређеног улазног концепта је заправо специфична схема активације мреже. Ако је мрежа обучена да повеже концепт А са одликом X, онда се концепт Б са сличном шемом активације мреже такође може довести у релацију са одликом X, што представља основу за симулирање механизма категоризације, генерализације, индукције, итд. [78]. Румелхартова мрежа спада у групу неуронских мрежа са пропагацијом сигнала унапред (енгл. *feed-forward*).

Знатан број модела семантичке меморије припада групи динамичких неуронских модела. Динамичке неуронске мреже су заправо рекурентне неуронске мреже које интегришу механизам повратне спреге и имају двосмерни ток информација. Посебну групу рекурентних динамичких мрежа, које су у великој мери коришћене за моделовање меморијских механизма, чине атракторске мреже које током времена конвергирају ка стабилним шемама активације. Оваква стабилна стања се називају атрактори и представљају кодоване меморијске ентитете. Један од раних неуронских модела меморије, заснован на рекурентним неуронским мрежама, јесте Хопфилдова мрежа [87]. Џон Хопфилд³ је доказао да његова мрежа конвергира у стабилно равнотежно стање. У овој мрежи алгоритам учења се заснива на Хебовом⁴ правилу по коме синхронизована активација неурона јача везу између њих. У међувремену је развијен велики број модела заснованих на атракторским мрежама, који су моделовали феномене везане за семантичку меморију, попут семантичког примовања [88], [89] и каузалног закључивања утемељеног на семантичком знању [90]. Румелхарт је користио

³ John Hopfield (1933 -) – физичар и неуронаучник.

⁴ Donald Hebb (1904-1985) – психолог.

атракторску неуронску мрежу за успостављање когнитивних шема на основу анализе динамике симултаних појављивања објеката [91].

У [83] је представљен самоорганизујући модел меморије, назван фузионом адаптивном резонантном теоријом мултимеморијског учења (енгл. *fusion adaptive resonance theory for multymemory learning*). Ова неурална архитектура представља и учи различите типове семантичког знања на униформни начин и адресира проблем интеракције семантичке меморије са осталим модулима дуготрајне меморије, укључујући епизодну и процедуралну меморију. У [92] ова теорија је коришћена да се постигне обједињени приступ скупу различитих парадигми учења, укључујући учење подударењем, асоцијативно учење, инструкисано учење и учење појачавањем. Учење се изводи у реалном времену на инкрементални и континуални начин. Још један модел семантичке меморије заснован на неуронским мрежама, који повезује сензори-моторне репрезентације објеката и лексичке репрезентације речи које их описују, предложен је у [93]. Неуронска мрежа за селектовање информација из дуготрајне меморије, заснована на Хопфилдовој неуронској мрежи, презентована је у [94].

Неуронске мреже су генерално подобне за парцијалну и постепену деградацију која се постиже уклањањем изабраних неурона и веза, тако да се користе за симулирања можданих лезија. У једном таквом приступу неуронска мрежа је коришћена за процену неуралних корелата декларативне дуготрајне меморије код пацијената са благим когнитивним оштећењима (енгл. *mild cognitive impairment*) [95]. Неуронска мрежа је такође коришћена за моделовање погоршања меморијског система код Алцхајмерове болести [96].

Ипак, неуронски модели ретко разматрају фундаментална питања попут кодовања и селектовања информација у људској меморији, моделовања ширег контекста интеракције, итд. [26]. Насупрот томе, у овој дисертацији је примењен репрезентациони приступ да би се адресирала ова фундаментална питања.



Аутоматско моделовање домена интеракције

To think is to forget differences, generalize, make abstractions.

- Jorge Luis Borges

Science may be described as the art of systematic over-simplification.

- Karl Popper

Ово поглавље описује рачунарски модел дуготрајне меморије који је на структуралном нивоу хијерархијски а на функционалном нивоу асоцијативан. Информације су ускладиштене у хијерархијску структуру која се састоји од семантичких јединица различитих комплексности, укључујући семантичке ентитете, семантичке конституенте и менталне репрезентације различитих типова. Предложен је алгоритам за аутоматско моделовање домена интеракције [97] који имплементира процесе (а) семантичке категоризације и асоцијативног учења менталних репрезентација и (б) кодовања и смештања информација у дуготрајну меморију. На слици 3.1 је приказан блок-дијаграм који илуструје ток извршавања алгоритма.

У првој фази извршавања овај алгоритам прихвата улазни скуп менталних репрезентација које само делимично описују један или више домена интеракције, класификује менталне репрезентације према доменима којима припадају и изводи менталне репрезентације које недостају. Као резултат извршавања овог алгоритма, успостављају се

семантичке релације између семантичких ентитета које нису биле априори дате. У другој фази алгоритма се почетне и изведене менталне репрезентације структурирају у форми графа, тј. као скуп тзв. семантичких мрежа.



Сл. 3.1. Блок-дијаграм који описује ток извршавања алгорита за аутоматско моделовање домена интеракције.

3.1 Основна нотација

У овом поглављу су дефинисани основни концепти потребни да би се објаснила хијерархијска структура предложеног модела дуготрајне меморије. Модел укључује семантичке јединице различитих комплексности: (i) семантички ентитет, (ii) семантички конституент и (iii) менталну репрезентацију.

(i) **Семантички ентитет** – представља основну и најједноставнију меморијску јединицу, која се не може даље разлагати у посматраном домену интеракције. Семантички ентитети се формално представљају симболима, који на униформни начин представљају информације произвољног сензорског модалитета. У овој дисертацији су, ради лакшег излагања, семантички ентитети најчешће представљени кључним речима или фразама, што не имплицира да су семантички ентитети ексклузивно везани за језички модалитет. Примери семантичких ентитета су:

$$e_1 = \{\text{квадрати}\}, e_2 = \{\text{велики црвени троугао}\}. \quad (1)$$

Успостављање скупа основних семантичких ентитета је контекстно зависно, тј. зависи од посматраног домена интеракције. Горе наведени пример подразумева домен интеракције у коме *квадрати* и *велики црвени троугао* представљају основне семантичке ентитете. Без обзира на то што је други ентитет представљен фразом која садржи три речи, он у посматраном домену интеракције није разложив. Фраза се користи само да би излагање било интуитивније читаоцу, док би за машинску репрезентацију било довољно да горњи ентитети буду представљени са два различита симбола. Ово је важна особина предложеног модела. Он манипулише скупом симбола који немају своје инхерентно значење (иако је читаоцу јасно семантичко значење именице „квадрат“, за систем је то само један симбол коме није придружено значење). Један од основних задатака модела је да на основу парцијалног скупа веза између основних ентитета конструише комплетни скуп веза између њих. У наставку излагања ћемо подразумевати да непразни скуп SE садржи све семантичке ентитете из једног или више домена интеракције, које је потребно моделовати.

(ii) **Семантички конституент** – јесте коначни и непразни скуп семантичких ентитета који припадају истој семантичкој категорији. Припадност истој категорији се може представити као релација еквиваленције над скупом SE тако да су семантичке категорије представљене међусобно дисјунктним скуповима SC_1, SC_2, \dots, SC_m , чија је унија једнака скупу SE :

$$sc = \{e_1, e_2, \dots, e_k \mid k \geq 1 \wedge (\forall 1 \leq i \leq k) e_i \in SE\} \subseteq SC_j. \quad (2)$$

Примери семантичких конституената су:

$$sc_1 = \{\text{квадраи}, \text{иџроуџао}\}, sc_2 = \{\text{улево}, \text{угесно}, \text{наџоре}, \text{наголе}\}. \quad (3)$$

Семантички конституент групише ентитете који припадају истој семантичкој категорији, тј. који имају исту семантичку улогу у оквиру исказа. Пошто су ове везе контекстно зависне, а модел је пројектован да буде независан од домена интеракције, ове везе нису унапред задате. Уместо тога, задатак модела је да успостави ове везе на основу почетног скупа задатих менталних репрезентација, што је објашњено у секцији 3.3.

(iii) Ментална репрезентација – јесте коначни и непразни скуп семантичких конституената који представља неко (валидно) значење у датом домену интеракције:

$$m = \{sc_1, sc_2, \dots, sc_k \mid k \geq 1\}. \quad (4)$$

Предложени приступ уводи два критеријума за класификацију менталних репрезентација. На основу кардиналности садржаних семантичких конституената, менталне репрезентације се деле на базичне и комплексне. На основу комплетности значења, менталне репрезентације се деле на комплетне и некомплетне. Треба напоменути да комплетност значења менталне репрезентације зависи од посматраног домена интеракције. Могуће је да једна ментална репрезентација у једном домену буде комплетна, а у другом не. У складу са уведеним критеријумима, дефинишу се следеће класе менталних репрезентација.

(iii) а) Базична комплетна ментална репрезентација, коју ћемо у наставку текста називати **пропозицијом** и означавати са p , мора да задовољи следеће услове:

- изражава комплетно значење исказа у посматраном домену,
- кардиналност свих садржаних семантичких конституената је 1,
- ниједан њен подскуп није пропозиција (тј. ниједан њен подскуп не изражава комплетно значење).

Пример пропозиције је:

$$p = \{\{\text{квадраи}\}, \{\text{иџомери}\}, \{\text{улево}\}\}, \quad (5)$$

у значењу „помери квадрат улево“.

(iii) б) Комплексна комплетна ментална репрезентација мора да задовољи следећа ограничења:

- изражава комплетно значење исказа у посматраном домену,
- кардиналност бар једног садржаног конституената већа је од 1.

Пример комплексне комплетне менталне репрезентације је:

$$m = \{\{\text{квадраџ}, \text{џроуџао}\}, \{\text{џомери}\}, \{\text{улево}\}\}, \quad (6)$$

у значењу „помери квадрат и троугао улево“. Свака комплексна комплетна ментална репрезентација може бити декомпонована на скуп одговарајућих базичних пропозиџија (у овом примеру, то су: „помери квадрат улево“ и „помери троугао улево“).

(iii) в) Базична некомплетна ментална репрезентација мора да задовољи следећа ограничења:

- не изражава комплетно значење исказа у посматраном домену,
- кардиналност свих садржаних семантичких конституената је 1.

Примери базичне некомплетне менталне репрезентације су:

$$m = \{\{\text{џомери}\}, \{\text{улево}\}\}, \quad m = \{\{\text{квадраџ}\}\}, \quad (7)$$

у значењима „помери улево“, односно „квадрат“. Као што се може видети у овом примеру, некомплетна ментална репрезентација може да садржи и само један семантички конституент.

(iii) г) Комплексна некомплетна ментална репрезентација мора да задовољи следећа ограничења:

- не изражава комплетно значење исказа у посматраном домену,
- кардиналност бар једног садржаног конституената већа је од 1.

Пример комплексне некомплетне менталне репрезентације је:

$$m = \{\{\text{џомери}\}, \{\text{квадраџ}, \text{џроуџао}\}\}, \quad (8)$$

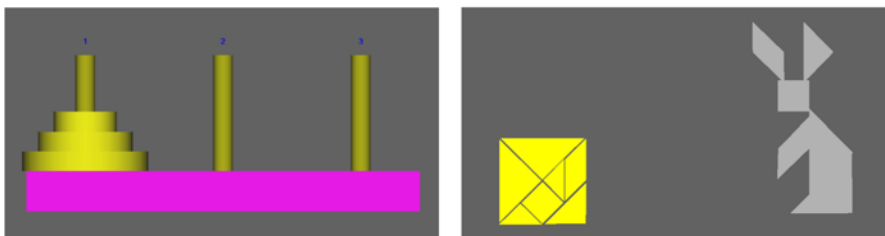
у значењу „помери квадрат и троугао“. Ова репрезентација је комплексна, јер може да се декомонује на две базичне репрезентације („помери троугао“ и „помери квадрат“) и непотпуна, јер није дефинисан смер померања фигура.

3.2 Представљање знања о домену интеракције

У предложеном меморијском моделу улазни скуп информација, који представља почетну базу знања, задаје се као прави подскуп пропозиција (тј. базичних комплетних менталних репрезентација) које делимично описују посматрани домен интеракције. Процес припреме улазног скупа који у општем случају обухвата аквизицију и обраду података различитих сензорских модалитета, не представља тему дисертације, односно предложени приступ претпоставља да је почетна база знања идеално позната. Треба приметити да почетна база знања није комплетна – она не садржи све пропозиције које су присутне у посматраном домену интеракције нити експлицитно задате информације о релацијама између семантичких ентитета. Комплетирање скупа пропозиција и извођење релација између ентитета представља резултат механизма учења предложеног у секцији 3.3. Осим тога, предложени формат за симболичку репрезентацију знања је амодалан, што значи да се стимуланси различитих сензорских модалитета кодују и складиште на униформни начин. Једна ментална репрезентација може бити формирана од семантичких ентитета који су прикупљени помоћу различитих типова сензора, нпр. микрофона или машинске визије.

Предложени модел меморије је илустрован на једноставном али реалистичном домену интеракције. Домен је преузет из корпуса НИМИТЕК, корпуса афективног понашања у интеракцији између човека и машине, на природном језику [98]. Овај корпус садржи снимке интеракције добијене применом симулационе технике „чаробњак из Оза“, у којој су људски субјекти ступали у интеракцију са симулираним конверзационим агентом, док су решавали графичке задатке приказане на екрану. Од низа задатака коришћених у овој студији, изабрана су два задатка као примери домена интеракције за илустрацију предложеног модела меморије. Први задатак се односи на слагалицу Танграм, а други

на проблем ханојских кула. За потребе илустрације, без губитка општости, у ова два домена су уведене извесне рестрикције. У слагалици Танграм се разматрају само два дводимензионална објекта (троугао и квадрат), који се могу транслирати у четири смера у равни (улево, удесно, нагоре, надоле) или ротирати у равни (у смеру кретања казаљке на сату и у супротном смеру). Редуковани проблем ханојских кула садржи само три диска. Увођење рестрикција не утиче на валидност изнетих закључака, него само на свођење изабраних примера на адекватну меру за презентацију. Ради јасноће излагања, на слици 3.2 су илустроване посматране слагалице. Илустрације су преузете из [99].



Сл. 3.2 Верзија проблема ханојских кула са три диска (лево) и пример слагалице Танграм (десно). Сlike су преузете из [99].

Пропозиције које представљају ова два домена интеракције су дате у табели 3.1. У табели је приказан комплетни скуп пропозиција који у потпуности описује дате домene, а у последњој колони табеле („Улаз“) назначен је подскуп пропозиција који је коришћен као улазни скуп за демонстрацију предложених алгоритама. Основна претпоставка алгоритама је да се улазни скуп пропозиција произвољно бира, што је детаљније дискутовано у 5. поглављу. У посматраном примеру изабрани подскуп садржи 12 од укупно 21 пропозиције (тј. приближно 51% укупног броја пропозиција). Модел нема унапред задату информацију о тачном броју различитих домена интеракције којима припадају пропозиције из улазног скупа и не поседује никакво екстерно знање о доменима (тј. кључне речи као што су *диск* и *улево* су разумљиве за човека, али немају никакво значење у оквиру модела који их третира само као симболе).

Табела 3.1 Комплетни скупи позиција који представља селективне домене интеракције повезане са Танграм слицицом (p_1 - p_{12}) и проблемом ханојских кула (p_{13} - p_{21}). У претпоследњој колони је означен додски позиција, који је коришћен као улаз за коинтерактивни ајенја.

ID	Пропозиција	Опис	Улаз
p_1	{троугао, помери, улево}	помери троугао улево	✓
p_2	{троугао, помери, удесно}	помери троугао удесно	✓
p_3	{троугао, помери, нагоре}	помери троугао нагоре	✓
p_4	{троугао, помери, надоле}	помери троугао надоле	
p_5	{троугао, ротирај, позитивно}	ротирај троугао у позитивном смеру	✓
p_6	{троугао, ротирај, негативно}	ротирај троугао у негативном смеру	✓
p_7	{квадрат, помери, улево}	помери квадрат улево	✓
p_8	{квадрат, помери, удесно}	помери квадрат удесно	
p_9	{квадрат, помери, нагоре}	помери квадрат нагоре	
p_{10}	{квадрат, помери, надоле}	помери квадрат надоле	✓
p_{11}	{троугао, ротирај, позитивно}	ротирај квадрат у позитивном смеру	
p_{12}	{троугао, ротирај, негативно}	ротирај квадрат у негативном смеру	
p_{13}	{диск ₁ , стави}	постави диск ₁ на штап ₁	✓
p_{14}	{диск ₁ , стави}	постави диск ₁ на штап ₂	✓
p_{15}	{диск ₁ , стави}	постави диск ₁ на штап ₃	
p_{16}	{диск ₂ , стави}	постави диск ₂ на штап ₁	✓
p_{17}	{диск ₂ , стави}	постави диск ₂ на штап ₂	
p_{18}	{диск ₂ , стави}	постави диск ₂ на штап ₃	✓
p_{19}	{диск ₃ , стави}	постави диск ₃ на штап ₁	✓
p_{20}	{диск ₃ , стави}	постави диск ₃ на штап ₂	
p_{21}	{диск ₃ , стави}	постави диск ₃ на штап ₃	

3.3 Моделовање домена интеракције – механизам учења

Предложени механизам учења обавља у реалном времену, аутоматско моделовање домена у интеракцији између човека и робота, само на основу некомплетног улазног скупа пропозиција. Функционалност алгорита, на којем се механизам заснива, укључује:

- категоризацију семантичких ентитета,
- асоцијативно учење у циљу извођења недостајућих пропозиција за сваки од домена интеракције.

Ови задаци се извршавају у два корака: у процесу категоризације семантичких ентитета и процесу асоцијативног учења.

3.3.1 Категоризација семантичких ентитета

Предложени механизам категоризације успоставља релације између семантичких ентитета, тј. групише ентитете који припадају истој семантичкој категорији. Припадност ентитета истој семантичкој категорији није експлицитно задата, него се утврђује током процеса интеграције менталних репрезентација предложеног у наставку овог поглавља. Пошто модел манипулише симболима којима није придружено специфично знање о домену интеракције, модел не утврђује природу семантичке категорије, него само утврђује који ентитети припадају *истој* семантичкој категорији.

У наставку су дефинисане релација *пропозиционе повезаности менталних репрезентација* и операције њихове *интеграције* и *семантичке категоризације*.

Пропозициона повезаност и интеграција менталних репрезентација.

Сматра се да су пропозиција p_i (тј. базична комплетна ментална репрезентација) и ментална репрезентација m_j (која није нужно базична) пропозиционо повезане ако су испуњени следећи услови:

- p_i и m_j су исте кардиналности,

- постоји само један конституент $sc_x \in p_i$ који није подскуп неког конституента у m_j ,
- постоји само један конституент $sc_y \in m_j$ који није надскуп неком конституенту у p_i .

Ови услови се могу формално изразити на следећи начин:

$$p_i \sim m_j \Leftrightarrow \left[\begin{array}{l} (|p_i| = |m_j|) \wedge (\exists! sc_x \in p_i) (\forall sc \in m_j) (sc_x \not\subset sc) \wedge \\ \wedge (\exists! sc_y \in m_j) (\forall sc \in p_i) (sc \not\subset sc_y) \end{array} \right]. \quad (9)$$

Имајући у виду да сваки семантички ентитет у менталној репрезентацији припада одређеној семантичкој класи, наведени услови одражавају „контекстуалну сличност“ менталних репрезентација која се заснива на претпоставци да семантички конституенти sc_x и sc_y , који не припадају пресеку посматраних менталних репрезентација, вероватно припадају истој семантичкој категорији те стога могу бити интегрисани. Другим речима, две пропозиционо повезане менталне репрезентације p_i и m_j могу се интегрисати у следећу менталну репрезентацију:

$$m := (m_j \setminus \{sc_y\}) \cup \{sc_x \cup sc_y\}, \quad sc_x \in p_i, \quad sc_y \in m_j. \quad (10)$$

Интеграција две пропозиционо повезане менталне репрезентације, p_1 и m_2 , илустрована је следећим примером:

пропозиција:

$$p_1 = \{\{\text{квадрант}\}, \{\text{помери}\}, \{\text{нагоре}\}\} \quad (11)$$

комплексна ментална репрезентација:

$$m_2 = \{\{\text{квадрант}, \text{троугао}\}, \{\text{помери}\}, \{\text{улево, удесно}\}\}$$

$$\begin{array}{l} sc_x = \{\text{нагоре}\}, (sc_x \in p_1) \wedge (\forall sc \in m_2) (sc_x \not\subset sc) \\ sc_y = \{\text{улево, удесно}\}, (sc_y \in m_2) \wedge (\forall sc \in p_1) (sc \not\subset sc_y) \end{array} \quad (12)$$

$$m = \{\{\text{квадрант}, \text{троугао}\}, \{\text{помери}\}, \{\text{улево, удесно, нагоре}\}\}. \quad (13)$$

Ако је познато да квадрат може да се помери нагоре и да се квадрат и троугао могу померити улево и удесно, закључује се да се троугао такође може померати нагоре. Формално речено, резултат интеграције је додељивање ентитета *нагоре* истој семантичкој категорији којој

припадају ентитети *улево* и *у десно*, што значи да је изведена једна интракатегоријална релација која није била дата у улазном скупу података. Интеграција пропозиционо повезаних менталних репрезентација пружа основу за њихову семантичку категоризацију.

Семантичка категоризација. Нека је M улазни скуп менталних репрезентација. Процес семантичке интеграције резултује скупом интегрисаних, у смислу јед. (10), менталних репрезентација M_i и може се описати на следећи начин:

- Корак 1: M_i се иницијализује празним скупом.
- Корак 2: Ако је скуп M празан, значи да је процес семантичке категоризације завршен и тада скуп M_i садржи интегрисане менталне репрезентације. У супротном се прелази на следећи корак.
- Корак 3: Произвољна ментална репрезентација m_x се бира и уклања из скупа M . Прелази се на корак 4.
- Корак 4: Док год постоји ментална репрезентација m_S у M која је пропозиционо повезана са m_x , репрезентација m_S се уклања из M и интегрише у m_x , у складу са јед. (10). У супротном, m_x се додаје у M_i и прелази се на корак 2.

Овај алгоритам је формално описан псеудокодом приказаним на слици 3.3.

Описани алгоритам је илустрован за изабране домене интеракције и улазни скуп пропозиција, назначен у табели 3.1. У овом примеру, M представља само подскуп свих пропозиција које описују посматране домене интеракције (тј. обухвата 12 од укупно 21 пропозиције). Резултујући скуп који садржи интегрисане менталне репрезентације је:

$$M_I = \{m_{x1}, m_{x2}, m_{x3}\}, \quad (14)$$

где су:

$$\begin{aligned}
m_{x1} &= \left\{ \left\{ \text{квадрант}, \text{проулао} \right\}, \left\{ \text{йомери} \right\}, \right. \\
&\quad \left. \left\{ \text{нагоре}, \text{надоле}, \text{улево}, \text{удесно} \right\} \right\} \\
m_{x2} &= \left\{ \left\{ \text{проулао} \right\}, \left\{ \text{роширај} \right\}, \left\{ \text{йозиийвно}, \text{нейайивно} \right\} \right\}. \quad (15) \\
m_{x3} &= \left\{ \left\{ \text{диск}_1, \text{диск}_2, \text{диск}_3 \right\}, \left\{ \text{шйай}_1, \text{шйай}_2, \text{шйай}_3 \right\} \right\}
\end{aligned}$$

Улаз: M – улазни скуп пропозиција
Изназ: M_I – резултујући скуп интегрисаних
менталних репрезентација
Алгоритам:
 $M_I = \emptyset$
while ($M \neq \emptyset$) {
 селектуј менталну репрезентацију $m_x \in M$;
 $M := M \setminus \{m_x\}$;
 while ($M \neq \emptyset$) \wedge ($\exists m_s \in M$) ($m_s \sim m_x$) {
 селектуј $m_s \in M \mid (m_s \sim m_x)$
 $m_x :=$ интегриши (m_s, m_x) ;
 $M := M \setminus \{m_s\}$;
 }
 $M_I := M_I \cup \{m_x\}$;
}
return M_I ;

Сл. 3.3 Алгоритам семантичке категоризације. Релација пропозиционе повезаности (\sim) дефинисана је једначином (9). Процедура интеракције менталних репрезентација дефинисана је једначином (10).

Предложени алгоритам интегрише пропозиције p_1, p_2, p_3, p_7 и p_{10} из улазног скупа у комплексну менталну репрезентацију m_{x1} . Треба приметити да би се интеграцијом већег подскупа пропозиција $\{p_1, p_2, p_3, p_4, p_7, p_8, p_9, p_{10}\}$ генерисала идентична ментална репрезентација m_{x1} . То значи да, иако су само p_1, p_2, p_3, p_7 , и p_{10} биле интегрисане, декомпозицијом m_{x1} могу се извести све пропозиције из скупа $\{p_1, p_2, p_3,$

$p_4, p_7, p_8, p_9, p_{10}$ (процес декомпозиције је разматран у следећој секцији). Другим речима, недостајуће пропозиције $\{p_4, p_8, p_9, p_{10}\}$ које нису део улазног скупа података могу се извести закључивањем. Слично опажање важи за менталну репрезентацију m_{x3} . Иако су интегрисане само пропозиције $\{p_{13}, p_{14}, p_{16}, p_{18}, p_{19}\}$, недостајуће пропозиције $\{p_{15}, p_{17}, p_{20}, p_{21}\}$ могу се посредно извести.

Треба истаћи још једно важно опажање везано за m_{x1} и m_{x3} . Ако две пропозиције припадају различитим доменима интеракције, оне ће бити интегрисане у две различите комплексне менталне репрезентације (нпр. m_{x1} се односи на слагалицу Танграм, а m_{x3} на проблем ханојских кула). Ипак, ова импликација није реверзибилна, тј. пропозиције које припадају истом домену интеракције не морају нужно бити интегрисане у исту комплексну менталну репрезентацију (нпр. m_{x1} и m_{x2} се односе на слагалицу Танграм). Другим речима, могуће је да се више интегрисаних менталних репрезентација односи на исти домен интеракције.

Коначно, резултујући скуп интегрисаних менталних репрезентација не обухвата нужно све пропозиције из посматраних домена интеракције. На пример, комплексна ментална репрезентација m_{x2} која интегрише p_5 и p_6 некомплетна је у таквом смислу да пропозиције p_{11} и p_{12} не могу бити изведене из ње. Како би се извеле недостајуће менталне репрезентације, предложен је механизам асоцијативног учења.

3.3.2 Асоцијативно учење менталних репрезентација

У оквиру модела је имплементиран механизам асоцијативног учења који додатно проширује семантичке конституенте новим семантичким ентитетима, на основу закључивања по аналогiji.

Асоцијативно учење. У складу са дефиницијом семантичког конституента изложеном у секцији 3.1, ако је пресек два семантичка конституента непразни скуп, може се закључити да сви семантички ентитети из њихове уније припадају истој семантичкој категорији. На основу овог опажања, асоцијативно учење је концептуализовано као транзитивно затварање скуповне операције пресека.

Нека је m_i^* скуп који садржи све семантичке конституенте који се појављују у интегрисаним менталним репрезентацијама у оквиру скупа M_i , односно:

$$m_i^* \in \bigcup_{m \in M_i} m. \quad (16)$$

Два семантичка конституента sc_1 и sc_2 који припадају m_i^* су асоцијативно повезана ако је задовољен један од наведених услова:

- sc_1 и sc_2 имају непразни пресек,
- постоји семантички конституент $sc_3 \in m_i^*$, такав да је sc_3 асоцијативно повезан са sc_1 и sc_2 .

Други, рекурзивни услов се може формулисати и у итеративној форми (примењеној у алгоритму приказаном на сл. 3.4). Нека је дат скуп $\{sc_1, sc_2, \dots, sc_k\} \subset m_i^*$ за који важи следеће:

$$\begin{aligned} & (sc_a \cap sc_1 \neq \emptyset) \wedge (sc_1 \cap sc_2 \neq \emptyset) \wedge (sc_2 \cap sc_3 \neq \emptyset) \wedge \dots \\ & \dots \wedge (sc_{k-1} \cap sc_k \neq \emptyset) \wedge (sc_k \cap sc_b \neq \emptyset) \end{aligned} \quad (17)$$

Лако се доказује да је релација асоцијативне повезаности релација еквиваленције (тј. да је рефлексивна, симетрична и транзитивна). Она формира класе еквиваленције над скупом m_i^* тако да су семантички конституенти у свакој класи асоцијативно повезани, тј. садрже семантичке ентитете који припадају истој семантичкој категорији. Стога, као резултат асоцијативног учења, сваки семантички конституент у скупу m_i^* проширен је тако да садржи све семантичке ентитете из своје класе еквиваленције. Предложени алгоритам за асоцијативно учење је илустрован псеудокодом на слици 3.4.

За потребе илустрације овог алгоритма користиће се пример из претходне секције. Скуп интегрисаних менталних репрезентација M_I [в. јед. (14-15)] представља улаз за алгоритам асоцијативног учења. Скуп m_i^* је иницијализован на следећу вредност:

$$\begin{aligned}
m_i^* &= m_{x1} \cup m_{x2} \cup m_{x3} \\
&= \{ \{ \text{квадрави, широуіао} \}, \{ \text{іомери} \}, \{ \text{наіоре, надоле, улево, удесно} \}, \\
&\quad \{ \text{іроуіао} \}, \{ \text{роіірај} \}, \{ \text{іозійіивно, неіайіивно} \}, \\
&\quad \{ \text{диск}_1, \text{диск}_2, \text{диск}_3 \}, \{ \text{ішйай}_1, \text{ішйай}_2, \text{ішйай}_3 \} \}. \quad (18)
\end{aligned}$$

Улаз: M_I – скуп интегрисаних менталних репрезентација

Излаз: M_L – скуп менталних репрезентација добијених асоцијативним учењем

Алгоритам:

$M_L = M_I$;

$m_I^* = \bigcup_{m \in M_I} m$

while ($m_I^* \neq \emptyset$) {

 селектуј семантички конституент $sc_x \in m_I^*$;

$m_I^* := m_I^* \setminus \{sc_x\}$;

while ($\exists sc_s \in m_I^* ((sc_s \cap sc_x) \neq \emptyset)$) {

 select $sc_s \in m_I^* | ((sc_s \cap sc_x) \neq \emptyset)$

$sc_x := sc_x \cup sc_s$;

$m_I^* := m_I^* \setminus \{sc_s\}$;

 }

for each $sc \in \bigcup_{m \in M_L} m$ {

 if $sc \cap sc_x \neq \emptyset$

$sc = sc_x$;

 }

}

return M_L ;

Сл. 3.4 Алгоритам асоцијативног учења.

Само су два семантичка конституента у скупу m_i^* асоцијативно повезана: $\{ \text{квадрави, широуіао} \} \in m_{x1}$ и $\{ \text{іроуіао} \} \in m_{x2}$. Према предложеном алгоритму, они су проширени тако да обухвате све семантичке

ентитете из својих класа еквиваленџије. Резултујући скуп менталних репрезентација проширен закључивањем по аналогији је:

$$M_L = \{m_{L1}, m_{L2}, m_{L3}\}, \quad (19)$$

где су:

$$\begin{aligned} m_{L1} &= \left\{ \left\{ \text{квадрај}, \text{џроуџао} \right\}, \left\{ \text{џомери} \right\}, \right. \\ &\quad \left. \left\{ \text{наџоре, надоле, улево, угесно} \right\} \right\} \\ m_{L2} &= \left\{ \left\{ \text{квадрај}, \text{џроуџао} \right\}, \left\{ \text{роџирај} \right\}, \left\{ \text{џозиџивно, неџаџивно} \right\} \right\}. \quad (20) \\ m_{L3} &= \left\{ \left\{ \text{диск}_1, \text{диск}_2, \text{диск}_3 \right\}, \left\{ \text{џџаџ}_1, \text{џџаџ}_2, \text{џџаџ}_3 \right\} \right\} \end{aligned}$$

Ако се упореде менталне репрезентације (15) и (20), може се приметити да је, као резултат асоџијативног учења у овом примеру, семантички конституент $\{\text{џроуџао}\}$ у m_{x2} проширен у $\{\text{квадрај}, \text{џроуџао}\}$, који припада у m_{L2} . Важно је истаћи да су пропозиџије p_{11} и p_{12} (в. табелу 3.1), које нису биле интегрисане у скупу M_i , сада обухваћене скупом менталних репрезентација добијених асоџијативним учењем M_L . У даљем тексту се разматра како се пропозиџије изводе из скупа M_L и како се реконструише комплетни скуп пропозиџија за посматране домене инџеракџије.

Извођење пропозиџија. Нека је $m_L = \{sc_1, sc_2, \dots, sc_k\}$ комплексна ментална репрезентација. Она се може разложити на скуп пропозиџија, тј. базичних менталних репрезентација. Да би се пропозиџија $p = \{sc_{b1}, sc_{b2}, \dots, sc_{bk}\}$ извела из m_L , мора да буде задовољен следећи услов:

$$(\forall 1 \leq i \leq k) (|sc_{bi}| = 1 \wedge sc_{bi} \subset sc_i). \quad (21)$$

Стога, број пропозиџија које се могу извести из m_L је једнак $\prod_{i=1}^k |sc_i|$, где

$sc_i \in m_L$. На пример, из комплексне менталне репрезентације:

$$m_{L2} = \left\{ \left\{ \text{квадрај}, \text{џроуџао} \right\}, \left\{ \text{роџирај} \right\}, \left\{ \text{џозиџивно, неџаџивно} \right\} \right\}, \quad (22)$$

могу се извести четири пропозиџије (в. табелу 3.1):

$$\begin{aligned}
 p_5 &= \{\bar{u}prouiao, ro\bar{u}ira\bar{a}j, \bar{u}ozii\bar{u}ivno\} \\
 p_6 &= \{\bar{u}prouiao, ro\bar{u}ira\bar{a}j, ne\bar{u}ai\bar{u}ivno\} \\
 p_{11} &= \{\bar{k}vadrai\bar{u}, ro\bar{u}ira\bar{a}j, \bar{u}ozii\bar{u}ivno\} \\
 p_{12} &= \{\bar{k}vadrai\bar{u}, ro\bar{u}ira\bar{a}j, ne\bar{u}ai\bar{u}ivno\}
 \end{aligned} \tag{23}$$

Из комплексних менталних репрезентација садржаних у скупу M_L [в. јед. (20)] се на овај начин могу извести све пропозиције из табеле 3.1, укључујући и оне које нису биле дате у улазном скупу.

Две битне одлике процеса извођења јесу да се не генеришу редундантне пропозиције нити се нарушава уведено ограничење, према ком ниједан подскуп пропозиције не сме сам бити пропозиција. Нека су m_{L1} и m_{L2} две различите комплексне менталне репрезентације, генерисане предложеним алгоритмима семантичке категоризације и асоцијативног учења и нека су p_1 и p_2 пропозиције изведене из m_{L1} и m_{L2} , респективно. Ако би p_1 било једнако p_2 , онда би обе репрезентације биле садржане у m_{L1} и у m_{L2} . Стога, према алгоритму асоцијативног учења, m_{L1} и m_{L2} би биле интегрисане у једну заједничку комплексну менталну репрезентацију, што је супротно претпоставци да су m_{L1} и m_{L2} различите. Слично се може показати да p_1 не може бити подскуп p_2 .

Важно је нагласити да предложени алгоритам учења извођењем аналогичја може извести неке пропозиције које нису нужно присутне у посматраном домену. На пример, претпоставимо да је систему саопштен следећи улазни скуп пропозиција:

$$\begin{aligned}
 p_1 &= \{\{\bar{u}o\bar{u}lega\bar{a}j\}, \{\bar{u}ore\}\} \\
 p_2 &= \{\{\bar{u}o\bar{u}lega\bar{a}j\}, \{\bar{u}ole\}\}, \\
 p_3 &= \{\{\bar{u}eg\bar{u}\}, \{\bar{u}ole\}\}
 \end{aligned} \tag{24}$$

Након процеса семантичке категоризације и асоцијативног учења, систем би извео следећу комплексну менталну репрезентацију:

$$m_{L1} = \{\{\bar{u}o\bar{u}lega\bar{a}j, \bar{u}eg\bar{u}\}, \{\bar{u}ore, \bar{u}ole\}\}. \tag{25}$$

Декомпоновањем ове комплексне репрезентације изводи се следећи скуп пропозиција:

$$\begin{aligned}
 p_1 &= \{\{\bar{i}o\bar{l}e\bar{d}a\bar{j}\}, \{\bar{i}o\bar{r}e\}\}, \\
 p_2 &= \{\{\bar{i}o\bar{l}e\bar{d}a\bar{j}\}, \{\bar{g}o\bar{l}e\}\}, \\
 p_3 &= \{\{\bar{s}e\bar{g}i\}, \{\bar{i}o\bar{r}e\}\}, \\
 p_4 &= \{\{\bar{s}e\bar{g}i\}, \{\bar{g}o\bar{l}e\}\}.
 \end{aligned}
 \tag{26}$$

Валидност научене пропозиције p_3 зависи од конкретног домена који се моделује, што у општем случају не представља методолошко ограничење, него је последица чињенице да је предложени приступ независан од домена интеракције који се моделује, тј. не користи унапред задато знање специфично за домен, него само врши обраду над скуповима симбола. Ово није ограничавајући фактор за практичне примене модела, јер се за конкретне домене интеракције могу дефинисати додатни, контекстно зависни критеријуми који би елиминисали пропозиције које нису дозвољене у посматраном домену.

3.4 Кодовање и складиштење информација

Описаним механизмом учења се успостављају релације између семантичких ентитета из задатог улазног скупа. Резултат извршавања овог алгоритма представља скуп научених, односно комплексних менталних репрезентација. Свака од њих садржи у себи скуп пропозиција које припадају истом домену интеракције. У овој секцији уводимо појам *усмерене семантичке мреже*, на којем су засновани кодовање и складиштење научених репрезентација, у предложеном моделу.

3.4.1 Усмерене семантичке мреже

Усмерена семантичка мрежа, коју ћемо у даљем тексту поједностављено називати семантичком мрежом, може се формално дефинисати као повезани, ациклични, усмерени граф:

$$SM = (V, E), \tag{27}$$

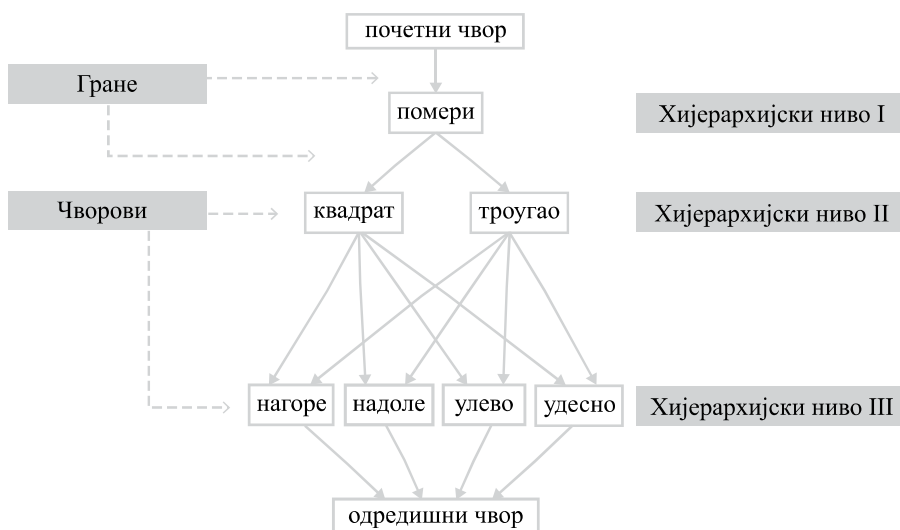
за који важи:

- V је скуп чворова,

- $E \subset V \times V$ је скуп грана такав да је граф G ацикличан,
- чвор v_p је једини чвор у V који има само излазне гране, чвор v_o је једини чвор у V који има само улазне гране, док сви остали чворови у V имају и улазне и излазне гране.

У контексту предложеног модела, чворови семантичке мреже представљају семантичке ентитете, док су релације између ентитета представљене усмереним гранама, а концептуализација ових релација је детаљније размотрена у следећој секцији. На слици 3.5 је дата семантичка мрежа која представља домен интеракције који је везан за редуковану верзију слагалице Танграм, разматрану у овом поглављу. Такође, назначени су основни елементи мреже.

У приступу предложеном у овој дисертацији, топологија семантичке мреже и релације између семантичких ентитета нису унапред дати, него произлазе из предложеног алгоритма учења, тј. аутоматског моделовања домена интеракције. Поступак аутоматског генерисања одговарајућег скупа семантичких мрежа, а на основу скупа научених комплексних менталних репрезентација, описан је у следећој секцији.



Сл. 3.5 Семантичка мрежа која одговара домену интеракције везаном за редуковану верзију слагалице Танграм са означеним основним елементима мреже.

3.4.2 Аутоматско генерисање семантичких мрежа

Алгоритам за аутоматско генерисање семантичких мрежа прихвата, као улазни податак, скуп научених, комплексних менталних репрезентација M_L [в. јед. (20)] и пресликава сваку менталну репрезентацију m_L из скупа M_L у одговарајућу семантичку мрежу. У општем случају, укупни садржај дуготрајне меморије представља знање о непразном скупу домена интеракције. У зависности од своје комплексности, сваки домен интеракције је представљен једном семантичком мрежом или скупом семантичких мрежа. Формалније исказано, алгоритам за аутоматско генерисање семантичких мрежа дефинише сурјективно пресликавање скупа генерисаних семантичких мрежа у скуп моделованих домена интеракције.

Описани процеси семантичке категоризације и асоцијативног учења успостављају релације између семантичких ентитета, које је потребно очувати током пресликавања менталне репрезентације у одговарајућу семантичку мрежу. Ово се постиже тако што се сви семантички конституенти (тј. ентитети садржани у њима) из комплексне менталне репрезентације пресликавају на засебне хијерархијске нивое семантичке мреже. Међутим, ментална репрезентација се дефинише као скуп конституената и не садржи специфичне информације о релацијама између самих конституената, а репрезентација у форми семантичке мреже захтева секвенцијално уређење семантичких конституената. Другим речима, неопходно је дефинисати бијективно пресликавање скупа конституената садржаних у менталној репрезентацији на скуп хијерархијских нивоа у семантичкој мрежи, која моделује посматрану менталну репрезентацију. Као критеријум за одређивање овог редоследа коришћен је један увид из модела фокусног стабла [1]. Усклађивање са моделом фокусног стабла мотивисано је циљем да се предложени модел дуготрајне меморије интегрише са моделом радне меморије заснованим на фокусном стаблу. Пошто у фокусном стаблу чворови на вишим нивоима хијерархије представљају општије семантичке ентитете у односу на своје потомке, исти однос је примењен и на семантичку мрежу, односно општији ентитети се пресликавају на више хијерархијске нивое у мрежи (тј. ближе почетном чвору), док се специфичнији ентитети пресликавају дубље (тј. ближе одредишном чвору).

У предложеном приступу општост семантичког конституента је обрнуто пропорционална његовој кардиналности. Ово правило рефлектује чињеницу да општост конституента зависи од домена интеракције, тј. да су општији они семантички ентитети који се јављају у већем броју пропозиција унутар једног домена интеракције. Због тога уводимо парцијално уређени скуп m_L^* који садржи исте семантичке конституенте као и дата комплексна ментална репрезентација m_L , али уређене у монотону неоппадајући низ, на основу њихових кардиналности:

$$m_L^* = \left\{ \{sc_1^*\}, \{sc_2^*\}, \dots, \{sc_n^*\} \mid |m_L| = n \wedge (\forall 1 < i \leq n), sc_i^* \in m_L \wedge (|sc_{i-1}^*| \leq |sc_i^*|) \right\}. \quad (28)$$

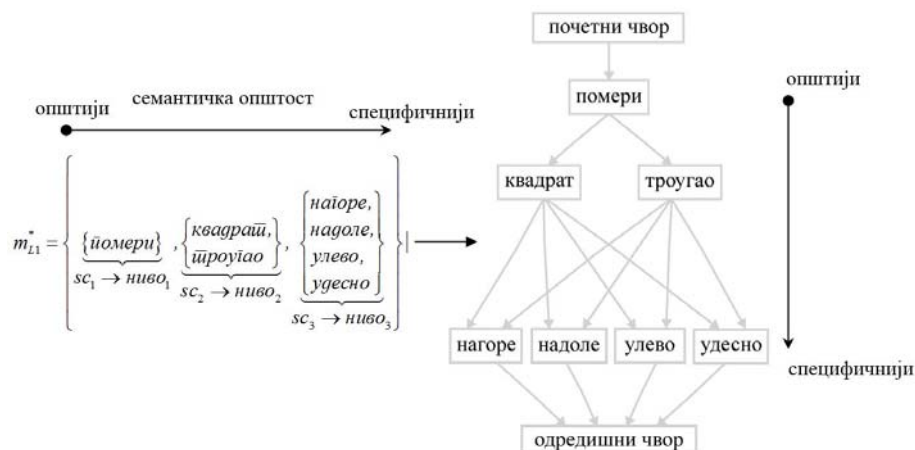
Одатле следи да конституент sc_i^* садржи семантички најопштије ентитете, док sc_n^* садржи најспецифичније ентитете. Ако два конституента имају исту кардиналност, њихов међусобни положај у уређеном скупу је произвољан. Семантичка мрежа која представља посматрану менталну репрезентацију m_L генерише се из парцијално уређеног скупа m_L^* , у складу са дефиницијом (27). Скуп чворова семантичке мреже је дефинисан на следећи начин:

$$V = \bigcup_{i=1}^n sc_i^* \cup \{v_p\} \cup \{v_o\}, \text{ где важи } |m_L| = n, (1 \leq i \leq n, sc_i^* \in m_L^*), \quad (29)$$

при чему је чвор v_p почетни чвор, а v_o одредишни чвор, тј. скуп чворова V садржи све семантичке ентитете садржане у менталној репрезентацији m_L^* , и два додатна помоћна чвора v_p и v_o , који представљају почетни, односно одредишни чвор. Скуп грана E је дефинисан на следећи начин:

$$E = \left\{ (x, y) \mid x \in sc_i^* \wedge y \in sc_{i+1}^*, i = 1, \dots, n-1 \right\} \cup \left(\{v_p, y\} \mid y \in sc_1^* \right) \cup \left\{ (x, v_o) \mid x \in sc_n^* \right\}, \text{ где важи } |m_L| = n, (1 \leq i \leq n, sc_i^* \in m_L^*) \quad (30)$$

Гране мреже повезују ентитете (чворове) из два суседна конституента у смеру од општијих ка специфичнијим ентитетима. Почетни чвор се директно повезује са свим ентитетима из најопштијег конституента, а сви ентитети из најспецифичнијег конституента су директно повезани са одредишним чвором.



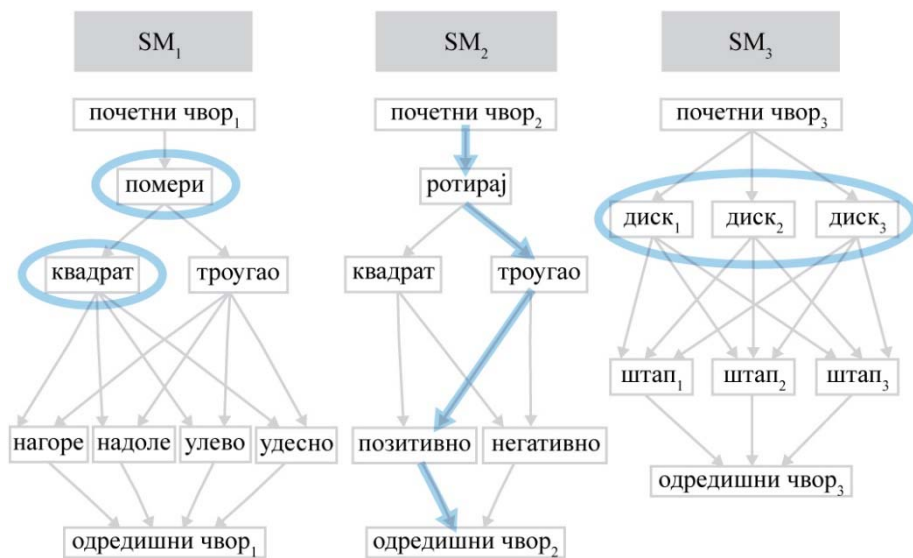
Сл. 3.6 Аутоматско генерисање семантичких мрежа. Процедура укључује одређивање одговарајућих хијерархијских нивоа у мрежи за семантичке конституенте из менталне репрезентације m_{L1} и успостављање веза између ентитета на суседним нивоима.

Аутоматско генерисање семантичких мрежа је илустровано на слици 3.6. За потребе илустрације и без губитка општости, изабрана је семантичка мрежа генерисана из менталне репрезентације m_{L1} [в. јед. (20)].

На основу скупа научених менталних репрезентација M_L [в. јед. (19-20)], за потребе илустрације је генерисан скуп семантичких мрежа који представља знање когнитивног агента о доменима интеракције описаним скупом пропозиција у табели 3.1. Овај скуп семантичких мрежа је илустрован на слици 3.7. На истој слици је приказана и графичка интерпретација свих уведених семантичких јединица, дата у контексту семантичких мрежа.

Семантички ентитети су представљени чворовима семантичке мреже, док су релације између њих представљене усмереним гранама. Усмерене путање у мрежи које повезују почетни и одредишни чвор представљају пропозиције. У процесу аутоматског генерисања семантичке мреже, пресликавање научене комплексне менталне репрезентације у семантичку мрежу представља инјективна пресликавања семантичких ентитета у чворове мреже, односно

пропозиција у путање. Репрезентација у форми семантичке мреже у потпуности чува семантички садржај комплексне менталне репрезентације из које је изведена. Као што је напоменуто раније, из једне комплексне менталне репрезентације може се извести скуп који садржи $\prod_{i=1}^k |sc_i|$ пропозиција, што произлази из правила да се пропозиције изводе комбиновањем по једног ентитета из сваког конституента. Аналогно томе, у мрежној репрезентацији се јавља тачно $\prod_{i=1}^k |sc_i|$ усмерених путања (од почетног до одредишног чвора), које покривају комплетни скуп пропозиција.



Сл. 3.7 Скупи семантичких мрежа генерисаних из скупа научених менталних репрезентација M_L [в. јед. (20)]. У оквиру приказаних семантичких мрежа назначена је графичка нотација дефинисаних семантичких јединица. а) Семантички ентитети су представљени чворовима, б) пропозиције су представљене усмереним путањама од почетној до одредишној чвора, в) семантички конституенти су представљени као групе чворова на истом хијерархијском нивоу мреже.

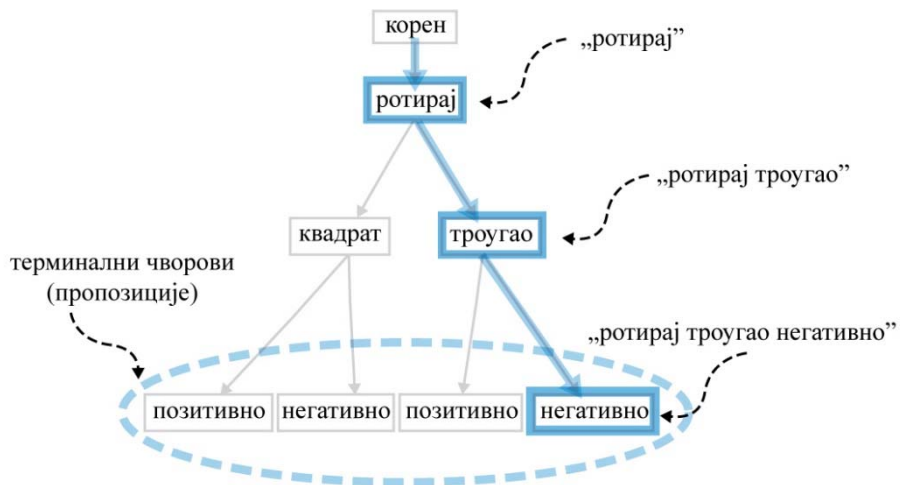
Овде треба нагласити још један аспект предложеног механизма за складиштење информација у меморији. Исти семантички ентитет се може појавити у више различитих мрежа које припадају различитим доменима интеракције. Ово је у складу са чињеницом да је семантичко значење ентитета контекстно зависно и може бити различито у различитим доменима. Вишеструко складиштење истог ентитета не узрокује редундансу, него се оправдава тиме што се ентитет складишти у оквиру различитих домена интеракције, што између осталог омогућава његову контекстно зависну интерпретацију.

3.4.3 Аутоматско генерисање фокусног стабла

Фокусно стабло је когнитивно инспирисани рачунарски модел радне меморије и механизма пажње, који укључује низ правила независних од домена интеракције за контекстно зависно интерпретирање дијалошких чинова и адаптивно управљање дијалогом у интеракцији између човека и машине. Детаљни опис модела фокусног стабла и приказ функционалности конверзационих агената, заснованих на овом моделу, доступан је у [1], [2], [3] и [100]. У овој секцији се разматрају само поједини аспекти модела фокусног стабла који су релевантни за дискусију. Фокусно стабло је хијерархијска структура која описује дијалошки домен. Пример фокусног стабла које описује домен интеракције, везан за редуковану слагалицу Танграм, дат је на слици 3.8.

Сваки чвор стабла представља један семантички ентитет из дуготрајне меморије, а свака усмерена путања од корена стабла ка листовима, представља менталну репрезентацију. Усмерена путања која почиње у корену стабла а завршава у неком терминалном чвору, представља пропозицију, тј. базичну менталну репрезентацију са комплетним значењем у посматраном домену. На слици 3.8 је подебљаном стрелицом означена пропозиција „ротирај троугао негативно“. Путања која се завршава са унутрашњим чвором одговара базичној менталној репрезентацији са некомплетним значењем [в. јед. (7)]. Као пример, наводимо менталну репрезентацију „ротирај троугао“. Чворови у оквиру истог хијерархијског нивоа фокусног стабла припадају истој семантичкој категорији, тј. представљају један семантички

конституент. Фокусно стабло на слици 3.8 има три хијерархијска нивоа, који представљају следеће семантичке конституенте: $sc_1 = \{\text{ротирај}\}$, $sc_2 = \{\text{квадрат}, \text{троугао}\}$ и $sc_3 = \{\text{позитивно}, \text{негативно}\}$.



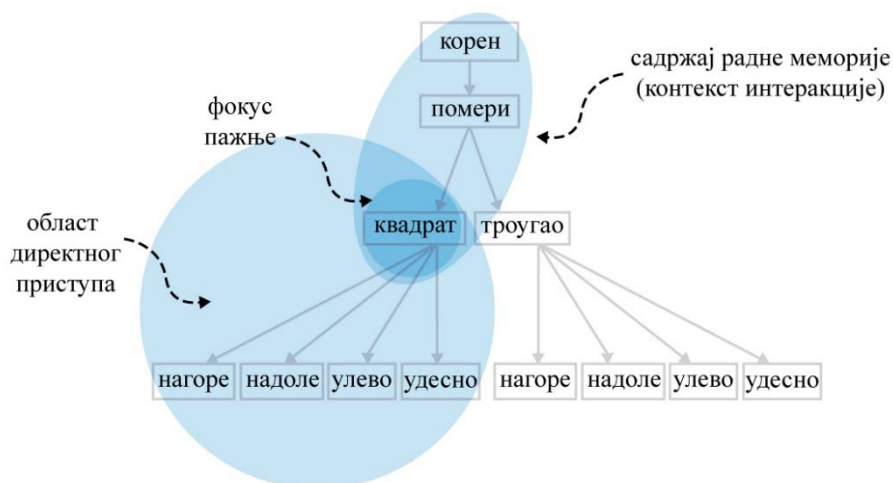
Сл. 3.8 Пример фокусног стабла које одговара семантичкој мрежи SM_2 и описује домен везан за редуковану слицицу Танграм.

Важно је нагласити да чвор у фокусном стаблу не представља само семантички ентитет (што му јесте основна интерпретација), него такође једнозначно представља менталну репрезентацију која садржи посматрани чвор и све његове претке у стаблу. Из тог разлога, ако је чвор n_A родитељ чвора n_B , тада ментална репрезентација придружена чвору n_B интегрисе и проширује менталну репрезентацију придружену чвору n_A . Ово је илустровано примером на слици 3.8. Уз сваки чвор из назначене пропозиције „ротирај троугао негативно“ наведена је ментална репрезентација коју представља. Може се приметити да у фокусном стаблу једино ментална репрезентација придружена терминалном чвору има комплетно значење.

У било ком тренутку интеракције фокус пажње је позициониран на тачно једном чвору стабла и одређује тренутни садржај радне меморије. Наиме, ако је чвор n_i тренутно у фокусу пажње, то значи да је ментална репрезентација која садржи чвор n_i и све његове претке тренутно активирана и представља садржај радне меморије. Област директног

приступа меморији садржи и чвор n_i , који је тренутно у фокусу пажње, и све његове потомке. Уведени концепти су илустровани на слици 3.9, за тренутни фокус пажње позициониран на чвору „квадрат“.

Као што је напоменуто у уводном поглављу, топологија фокусног стабла је у претходним применама била унапред задата, а на основу анализе домена или језичких корпуса [1], [2], [3]. Један од доприноса ове дисертације се односи на аутоматско генерисање фокусног стабла на основу скупа семантичких мрежа које представљају научене, комплексне менталне репрезентације из скупа M_L . У наставку описујемо како се генеришу појединачна стабла из одговарајућих семантичких мрежа.



Сл. 3.9 Фокусно стабло које одговара семантичкој мрежи SM_1 [в. сл 3.7]. Чвор „квадрат“ је произвољно изабран као тренутни фокус пажње.

Нека је m_L комплексна ментална репрезентација. Слично као у претходном поглављу [в. јед. (28)], уводимо парцијално уређени скуп m_L^* који садржи исте семантичке конституенте као и дата комплексна ментална репрезентација m_L , али уређене у монотонно неоппадајући низ, на основу њихових кардиналности:

$$m_L^* = \left\{ \left\{ sc_1^* \right\}, \left\{ sc_2^* \right\}, \dots, \left\{ sc_n^* \right\} \mid |m_L| = n \wedge (\forall 1 < i \leq n), \right. \\ \left. sc_i^* \in m_L \wedge (|sc_{i-1}^*| \leq |sc_i^*|) \right\}. \quad (31)$$

Фокусно стабло FS_{sub_i} које представља менталну репрезентацију m_L дефинисано је скупом чворова V_i и скупом грана E_i :

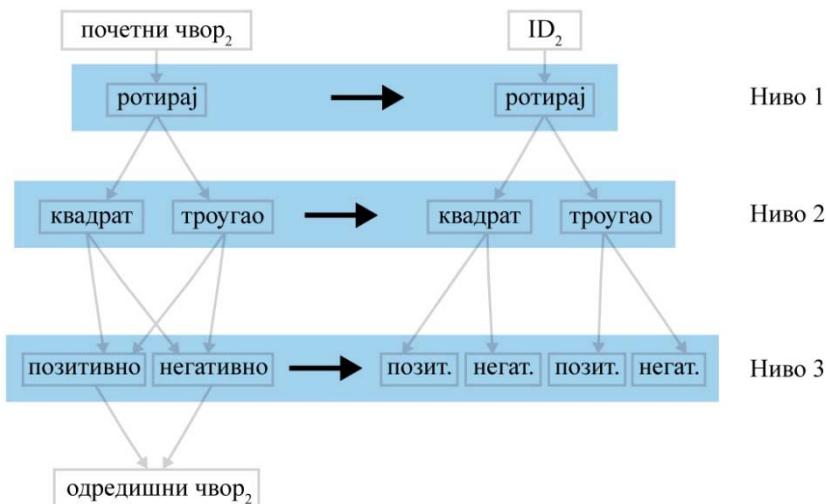
$$FS_{sub_i} = (V_i, E_i). \quad (32)$$

Скуп чворова V_i у стаблу дефинисан је Декартовим производима семантичких конституената садржаних у m_L^* , на следећи начин:

$$V_i = \left\{ \begin{aligned} & \{ID_i\} \cup (sc_1^*) \cup (sc_1^* \times sc_2^*) \cup (sc_1^* \times sc_2^* \times sc_3^*) \cup \\ & \dots \cup (sc_1^* \times sc_2^* \times \dots \times sc_K^*) \end{aligned} \right\}. \quad (33)$$

где су: $sc_1^*, sc_2^*, \dots, sc_K^* \in m_{L_i}^*$

У горњој формулацији, чвор ID_i представља корен стабла, а његов назив једнозначно одређује менталну репрезентацију (тј. семантичку мрежу) коју представља посматрано стабло. Сваки од преосталих Декартових производа представља чворове на једном хијерархијском нивоу у подстаблу. Ова дефиниција је илустрована на слици 3.10, на примеру превођења семантичке мреже SM_2 (в. сл. 3.7) у одговарајуће фокусно стабло.



Сл. 3.10 Превођење семантичке мреже SM_2 у фокусно стабло.

Кад се занемаре почетни и одредишни чвор, посматрана семантичка мрежа има три хијерархијска нивоа и представљена је парцијално уређеним скупом:

$$m_{L2}^* = \left\{ \underbrace{\{роширај\}}_{sc_1^*}, \underbrace{\{кваграи, ипроуао\}}_{sc_2^*}, \underbrace{\{иозииивно, неіаииивно\}}_{sc_3^*} \right\}. \quad (34)$$

У складу са изразом (33), чворови на сваком од три хијерархијска нивоа у фокусном стаблу су:

$$\begin{aligned} \text{ниво}_1 &\rightarrow sc_1^* = \{роширај\} \\ \text{ниво}_2 &\rightarrow sc_1^* \times sc_2^* = \\ &\quad \{роширај\} \times \{кваграи, ипроуао\} = \\ &\quad \{(роширај, кваграи), (роширај, ипроуао)\} \\ \text{ниво}_3 &\rightarrow sc_1^* \times sc_2^* \times sc_3^* = \\ &\quad \{роширај\} \times \{кваграи, ипроуао\} \times \{иозииивно, неіаииивно\} = \\ &\quad \{(роширај, кваграи, иозииивно), (роширај, кваграи, неіаииивно), \\ &\quad \{(роширај, ипроуао, иозииивно), (роширај, ипроуао, неіаииивно)\} \end{aligned} \quad (35)$$

Комплетни скуп који садржи све чворове из датог примера изгледа овако:

$$V_2 = \{ID_2\} \cup \left\{ \begin{array}{l} (роширај), \\ (роширај, кваграи), \\ (роширај, ипроуао), \\ (роширај, кваграи, иозииивно), \\ (роширај, кваграи, неіаииивно), \\ (роширај, ипроуао, иозииивно), \\ (роширај, ипроуао, неіаииивно) \end{array} \right\}. \quad (36)$$

Сваки елемент у овом скупу је представљен k -торком. Последњи члан k -торке представља семантички ентитет који се придружује посматраном чвору фокусног стабла, док преостали чланови k -торке представљају семантичке ентитете придружене његовим прецима у стаблу, тј. представљају комплетну путању од корена до тог чвора.

Оваква репрезентација једнозначно одређује сваки чвор у стаблу. На слици 3.10 се види како више чворова у фокусном стаблу може бити означено истим називом, тј. да представљају исти семантички ентитет – нпр. два чвора $\{\bar{u}ozi\bar{t}ivno\}$ на трећем нивоу стабла. Али ово нису исти чворови, пошто су представљени различитим k -торкама. Један чвор $\{\bar{u}ozi\bar{t}ivno\}$ је придружен менталној репрезентацији „ротирај квадрат позитивно”, док је други чвор $\{\bar{u}ozi\bar{t}ivno\}$ придружен менталној репрезентацији „ротирај троугао позитивно”. Другим речима, ова два чвора представљају исти семантички ентитет, али у различитим менталним репрезентацијама.

Последица овакве топологије јесте да је укупан број чворова у фокусном стаблу већи од броја чворова у семантичкој мрежи, коју представља посматрано стабло. Број чворова у стаблу је:

$$\begin{aligned} |V_i| &= 1 + |sc_1^*| + |sc_1^*||sc_2^*| + |sc_1^*||sc_2^*||sc_3^*| + \dots + |sc_1^*||sc_2^*|\dots|sc_n^*| = \\ &= 1 + |sc_1^*| + |sc_1^* \times sc_2^*| + |sc_1^* \times sc_2^* \times sc_3^*| + \dots + |sc_1^* \times sc_2^* \times \dots \times sc_n^*|. \end{aligned} \quad (37)$$

За реалне примере, повећање броја чворова у односу на семантичку мрежу може бити значајно, на шта ћемо се осврнути у поглављу 5.

Да бисмо дефинисали скуп грана фокусног стабла, потребно је приметити да се k -торке које садрже више од два семантичка ентитета могу представити као уређени пар на следећи начин:

$$n_i = (e_1, e_2, \dots, e_{k-1}, e_k) \equiv ((e_1, e_2, \dots, e_{k-1}), e_k) \equiv (n_j, e_k), \quad (38)$$

при чему, први елемент овог пара такође представља чвор стабла. Сваки такав пар у фокусном стаблу представља једну усмерену грану, која спаја чвор $(e_1, e_2, \dots, e_{k-1})$ са чвором $(e_1, e_2, \dots, e_{k-1}, e_k)$. Скуп грана E_i дефинише се на следећи начин:

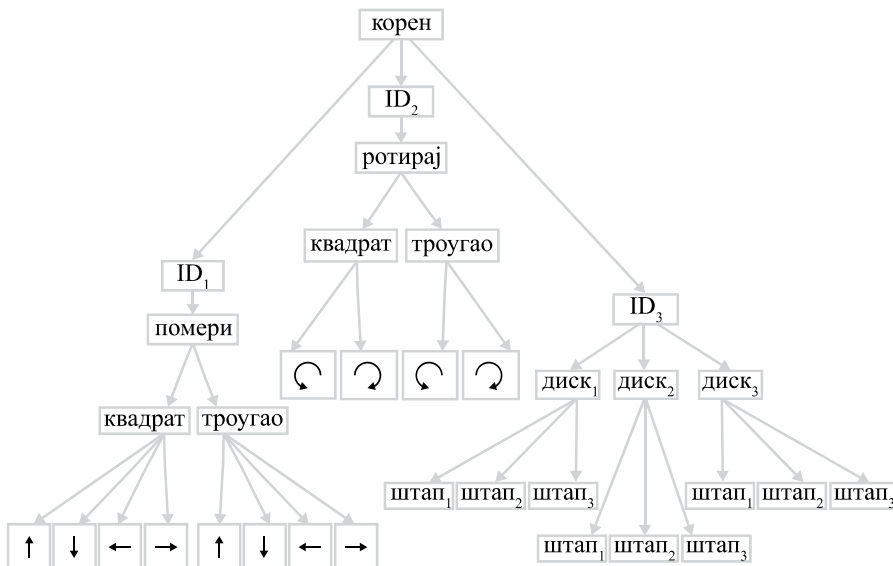
$$E_i = \{(ID_i, (e)) \mid (e) \in V_i\} \cup \{(n_i, n_j) \mid n_i, n_j \in V_i \wedge n_i = (n_j, e_i)\}. \quad (39)$$

Лако се показује да је скуп грана у фокусном стаблу једнак скупу грана у семантичкој мрежи коју стабло представља и да су скупови пропозиција, садржаних у обе структуре, идентични.

На сличан начин се и скуп семантичких мрежа може превести у фокусно стабло које се генерише спајањем (тј. додавањем надређеног корена) подстабала која, при том, представљају појединачне мреже. Овакво фокусно стабло FS које садржи S подстабала $FS_{sub1}, FS_{sub2}, \dots, FS_{subS}$, формално се представља на следећи начин:

$$\begin{aligned} FS &= (V, E), \\ V &= \{\text{корен}\} \cup V_1 \cup V_2 \cup \dots \cup V_S, \\ E &= \{(\text{корен}, ID_i) \mid \forall i = 1, 2, \dots, S\} \cup E_1 \cup E_2 \cup \dots \cup E_S \end{aligned} \quad (40)$$

где су: V и E скупови чворова и грана у фокусном стаблу FS , а V_i и E_i скупови чворова и грана у фокусном подстаблу FS_{sub_i} . Фокусно стабло генерисано на основу скупа семантичких мрежа датих на слици 3.7, приказано је на слици 3.11.



Сл. 3.11 Фокусно стабло генерисано на основу скупа семантичких мрежа датих на слици 3.7.



Валидација модела

Science is what you know, philosophy is what you don't know.

- Bertrand Russel

Science is built up of facts, as a house is with stones. But a collection of facts is no more a science than a heap of stones is a house.

- Henri Poincare

У овом поглављу је описана валидација функционалности прототипског система заснованог на предложеном моделу меморије. Прототипски систем је тестиран у две фазе: (i) на реалистичном корпусу снимака интеракције између људи и машине, а потом и (ii) на домену интеракције између терапеута и асистивног конверзационог робота, у контексту терапије за децу са сметњама у развоју.

4.1 Валидација модела на корпусу

У првој фази валидације је коришћен корпус НИМИТЕК који садржи снимке афективног понашања у вербалној интеракцији између човека и машине [98]. Овај корпус је произведен применом симулационе технике „чаробњак из Оза“. Десет наивних, здравих субјеката (7 жена, 3 мушкарца, старости од 18 до 27, средња вредност 21.7) учествовало је у студији у којој су решавали 14 графичких задатака (задаци представљају 6 различитих дијалогских домена), приказаних на екрану, при чему су субјекти са системом могли да комуницирају искључиво вербално.

Субјектима није наметнуто било какво језичко ограничење у вези са синтаксном структуром, избором речи или пропозиционим садржајем њихових дијалošких чинова, тј. било им је дозвољено да се спонтано изражавају. У [98] је показано да је корпус индикативан у односу на начин на који корисници комуницирају са дијалošким системима. Поред тога, у снимљеним интеракцијама, субјекти и систем деле не само језички, него и просторни контекст, што је карактеристично за интеракцију између човека и робота.

Субјекти су током снимања генерисали 1847 дијалošких размена, са просечним бројем од 17,19 речи по размени, стандардном девијацијом од 24,37 и речником од приближно 900 лема. Имајући на уму да једна дијалošка размена може да садржи више дијалošких чинова, анализа корпуса је показала да је 8915 вербалних дијалošких чинова (тј. 98,09 %) спонтано генерисано. Ови дијалošки чинови су класификовани у три групе: команде (6798), питања (390) и изјаве (1727). За потребе валидације је коришћена најдоминантнија класа, спонтано генерисане команде. У оквиру ове класе постоји 5469 команди које се могу сурјективно мапирати у базичне (комплетне и некомплетне) менталне репрезентације. Изабране комплетне базичне репрезентације (тј. пропозиције) коришћене су као улазни подаци на којима је валидиран прототипски систем.

Основна идеја ове фазе валидације је да се прототипском систему, као улаз, саопште изабрани подскупови пропозиција заступљених у корпусу, након чега људски субјекат проверава валидност аутоматски генерисаних фокусних стабала.

У првом делу ове фазе, дијалošки домени заступљени у корпусу НИМИТЕК посматрани су засебно. За сваки домен, прототипском систему је саопштен прво комплетни скуп пропозиција који представља посматрани домен, а потом само случајно изабрани подскуп пропозиција, који садржи приближно половину целог скупа. Ово је илустровано на слагалици Танграм, која је описана скупом од 42 пропозиције. Кад је комплетни скуп пропозиција саопштен прототипском систему, резултујуће фокусно стабло је садржало све пропозиције, и то без сувишних, а топологија стабла је одговарала очекиваној. Када је

прототипском систему саопштен случајно изабрани подскуп који садржи 22 од 42 пропозиције, прототипски систем је извео недостајуће пропозиције без генерисања сувишних пропозиција и генерисао валидно фокусно стабло, као и у претходном случају. Слично је и са верзијом проблема ханојских кула са три диска која је описана са 9 пропозиција. Валидно фокусно стабло је генерисано у случајевима када је прототипском систему саопштен комплетни скуп пропозиција, као и подскуп од 5 пропозиција. На исти начин, прототипски систем је тестиран за преостале домене (бирање фигуре која наставља дати низ фигура, класификација фигура, проблем пресипања са три посуде, уређење бројева у решеткастој слагалици – види [98]).

У другом делу ове фазе, прототипском систему су као улаз саопштавани некомплетни скупови пропозиција, које делимично описују два или више домена интеракције садржаних у корпусу. Генерисана фокусна стабла су обухватала домене, који су представљени у улазном скупу, а по садржају и топологији стабла су се слагала са фокусним стаблима која су, на основу истих улазних података, генерисали људски субјекти. Пример оваквог улаза је дат у табели 3.1, а резултујуће фокусно стабло је илустровано на слици 3.11.

4.2. Валидација у оквиру интегрисаног конверзационог ајенџа

У другој фази валидације прототипски систем је интегрисан у дијалошки систем који управља вербалном интеракцијом између корисника и два роботска система, приказана на сликама 4.1 и 4.2. Такви роботски системи су: (i) антропоморфни индустријски робот АBB IRB 140 (в. сл. 4.1), са 6 степени слободе и (ii) хуманоидни робот MAPKO (в. сл. 4.2), дизајниран да буде асистивно средство у терапији за децу са сметњама у развоју [3], [101] и [102]. Поред посматраног прототипског система, дијалошки систем обухвата и модул за управљање дијалогом, заснован на моделу фокусног стабла, затим модуле за синтезу и препознавање говора на српском језику и модул за машинску визију [100], [3], [103], [102].



Сл. 4.1 Индустијски робот ABB IRB140 са две мрежне камере за стереовизију AXIS 211 IP. На столу се налази скуп дрвених објеката који представљају просторни контекст који робот дели са корисником приликом интеракције. Слика преузета из [102].



Сл. 4.2 Хуманоидни робот MARKO. Слика преузета из [104].

Домен интеракције за оба роботска система односи се на вежбе fine моторике. Просторни контекст се састоји од 48 дрвених објеката (36 призми и 12 ваљака), који се међусобно разликују по:

- *облику основе* – троугао, круг, квадрат и правоугаоник,
- *висини* – танки (7 mm) и дебели (15 mm),
- *величини основе* – мали и велики
- *боји* – црвена, плава и жута.

Сваки објекат је представљен засебним семантичким ентитетом (нпр. „велики црвени ваљак“ представља један ентитет). Над објектима је могуће вршити 5 акција: показивање објекта и померање објекта у једном од четири могућа просторна правца у равни стола, релативно у односу на корисника – улево, удесно, нагоре и надоле. Комбиновањем објеката и акција може се дефинисати 240 (=48x5) пропозиција које у потпуности описују овај просторни домен.

Да би се проценила валидност предложеног алгорита учења, спроведена је серија од 100 експеримената. У сваком појединачном експерименту насумично је изабрано 9 подскупова комплетног скупа који садржи 240 пропозиција. Ови подскупови су садржали 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% и 90% комплетног скупа, тим редом. Сваки од изабраних подскупова је појединачно саопштен, као улаз, прототипском систему. Током читавог експеримента, прототипском систему је саопштено 900 скупова улазних података.

За квантитативно процењивање учинка прототипског система користе се следећи параметри: прецизност, одзив и Φ -мера [105]. У циљу јасног излагања, уводимо следеће скраћенице:

- T – скуп пропозиција који у потпуности описује домен интеракције,
- S – скуп научених пропозиција који у општем случају може да садржи и пропозиције које не припадају посматраном домену интеракције,
- P – прецизност,
- R – одзив,
- F – Φ -мера.

Прецизност се дефинише следећим изразом:

$$P = \frac{|T \cap S|}{|S|} \times 100. \quad (41)$$

Пресек скупова $T \cap S$ садржи научене пропозиције које су заиста садржане у посматраном домену тако да прецизност представља удео исправно научених пропозиција у скупу научених пропозиција. Одзив се дефинише следећим изразом:

$$R = \frac{|T \cap S|}{|T|} \times 100. \quad (42)$$

Одзив квантитативно изражава меру у којој је алгоритам учења комплетирао знање о домену интеракције, тј. представља однос броја исправно научених пропозиција и укупног броја пропозиција које описују посматрани домен. Вредности прецизности и одзива су, по правилу, обрнуто пропорционалне, због чега се уводи Φ -мера која обједињује ова два параметра. Φ -мера се дефинише следећим изразом:

$$F = \frac{1}{\alpha \frac{1}{P} + (1 - \alpha) \frac{1}{R}}, \quad (43)$$

где параметар α узима вредност из опсега $[0,1]$, у зависности да ли Φ -мера наглашава прецизност или одзив. Овде је изабрана вредност $\alpha=0,5$, која одговара тзв. балансираној Φ -мери која се дефинише следећим изразом:

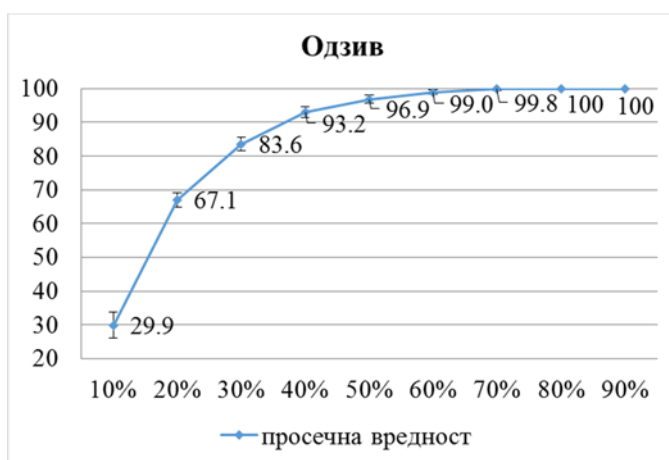
$$F = \frac{2PR}{P + R}. \quad (44)$$

Након спровођења свих 900 појединачних тестирања алгоритма учења, сви улазни скупови су груписани по величини, тј. 900 улазних скупова је груписано у 9 категорија (у првој категорији се налази 100 улазних скупова, од којих сваки садржи по 10% насумично изабраних пропозиција, у другој групи 100 улазних скупова, од којих сваки садржи 20% насумично изабраних пропозиција, итд.). За сваку категорију су израчунате средње вредности и стандардне девијације прецизности,

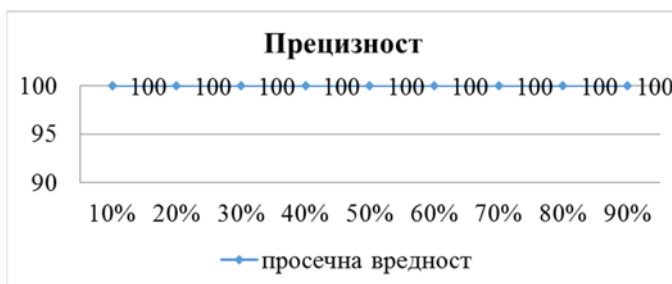
одзива и Ф-мере, тим редом. Резултати су сумирани у табели 4.1 и илустровани на сликама 4.3, 4.4 и 4.5.

Табела 4.1 Резултати експеримента: средње вредности и стандардне девијације прецизности, одзива и Ф-мере. Вредности су израчунаће за 9 различитих величина улазних скупова, на узорку од по 100 вредности.

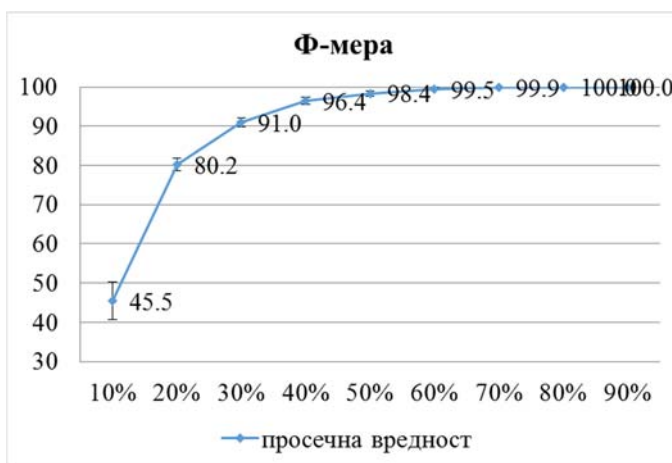
Улазни скуп (поодскупи од T)	Одзив		Прецизност		Ф-мера	
	ср. вредност	ст.дев.	ср. вредност	ст.дев.	ср. вредност	ст.дев.
10%	29,90	7,85	100	0	45,47	9,40
20%	67,06	4,23	100	0	80,21	3,03
30%	83,58	3,88	100	0	91,01	2,33
40%	93,17	3,20	100	0	96,43	1,72
50%	96,85	2,52	100	0	98,39	1,31
60%	98,98	1,37	100	0	99,48	0,70
70%	99,83	0,57	100	0	99,92	0,29
80%	99,98	0,21	100	0	99,99	0,11
90%	100,00	0,00	100	0	100,00	0,00



Сл. 4.3 Средња вредности и стандардна девијација одзива



Сл. 4.4 Средња вредности и стандардна девијација прецизности



Сл. 4.5 Средња вредности и стандардна девијација Ф-мере

Из приказаних резултата се може видети да су перформансе алгоритма учења задовољавајуће. У случају малих улазних скупова, који садрже између 10% и 30% укупног броја пропозиција, резултати алгоритма се знатно разликују. За улазни скуп који садржи 10% комплетног домена, систем је у просеку успео да научи приближно 30% укупног броја пропозиција, док је за дати улазни скуп од 30% домена, систем успео у просеку да научи чак 83,5% домена. Из табеле 4.1 за посматрани домен интеракције се види да кад се алгоритму саопшти улазни скуп који представља 60% комплетног знања о домену, резултат његовог извршавања садржи скоро све пропозиције из домена. За улазне скупове који садрже између 60% и 90% комплетног знања о домену, резултати се не мењају у значајној мери. Интересантно је да у овом експерименту систем није научио ниједну менталну репрезентацију која

не припада домену. Ово се рефлектује у чињеници да прецизност има вредност 100% за све категорије улаза.

Треба напоменути да изабрани домен интеракције не представља специфичан случај. Високи степен униформности који је присутан у овом домену карактеристичан је за широки спектар различитих домена интеракције између човека и сервисних робота, чија је функционалност примарно оријентисана на манипулисање физичким ентитетима у непосредном окружењу.

Моделовање односа између дуготрајне и радне меморије

Не може постојати неки други језик који би био општији и једноставнији, поштеђенији грешака и нејасноћа... примеренији изражавању непроменљивих односа природних ствари (него што је то математика). Она тумачи (све појаве) истим језиком, као да тиме жели да потврди јединство и једноставност устројства Васељене и да учини још уочљивији ненарушиви поредак који влада свим природним узроцима.

- Жан Батист Фурије

Without context words and actions have no meaning at all.

- Gregory Bateson

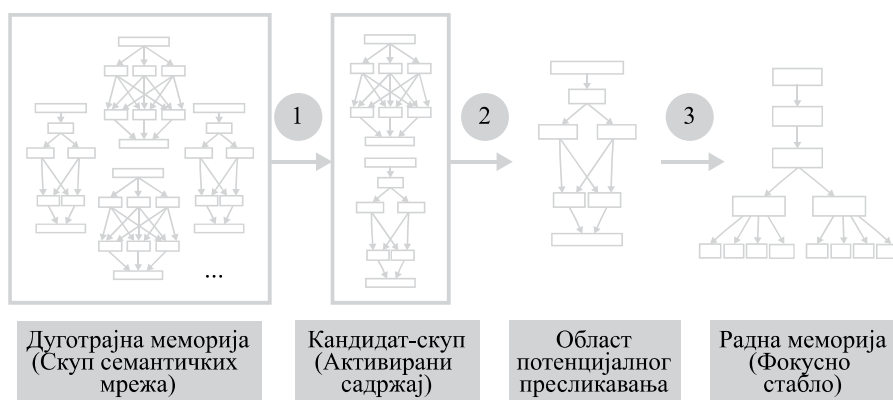
You shall know a word by the company it keeps.

- John Rupert Firth

На крају 3. поглавља је дискутовано како се превођењем семантичких мрежа у фокусно стабло знатно увећава број чворова. За потребе практичних имплементација модела, које укључују велике количине података у дуготрајној меморији, прикладније је да се само подскуп семантичких мрежа који је релевантан за обраду текућег стимуланса, преведе у фокусно стабло, у оквиру ког би се текући стимуланс интерпретирао. Овај захтев је конзистентан са хипотезом да је радна меморија функционално стање (тј. активирани део) дуготрајне меморије, а не засебни анатомски ентитет. Из тог разлога, посебно је наглашено моделовање функционалних веза између дуготрајне и радне меморије,

укључујући и пресликавање активiranог дела семантичких мрежа у фокусно стабло.

У суштини, функционални аспект модела се односи на проблеме адекватног активирања делова дуготрајне меморије и контекстно зависног селектовања информација из активiranог дела, у складу са спољашњим стимулансима, историјом интеракције и тренутним контекстом интеракције. Подскуп активiranог садржаја дуготрајне меморије, чији ниво активације прелази задати праг, проглашава се за тзв. област потенцијалног пресликавања. Затим се садржај ове области представља у форми фокусног стабла и постаје доступан когнитивним механизмима радне меморије за манипулацију и обраду. У оквиру фокусног стабла се обавља контекстно зависно интерпретирање стимуланса. Функционални аспект меморије је шематски приказан блок-дијаграмом на слици 5.1.



Сл. 5.1 Блок-дијаграм који илустрирује функционални аспекти модела меморије. (1) Механизам асоцијативног активирања меморије, који издваја подскуп садржаја дуготрајне меморије, (2) Механизам приоритизације активiranог садржаја, који издваја најрелевантнији подскуп активiranог садржаја, (3) Аутоматско генерисање фокусног стабла и контекстно зависно интерпретирање стимуланса.

У овом поглављу се разматра класификација стимуланса у односу на њихов пропозициони садржај, предлаже се механизам асоцијативног активирања меморије и уводи се мера семантичке сличности између

менталних репрезентација, на основу чега се дефинише механизам приоритизације активираних садржаја.

5.1 Класификација стимуланса

Стимуланс представља улазни податак произвољног сензорског модалитета који током интеракције доспева у меморијски систем. То може бити вербални дијалогски чин генерисан од стране корисника, невербална активност, интерни стимуланс генерисан од стране самог когнитивног агента, итд. У општем случају, стимуланс се формално дефинише као коначни и непразни скуп семантичких конституената:

$$s = \{sc_1, sc_2, \dots, sc_k \mid k \geq 1\}. \quad (45)$$

Дефиниција стимуланса је врло слична дефиницији менталне репрезентације [в. јед. (4)], са том разликом што стимуланс не мора нужно да изражава валидно значење у посматраном домену интеракције. У поглављу 3.1 уведена је класификација менталних репрезентација, заснована на кардиналности садржаних конституената, која се може применити и на стимулансе. Ако стимуланс садржи барем један конституент чија је кардиналност већа од 1, такав стимуланс се сматра комплексним. Међутим, пошто се сваки комплексни стимуланс може декомпоновати на скуп базичних стимуланса, у посматраном моделу, без губитка општости, сви стимуланси се сматрају базичним и идеално познатим.

Интерпретација и класификација одређеног стимуланса су у општем случају контекстно зависне. У овој секцији размотрићемо интерпретацију изолованих стимуланса у датом домену интеракције, само на основу њиховог експлицитног садржаја, не узимајући у обзир историју интеракције (контекстно зависна интерпретација биће касније размотрена у поглављу).

Пресликавање стимуланса на дату семантичку мрежу подразумева одређивање скупа свих путања у мрежи за које важи:

- путања почиње од почетног чвора и усмерена је ка специфичнијим чворовима,
- путања садржи чворове који покривају све семантичке ентитете садржане у стимулансу,
- путања је минимална, односно ако би се из путање уклонио последњи чвор, путања не би испуњавала услов да покрива све ентитете садржане у стимулансу,
- ако путања садржи чвор који је директно повезан са одредишним чвором, тада и одредишни чвор постаје део путање.

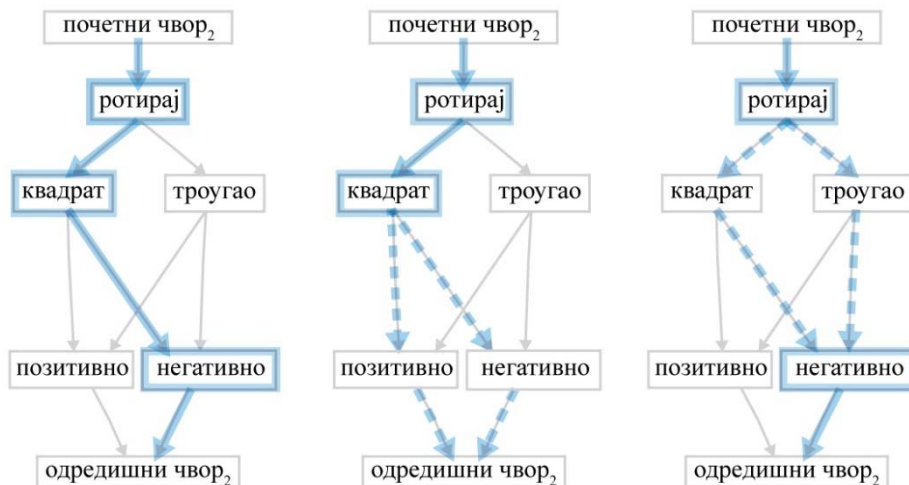
Резултат пресликавања стимуланса може да буде: празни скуп, скуп који садржи једну путању или скуп који садржи више путања. У наставку текста се анализира сваки од ова три случаја и дефинишу се следећи типови стимуланса: (i) комплетни, (ii) некомплетни, (iii) вишесмислени и (iv) семантички некоректни, по узору на [1]. За потребе илустрације и без губитка општости, за референтни домен се усваја семантичка мрежа SM_2 (в. сл. 3.7), која описује домен редуковане слагалице Танграм. Слика 5.2 даје графичку репрезентацију различитих типова стимуланса.

Четири наведена типа стимуланса се дефинишу на следећи начин:

(i) Комплетни стимуланс - Стимуланс се сматра комплетним ако је резултат његовог пресликавања само једна путања, која садржи одредишни чвор (тј. пролази кроз све нивое семантичке мреже). Овакву путању називамо комплетном и она је еквивалентна пропозицији, која, по дефиницији, носи комплетно значење у посматраном домену интеракције. Пример комплетног стимуланса, за претпостављени садржај меморије, дат је следећим скупом ентитета:

$$s = \{\{\text{роширај}\}, \{\text{квадрај}\}, \{\text{негативно}\}\}. \quad (46)$$

Пресликавање овог улаза на семантичку мрежу илустровано је у левом делу слике 5.2.



Сл. 5.2 Графичка репрезентација различитих типова стимуланса. Чворови мреже, који су уоквирени задебљаном линијом, представљају експлицитни садржај стимуланса, док су задебљаним сирелицама означене менталне репрезентације на које је могуће пресликавати стимуланс. Пуне сирелице означавају једнозначно пресликавање, а исирекдане сирелице представљају поштенцијалне улазе за пресликавање некомплетних или вишесмислених стимуланса. Пример некомплетног стимуланса је приказан у левом делу слике, некомплетног стимуланса у средини, а вишесмисленог стимуланса у десном делу слике.

(ii) **Некомплетни стимуланс** – Стимуланс се сматра некомплетним ако је резултат његовог пресликавања само једна путања која не садржи одредишни чвор (тј. путања не пролази кроз све нивое семантичке мреже). Овакву путању називамо некомплетном и она представља менталну репрезентацију са непотпуним значењем у посматраном домену. Пример некомплетног стимуланса је:

$$s = \{\{\overline{\text{квадрат}}\}, \{\overline{\text{ротирај}}\}\}. \quad (47)$$

Као што је илустровано у средишњем делу слике 5.2, овакав улаз се може пресликати на две различите путање у семантичкој мрежи SM_2 , али захтева додатне информације да би се употпунило значење.

(iii) Вишесмислени стимуланс – Стимуланс се сматра вишесмисленим ако је резултат његовог пресликавања скуп од минимално две путање, без обзира да ли су комплетне или некомплетне. Свака од ових путања представља могућу интерпретацију посматраног стимуланса. Пример вишесмисленог стимуланса је:

$$s = \{\{\text{роширај}\}, \{\text{неакиивно}\}\}. \quad (48)$$

Као што је илустровано у левом делу слике 5.2, вишесмислени стимуланс се може прсликати на две различите пропозиције у семантичкој мрежи SM_2 . Додатне информације су потребне да би се разрешила вишесмисленост у значењу.

(iv) Семантички некоректни стимуланс – Стимуланс се сматра семантички некоректним ако је резултат његовог пресликавања празни скуп. То значи да стимуланс или не садржи ниједан познати ентитет, или представља комбинацију познатих ентитета која нема валидно значење у посматраном домену интеракције. Пример семантички некоректног стимуланса је:

$$s = \{\{\text{диск}_1\}, \{\text{роширај}\}, \{\text{иозиивно}\}\}. \quad (49)$$

где је диск_1 познати ентитет који припада мрежи SM_3 (в. сл. 3.7).

5.2 Асоцијативно активирање меморије

За имплементацију асоцијативног активирања меморије користи се инвертовани индекс, структура која се користи за ефикасно проналажење информација у великим колекцијама докумената [105]. У стандардном приступу, за сваки индексни термин који припада речнику термина, на нивоу колекције генерише се листа једнозначних идентификатора докумената у којима се посматрани термин јавља. У контексту предложеног модела меморије, индексни термини су представљени семантичким ентитетима, садржаним у дуготрајној меморији когнитивног агента, а улогу докумената преузимају семантичке мреже. Садржај дуготрајне меморије може да чини потенцијално велики број семантичких мрежа, а семантички ентитети, у општем случају, могу да се јављају

у више различитих семантичких мрежа. Структура инвертованог индекса се формира тако што се за сваки семантички ентитет дефинише уређена листа семантичких мрежа у којима се ентитет јавља (тзв. инвертоване листе). У табели 5.1 је приказана структура инвертованог индекса, формирана за домене интеракције који су представљени на слици 3.7.

Табела 5.1 Структура инвертованог индекса, генерисана за скупи семантичких мрежа приказан на слици 3.7. Семантички ентитети у првој колони су уређени према растућим вредностима. Сваки ентитет има придружену уређену листу једнозначних идентификатора семантичких мрежа у којима се јавља. Инвертоване листе су сортиране у растућем низу.

Структура инвертованог индекса	
Индексни термини (семантички ентитети)	Инвертоване листе (идентификатори мрежа)
<i>диск₁</i>	<i>SM₃</i> .
<i>диск₂</i>	<i>SM₃</i> .
<i>диск₃</i>	<i>SM₃</i> .
<i>квадрати</i>	<i>SM₁ → SM₂</i> .
<i>нагоре</i>	<i>SM₁</i> .
<i>нагоре</i>	<i>SM₁</i> .
<i>негативно</i>	<i>SM₂</i> .
<i>помери</i>	<i>SM₁</i> .
<i>позитивно</i>	<i>SM₂</i> .
<i>ротирај</i>	<i>SM₂</i> .
<i>проутао</i>	<i>SM₁ → SM₂</i> .
<i>удесно</i>	<i>SM₁</i> .
<i>улево</i>	<i>SM₁</i> .
<i>ширај₁</i>	<i>SM₃</i> .
<i>ширај₂</i>	<i>SM₃</i> .
<i>ширај₃</i>	<i>SM₃</i> .

Кад се скуп семантичких мрежа, које представљају садржај дуготрајне меморије, преведе у инвертовани индекс, свака наредна обрада корисничког дијалошког чина (или другог стимуланса), обавља се над инвертованим индексом а не над семантичким мрежама, чиме се постиже ефикасност обраде. Битно је напоменути да су индексни

термини, као и шифре индекса у инвертованим листама, сортирани у растућем низу како би се оптимизовала примена логичких операција (тј. операције уније, пресека и разлике) над инвертованим листама током претраге.

У случају промене саме колекције, нпр. приликом учења нових ентитета или додавања нових семантичких мрежа, потребно је поново урадити пресликавање колекције у инвертоване листе. Треба истаћи да се приликом обраде корисничког упита, превођење семантичких мрежа у инвертовани индекс не одвија у реалном времену, већ унапред и да је инвертовани индекс потребно ажурирати сваки пут када се садржај дуготрајне меморије промени. Ово је аналогно процесу консолидације људске меморије која се у највећој мери обавља током сна [106].

Структура инвертованог индекса је употребљена у предложеном моделу дуготрајне меморије као основа механизма асоцијативне активације. Наиме, за дати стимуланс, из инвертованог индекса се издвајају оне инвертоване листе које су придружене ентитетима садржаним у стимулансу. Над тим инвертованим листама се примењује операција пресека скупова. Овај пресек садржи идентификаторе свих семантичких мрежа (уколико такве постоје) које садрже све ентитете из стимуланса. Идентификоване мреже постају активирани део дуготрајне меморије и конституишу тзв. *кандидат-скуп* (означен као S_{kand}), у оквиру којег је могуће прсликати стимуланс. Важно је нагласити да се процес активације, у предложеном приступу, обавља на нивоу семантичких мрежа. То значи да се приликом обраде стимуланса не активирају само семантички ентитети који су експлицитно садржани у њему, него читаве семантичке мреже. Самим тим, активирају се и други контекстно повезани семантички ентитети и менталне репрезентације, садржани у активираним мрежама. Асоцијативна природа предложеног механизма активације проистиче из предложеног механизма учења (описаног у претходном поглављу), који је претходно успоставио семантичке релације међу ентитетима и менталним репрезентацијама.

Зависно од типа стимуланса, кандидат-скуп, добијен пресеком инвертованих листа, може садржати једну семантичку мрежу, више различитих мрежа или да буде празан. За потребе илустрације,

претпоставићемо да је садржај меморије представљен инвертованим индексом, датим у табели 5.1.

(i) Кандидат-скуп са једном семантичком мрежом – Као пример наводи се следећи стимуланс:

$$s = \{\{\overline{\text{квадрави}}\}, \{\overline{\text{иомери}}\}, \{\overline{\text{улево}}\}\}. \quad (50)$$

Обрадом овог стимуланса над датим инвертованим индексом добијају се три листе, чији пресек одређује резултујућу семантичку мрежу:

- Листа1: *улево* $\rightarrow SM_1$.
- Листа2: *иомери* $\rightarrow SM_1$.
- Листа3: *квадрави* $\rightarrow SM_1 \rightarrow SM_2$.
- Резултат: Листа1 \cap Листа2 \cap Листа3 $\rightarrow SM_1$.

Пресек инвертованих листа издваја семантичку мрежу SM_1 која постаје активирани део дуготрајне меморије. Анализирани стимуланс има комплетно значење у посматраном домену интеракције и једнозначно се пресликава на одговарајућу путању у мрежи.

(ii) Кандидат-скуп са више семантичких мрежа – Пример стимуланса чија обрада, у посматраном контексту, резултује оваквим кандидат-скупом гласи:

$$s = \{\{\overline{\text{квадрави}}\}\}. \quad (51)$$

Семантички ентитет „квадрат“ је познат систему, али је садржан у две семантичке мреже, SM_1 и SM_2 , које се обе активирају. Као што се види на слици 3.7, постоји чак шест различитих пропозиција у оквиру ове две мреже, на које се стимуланс потенцијално може прсликати.

(iii) Празни кандидат-скуп – Ако пресек инвертованих листа врати празан скуп, значи да не постоји ниједна семантичка мрежа у оквиру које је могуће прсликати дати стимуланс. Такав стимуланс је семантички некоректан, што је већ описано у секцији 5.1.

5.3 Приоритизација активираних садржаја

У овој секцији је представљен когнитивно инспирисани механизам за приоритизацију активираних садржаја меморије. Задатак овог механизма је да издвоји подскуп кандидат-скупа који садржи семантичке мреже, чији је активациони потенцијал изнад задатог прага. На концептуалном нивоу, то значи да се издвајају оне семантичке мреже које су најрелевантније за посматрани стимуланс и тренутни контекст интеракције. Издвојене мреже чине *област појенцијалног пресликавања* и преводе се у форму фокусног стабла, чиме постају доступне за когнитивну обраду и манипулацију од стране механизма радне меморије.

У наставку секције је објашњено израчунавање активационих потенцијала семантичких мрежа и адаптивно одређивање вредности прага, након чега је предложен механизам за утврђивање промена у контексту интеракције.

5.3.1 Мера семантичке сличности менталних репрезентација

Мера семантичке сличности између комплексних менталних репрезентација (тј. семантичких мрежа које их представљају) уводи се са циљем да би се идентификовале групе асоцијативно повезаних менталних репрезентација. Предложена мера сличности је заснована на косинусној сличности [107]. Да би се применила мера косинусне сличности, свака ментална репрезентација је представљена одговарајућим вектором, чији се елементи израчунавају на основу *релативне учесталости* појављивања семантичких ентитета у оквиру посматране комплексне менталне репрезентације.

Апсолутна учесталост појављивања ентитета у комплексној менталној репрезентацији дефинише се као број различитих пропозиција које се могу извести из те комплексне репрезентације и садрже посматрани ентитет. Формална дефиниција за апсолутну учесталост ентитета e у менталној репрезентацији m_L , дата је следећим изразом:

$$f_{abs}(e, m_L) = \prod |sc_i|, \text{ где важи } (sc_i \in m_L \wedge e \notin sc_i). \quad (52)$$

Релативна учесталост појављивања ентитета у комплексној менталној репрезентацији представља нормализовану вредност апсолутне учесталости и дефинише се као количник апсолутне учесталости ентитета и укупног броја пропозиција које се могу извести из дате комплексне менталне репрезентације. Лако се показује да је релативна учесталост ентитета обрнуто пропорционална кардиналности семантичког конституента који га садржи. Формална дефиниција за релативну учесталост ентитета e у менталној репрезентацији m_L , дата је следећим изразом:

$$f_{rel}(e, m_L) = \frac{1}{|sc|}, \text{ где важи } (e \in sc, sc \in m_L). \quad (53)$$

За потребе илустрације, наводимо менталну репрезентацију $m_{L1} \in M_L$:

$$m_{L1} = \left\{ \left\{ \begin{array}{l} \text{квадрати} \\ \text{широкиао} \end{array} \right\}, \{ \text{помери} \}, \left\{ \begin{array}{l} \text{нагоре,} \\ \text{надоле,} \\ \text{улево,} \\ \text{угесно} \end{array} \right\} \right\}, \quad (54)$$

Апсолутна учесталост семантичког ентитета „квадрат“ у менталној репрезентацији m_{L1} има вредност 4, док је релативна учесталост једнака 1/2.

Апсолутна учесталост ентитета у менталној репрезентацији квантификује значај неког семантичког ентитета за описивање одређене менталне репрезентације. Семантички ентитет који је заступљен у више пропозиција, изведених из неке менталне репрезентације, сматра се општијим. Релативна учесталост појављивања ентитета одговара апсолутној учесталости нормализованој тако да припада опсегу [0,1], што омогућава поређење учесталости истог ентитета у различитим комплексним менталним репрезентацијама.

Вектор који представља менталну репрезентацију m_L формира се на следећи начин: нека је SE скуп свих семантичких ентитета садржаних у дуготрајној меморији, при чему је укупни број ентитета једнак $|SE|=N$. Претпоставимо да су семантички ентитети из SE уређени према неком

задатом критеријму, у низ: e_1, e_2, \dots, e_N . Ментална репрезентација m_L се представља N -димензионалним вектором \mathbf{v} , дефинисаним на следећи начин:

$$\mathbf{v} = [v_1, v_2, \dots, v_i, \dots, v_N], \text{ где важи} \\ (\forall 1 \leq i \leq N), v_i = \begin{cases} f_{rel}(e_i, m_L), & (\exists sc \in m_L)(e_i \in sc) \\ 0, & (\forall sc \in m_L)(e_i \notin sc) \end{cases} \quad (55)$$

У вектору \mathbf{v} , i -та компонента вектора одговара i -том елементу у уређеном низу ентитета. Вредност компоненте вектора је једнака релативној учесталости одговарајућег ентитета у посматраној менталној репрезентацији, ако је ентитет садржан у њој, или 0, у супротном. За потребе илустрације, у табели 5.2 су приказани одговарајући вектори за све комплексне менталне репрезентације, представљене семантичким мрежама на слици 3.7.

Табела 5.2 Векторска репрезентација семантичких мрежа представљених на слици 3.7. Компоненте вектора одговарају релативним учесталостима појављивања ентитета у оквиру семантичких мрежа.

Мера семантичке сличности			
Семантички ентитети	Векторска репрезентација		
	$v_1(SM_1)$	$v_2(SM_2)$	$v_3(SM_3)$
<i>диск1</i>	0	0	0.333
<i>диск2</i>	0	0	0.333
<i>диск3</i>	0	0	0.333
<i>квадрат</i>	0.5	0.5	0
<i>нагоре</i>	0.25	0	0
<i>нагоре</i>	0.25	0	0
<i>негативно</i>	0	0.5	0
<i>помери</i>	1	0	0
<i>позитивно</i>	0	0.5	0
<i>ротирај</i>	0	1	0
<i>пројекто</i>	0.5	0.5	0
<i>удесно</i>	0.25	0	0
<i>улево</i>	0.25	0	0
<i>шија1</i>	0	0	0.333
<i>шија2</i>	0	0	0.333
<i>шија3</i>	0	0	0.333

Мера (косинусне) семантичке сличности између две менталне репрезентације дефинише се као количник скаларног производа вектора који их представљају и производа интензитета тих вектора [107]:

$$\text{sim}(m_{Li}, m_{Lj}) = \cos(\varphi_{i,j}) = \frac{\mathbf{v}_i \mathbf{v}_j}{|\mathbf{v}_i| |\mathbf{v}_j|} = \frac{\sum_{k=1}^n v_{i_k} v_{j_k}}{\sqrt{\sum_{k=1}^n v_{i_k}^2} \sqrt{\sum_{k=1}^n v_{j_k}^2}}. \quad (56)$$

Уведена мера се може геометријски интерпретирати као косинус угла између два N -димензиона вектора. Компоненте вектора који представљају семантичке ентитете, по дефиницији, не могу бити негативне вредности, па је вредност косинусне сличности увек у опсегу $[0,1]$, при чему, вредност 0 означава да две менталне репрезентације немају ниједан заједнички ентитет (тј. представљене су ортогоналним векторима), а вредност 1 означава да се менталне репрезентације апсолутно поклапају (тј. представљене су паралелним векторима).

Табела 5.3 Мера семантичке сличности израчунава за сваки пар семантичких мрежа представљених на слици 3.7.

Семантичке мреже	Семантичка сличност
$SM_1 \sim SM_2$	0.2673
$SM_1 \sim SM_3$	0
$SM_2 \sim SM_3$	0

Мера семантичке сличности је израчуната за сваки пар генерисаних семантичких мрежа и приказана је у табели 5.3. Из табеле се може видети да постоји извесни ниво семантичке сличности између семантичких мрежа SM_1 и SM_2 , док мрежа SM_3 нема никакве сличности са прве две мреже. Овај резултат је усаглашен са чињеницом да семантичке мреже SM_1 и SM_2 припадају истом домену интеракције (тј. редукованој слагалици Танграм), док SM_3 припада другом домену (тј. проблему ханојских кула). Ипак, треба нагласити да чак и семантичке мреже које припадају различитим доменима интеракције, могу у општем случају да имају не-нулту вредност семантичке сличности. Разлог за то је што различити домени могу да деле скуп заједничких семантичких ентитета,

што омогућава примену мере сличности у предложеном механизму приоритизације активiranог садржаја меморије, који има извесну аналогију са механизмом ширеће активације. Приликом активирања одређене менталне репрезентације, такав механизам побуђује и скуп семантички сличних менталних репрезентација, што ће бити размотрено у следећој секцији.

5.3.2 *Активациони потенцијал семантичке мреже*

Активациони потенцијал семантичке мреже се израчунава за сваку активiranу мрежу из кандидат-скупа на основу два критеријума: њене релевантности током историје интеракције и њене релевантности за тренутни контекст интеракције. Ово је формално исказано следећим изразом:

$$P_A(SM_i) = k_A P_I(SM_i) + (1 - k_A) P_K(SM_i), \quad (SM_i \in S_{kand}). \quad (57)$$

где су: $P_A(SM_i)$ активациони потенцијал мреже SM_i , $P_I(SM_i)$ компонента потенцијала мреже SM_i која се односи на релевантност мреже током историје интеракције, а $P_K(SM_i)$ компонента потенцијал мреже SM_i која се односи на релевантност мреже за тренутни контекст интеракције. Параметар $k_A \in [0,1]$ одређује узајамни однос између ове две компоненте активационог потенцијала мреже. У наставку текста се детаљније разматрају компоненте активационог потенцијала.

5.3.2.1 *Релевантност мреже током историје интеракције*

Да би се квантификовала релевантност мреже током историје интеракције, коришћени су увиди из когнитивне психологије, у тзв. меморијску пристрасност, врсту когнитивне пристрасности која поспешује или слаби присећање, тј. могућност извлачења одређене информације из меморије [22], [37]. Постоји више различитих типова меморијске пристрасности, од којих се овде разматрају два најрелевантнија за предложени механизам приоритизације: (i) учесталост активације меморијске јединице и (ii) време последњег приступа меморијској јединици. Стога, коначна вредност компоненте $P_I(SM_i)$ рачуна се као тежинска сума два поменутог фактора:

$$P_l(SM_i) = k_l P_f(SM_i) + (1 - k_l) P_r(SM_i), \quad (58)$$

где су: $P_f(SM_i)$ потенцијал мреже SM_i у односу на учесталост њене активације током историје интеракције и $P_r(SM_i)$ потенцијал у односу на тренутак последњег приступа. Параметар $k_l \in [0,1]$ одређује узајамни однос између ове две компоненте.

(i) *Критеријум заснован на учесталости активације*

Овај критеријум је заснован на претпоставци да је вероватноћа да ће семантичка мрежа бити активирана пропорционална учесталости њеног претходног активирања током историје интеракције. Одговарајућа компонента потенцијала дефинише се следећом формулом:

$$P_f(SM_i) = \frac{br(SM_i)}{\sum_{SM_k \in M_L} br(SM_k)}, \quad (59)$$

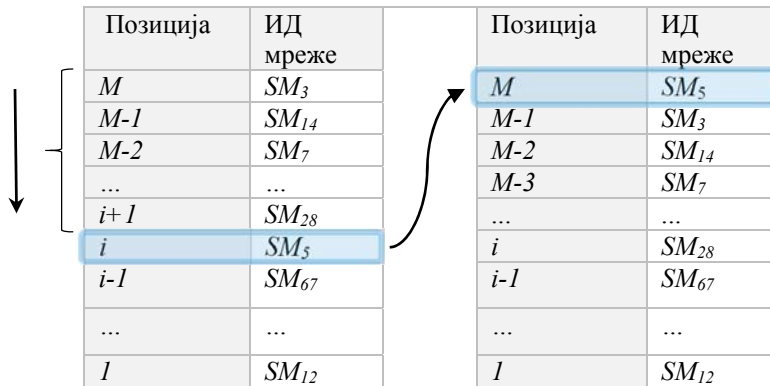
где су: $br(SM_i)$ број претходних активирања мреже SM_i , M_L скуп комплексних менталних репрезентација, које чине садржај дуготрајне меморије, а сума у делиоцу представља укупан број активирања семантичких мрежа.

Вредност потенцијала $P_f(SM_i)$ увек је у опсегу $[0,1]$, а сума потенцијала (овог типа) свих мрежа у дуготрајној меморији је увек једнака 1. На почетку интеракције, сви потенцијали се иницијализују на вредност $1/|M_L|$. Приликом сваке активације, вредности потенцијала се ажурирају.

(ii) *Критеријум заснован на времену последњеј присијуа*

Овај критеријум је заснован на претпоставци да је вероватноћа активације семантичке мреже обрнуто пропорционална времену које је протекло од њене последње активације. Потенцијал мреже P_r , заснован на времену последњег приступа, израчунава се на следећи начин: креира се M -димензионални линеарни бафер, где је M једнако укупном броју семантичких мрежа у дуготрајној меморији, тј. $M=|M_L|$. Овај бафер садржи једнозначне идентификаторе семантичких мрежа, уређене тако да се на првој позицији бафера налази идентификатор мреже која најдуже није активирана, а на последњој позицији идентификатор мреже која је

последња активирана. Треба истаћи да бафер не бележи апсолутно време активирања мрежа, него да редослед у баферу одражава само релативни временски след активација мрежа током историје интеракције. Почетни редослед идентификатора у баферу је произвољан, а током интеракције се њихов редослед индекса ажурира приликом сваке активације тако што се индекс активираних мрежа помера на последње место у баферу, а сви елементи бафера који су се налазили на вишим позицијама од тренутно активираних мрежа, померају се за једно место ка почетку бафера. За потребе илустрације, претпоставимо, без губитка општости, да је тренутни садржај бафера приказан у левом делу слике 5.3.



Сл. 5.3 Илустрација ажурирања бафера приликом активирања семантичке мреже. Идентификатор на последњем месту бафера односи се на најскорије активiranу мрежу.

Потенцијал $P_r(SM_i)$ посматране мреже SM_i дефинише се као количник тренутне редне позиције идентификатора мреже SM_i у баферу и укупног броја елемената у баферу. Вредност потенцијала $P_r(SM_i)$ је увек у опсегу $[0,1]$, а сума свих потенцијала овог типа, свих мрежа у дуготрајној меморији, увек је једнака 1.

5.3.2.2 Релевантност мреже за тренутни контекст интеракције

Релевантност семантичке мреже за тренутни контекст интеракције процењује се на основу уведене мере семантичке сличности између менталних репрезентација. Појам тренутног контекста интеракције овде је представљен као скуп семантичких мрежа које тренутно чине садржај области потенцијалног пресликавања, тј. скуп мрежа које су биле

селектоване приликом обраде актуелног стимуланса, при чему се једна од њих на коју је преликан актуелни стимуланс, налази у тренутном фокусу пажње. Као што ће бити приказано у наставку поглавља, управо ће садржај ове области бити представљен у форми фокусног стабла, у оквиру којег ће се преликати стимуланс.

Када се обрадом новог стимуланса одреди нови кандидат-скуп, могуће је проценити ниво његовог контекстуалног преклапања са тренутним контекстом интеракције (тј. са облашћу потенцијалног преликавања). Ниво преклапања је концептуално повезан са компонентом активационог потенцијала која се односи на релевантност мреже за тренутни контекст. Релевантност мреже SM_i из кандидат-скупа, за текући контекст интеракције, дефинише се на следећи начин:

$$P_K(SM_i) = \begin{cases} \frac{\sum_{SM_j \in OPP} w_j \text{sim}(SM_i, SM_j)}{\sum_{SM_p \in S_{kand}} \sum_{SM_j \in OPP} w_j \text{sim}(SM_p, SM_j)}, & SM_i \in S_{kand} \\ 0, & SM_i \notin S_{kand} \end{cases}, \quad (60)$$

где су: w_j тежински коефицијенти, $\text{sim}(SM_i, SM_j)$ семантичка сличност између мрежа SM_i и SM_j , S_{kand} кандидат-скуп, а OPP област потенцијалног преликавања. Из ове дефиниције следи да је вредност потенцијала $P_K(SM_i)$ увек у опсегу $[0,1]$, а сума свих потенцијала овог типа, свих мрежа у дуготрајној меморији, увек је једнака 1. Иницијалне вредности овог потенцијала се на почетку интеракције постављају на вредност $1/|M_L|$, где је $|M_L|$ број семантичких мрежа (тј. комплексних менталних репрезентација) у дуготрајној меморији.

Одређивање тежинских коефицијената w_{ij} засновано је на претпоставци, да тежински коефицијент, који се примењује кад се одређује сличност дате мреже SM_i са семантичком мрежом из области потенцијалног преликавања у оквиру које је преликан претходни стимуланс⁵, треба да буде x пута већи од осталих тежинских коефицијената, који имају међусобно једнаке вредности, где је $x > 1$.

⁵ У наставку ћемо користити формулацију да је ова мрежа у фокусу пажње.

Такође, збир свих тежинских коефицијената треба да буде једнак 1, па се ови коефицијенти дефинишу на следећи начин:

$$w_j = \begin{cases} \frac{x}{x + |OPP| - 1}, & SM_j \text{ у фокусу пажње} \\ \frac{1}{x + |OPP| - 1}, & SM_j \text{ није у фокусу пажње} \end{cases}, \quad (61)$$

где је $|OPP|$ број менталних репрезентација у области потенцијалног мапирања.

5.3.2.3 Адаптивно одређивање вредности прага активације

Основни циљ механизма за приоритизацију активираних семантичких мрежа је идентификација оних мрежа чији активациони потенцијали прелазе задату вредност прага. У претходним секцијама смо размотрили израчунавање активационих потенцијала семантичких мрежа, а у овој секцији представимо адаптивно одређивање вредности прага активације [108].

Пошто се одреде активациони потенцијали свих мрежа из кандидатског скупа [према јед. (57)], израчунава се аритметичка средина ових потенцијала и проглашава се за текућу вредност прага α_1 . Све мреже из кандидатског скупа се деле у два дисјунктна подскупа. Један подскуп садржи мреже чији су активациони потенцијали мањи од текуће вредности прага, док други садржи мреже чији су потенцијали једнаки текућој вредности прага или већи од ње. Након тога се израчунава аритметичка средина активационих потенцијала за сваки од ових подскупова, m_{ispod} и m_{iznad} , а за нову текућу вредност прага се проглашава следећа вредност:

$$\alpha_2 = \frac{m_{ispod} + m_{iznad}}{2}. \quad (62)$$

Овај поступак се итеративно понавља, при чему текућа вредност прага конвергира ка својој крајњој вредности. У практичним применама, гранична вредност се не одређује аналитички, већ итеративно. Поступак се понавља све док апсолутна вредност разлике две сукцесивне текуће вредности прага не постане мања од унапред задате, ниске, позитивне вредности ϵ , што се формално може записати као:

$$|\alpha_k - \alpha_{k-1}| < \varepsilon. \quad (63)$$

Вредност прага се тад поставља на α_k .

5.3.3 Процена промене контекста интеракције

У претходним секцијама смо дефинисали поступак за израчунавање активационих потенцијала и прага активације. Предложене формуле укључују низ параметара који омогућавају наглашавање појединих компоненти активационих потенцијала. Вредности ових параметара нису разматране, јер оне зависе од спецификације конкретног когнитивног система у оквиру којег се имплементира предложени меморијски модел. У наставку излагања се подразумева да су вредности коефицијената познате.

За управљање интеракцијом од значаја је да систем може да квантификује меру промене контекста интеракције изазване обрадом текућег стимуланса. Промена контекста се своди на поређење скупова, који представљају области потенцијалног пресликавања пре обраде актуелног стимуланса (OPP_1) и после ње (OPP_2). Сличност ових скупова се дефинише на основу Жакаровог⁶ индекса сличности [105], који је дефинисан као количник кардиналности пресека скупова који се пореде и кардиналности њихове уније:

$$J(OPP_1, OPP_2) = \frac{|OPP_1 \cap OPP_2|}{|OPP_1 \cup OPP_2|}. \quad (64)$$

Вредност Жакаровог индекса је увек у опсегу [0,1]. Ако су два скупа идентична, одговарајућа вредност индекса је 1, што у контексту модела значи да се контекст интеракције није променио. Ако два скупа немају ништа заједничко, одговарајућа вредност индекса је 0, што значи да се целокупни контекст интеракције променио.

⁶ Paul Jaccard (1868-1944) – професор ботанике и физиологије биљака.

5.4 Концептуализација радне меморије

Механизми активације и приоритизације, изложени у секцијама 5.2 и 5.3, издвајају подскуп семантичких мрежа (тј. комплексних менталних репрезентација) дуготрајне меморије који је од значаја за интерпретирање стимуланса. Овај скуп дефинише област потенцијалног пресликавања. У приступу разматраном у овом поглављу, само се семантичке мреже, садржане у области потенцијалног пресликавања, преводе у фокусно стабло, које представља радну меморију и у оквиру којег се врши интерпретација текућег стимуланса. На овај начин се успоставља веза између предложеног модела дуготрајне меморије, са једне стране, и модела радне меморије заснованог на фокусном стаблу, са друге стране. Ово је у складу са уведеном претпоставком да дуготрајна и радна меморија нису засебни анатомски ентитети, него да се механизми радне меморије извршавају директно над активираним садржајима дуготрајне меморије.

За разлику од претпоставки изнетих и коришћених у секцијама 5.1 и 5.2, интерпретација стимуланса у фокусном стаблу је контекстно зависна, тј. текући стимуланс се интерпретира не само у односу на свој пропозициони садржај, него и у односу на тренутни контекст интеракције и историју интеракције, који су изражени кроз топологију и садржај фокусног стабла и фокус пажње. Поступак за контекстно зависно интерпретирање стимуланса у оквиру фокусног стабла детаљно је представљен у [1], [2], [3].

Закључак

Every science begins as philosophy and ends as art.

- *Will Durant*

Талентован човек остварује циљ који нико други не може да оствари. Геније остварује циљ који нико други не може да види.

- *Артур Шопенхауер*

У дисертацији је представљен нови рачунарски модел дуготрајне меморије намењен за примене у техничким когнитивним агентима са циљем да допринесе симулацији интелигентног понашања робота. Предложени модел је инспирисан изабраним когнитивним механизмима људског меморијског система, који укључују интеграцију менталних репрезентација, семантичку категоризацију, асоцијативно учење и контекстно зависно селектовање информација. У фокус истраживања су постављени проблеми моделовања когнитивних механизма који су релевантни за интеракцију између човека и робота на природном језику и који проширују скуп функционалности претходно развијеног дијалогског система, заснованог на моделу фокусног стабла.

Ниво апстракције у спецификацији модела довољан је да омогући примену модела у широком спектру просторних, униформних домена, који су карактеристични за интеракцију између човека и робота и у којима корисник задаје говорне инструкције роботу у вези са манипулацијом физичких ентитета у радном окружењу робота. Са друге стране, ниво детаља у спецификацији довољан је за рачунарску имплементацију

модела. На методолошком нивоу, предложени модел је симболички, што га чини следљивим, а резултати извршавања предложених алгоритама могу се објаснити у оквиру претпостављених концепата на којима се заснива модел, што значи да модел има експланаторну моћ.

У основи модела се налази симболички приступ за аутоматско моделовање домена интеракције између човека и робота. У процесима семантичке категоризације и асоцијативног учења, систем на основу почетног скупа пропозиција класификује пропозиције у односу на домене којима припадају и изводи недостајуће знање о доменима. Прототипски систем заснован на предложеном моделу и интегрисан у дијалогски систем експериментално је валидиран. Резултати су показали високу успешност предложеног механизма учења, као и његову скалабилност. У низу експеримената, прототипском систему су саопштавани улазни скупови различитих обима, у опсегу од 10% до 90% укупног знања о домену. Показано је да систем успева, са улазним скупом од приближно 50% свих пропозиција које описују домен, да комплетира готово 100% знања о домену.

Адекватна пажња је посвећена и потребама практичних имплементација модела које укључују велике количине података у дуготрајној меморији. Посебно је наглашено моделовање функционалних веза између дуготрајне и радне меморије. Релевантни функционални аспект предложеног модела односи се на проблеме адекватног активирања делова дуготрајне меморије и контекстно зависног селектовања информација из активiranог дела, у складу са спољашњим стимулансима, историјом интеракције и тренутним контекстом интеракције.

Научни доприноси ове тезе представљају основу за даље истраживање теоријских и практичних аспеката у области моделовања меморије и интеракције између човека и робота. Будући правци истраживања укључују (али се не ограничавају на) истраживачке проблеме одређивања невалидних пропозиција генерисаних у случајевима нетипичних и неуниформних домена, учење из интеракције, и моделовање епизодне меморије.

Библиографија

- [1] M. Gnjatović, M. Janev and V. Delić, "Focus tree: modeling attentional information in task-oriented human-machine interaction," *Applied Intelligence*, vol. 37, no. 3, p. 305–320, 2012.
- [2] M. Gnjatović and V. Delić, "Cognitively-inspired representational approach to meaning in machine dialogue," *Knowledge-Based Systems*, vol. 71, pp. 25-33, 2014.
- [3] M. Gnjatović, "Therapist-Centered Design of a Robot's Dialogue Behavior," *Cognitive Computation*, vol. 6, no. 4, pp. 775-788, 2014.
- [4] M. Gnjatović and V. Delić, "Electrophysiologically-inspired evaluation of dialogue act complexity," in *Proc. of the 4th IEEE International Conference on Cognitive InfoCommunications, CogInfoCom*, Budapest, Hungary, 2013.
- [5] M. Gnjatović, D. Mišković, S. Savić, B. Borovac, N. Maček and B. Trenkić, "A Novel Modular Architecture for Conversational Robotic Agents," in *4th International Conference on Electrical, Electronic and Computing Engineering, IcETRAN 2017*, Kladovo, Serbia, 2017.
- [6] "Америчка психолошка асоцијација, АПА," [Online]. Available: <http://www.apa.org/research/action/glossary.aspx?tab=13>. [Accessed 9 September 2016].
- [7] "Стенфордова енциклопедија филозофије," [Online]. Available: <http://plato.stanford.edu/entries/memory/>. [Accessed 9 September 2016].

-
- [8] D. Matsumoto, *The Cambridge Dictionary of Psychology*, D. Matsumoto, Ed., Cambridge, UK: Cambridge University Press, 2009.
- [9] E. Tulving, "Episodic and semantic memory," in *Organization of memory*, 2 ed., E. Tulving and W. Donaldson, Eds., New York, Academic Press, 1972, pp. 381- 403.
- [10] D. Filipović Đurđević and S. Zdravković, *Uvod u kognitivne neuronauke*, Zrenjanin: Gradska narodna biblioteka Zrenjanin, 2013.
- [11] R. J. Sternberg and K. Sternberg, *Cognitive Psychology*, 6 ed., Wadsworth: Cengage Learning, 2012.
- [12] E. R. Kandel, I. Kupferman and S. Iverson, "Learning and Memory," in *Principles of Neural Science*, E. R. Kandel, J. H. Schwartz and T. M. Jessell, Eds., New York, McGraw-Hill, 2000, pp. 1227- 1246.
- [13] Z. Pylyshyn, "What the mind's eye tells the mind's brain: A critique of mental imagery," *Psychological Bulletin*, vol. 80, no. 1, pp. 1- 24, 1973.
- [14] A. Kostić, *Kognitivna psihologija*, Beograd, Srbija: Zavod za udžbenike, 2010.
- [15] J. R. Anderson and G. H. Bower, *Human associative memory*, New York: Wiley, 1973.
- [16] J. R. Anderson, *Cognitive Psychology and its Implications*, 2 ed., New York: Freeman, 1985.
- [17] N. Chomsky, *Syntactic Structures*, Hague: Mouton, 1957.
- [18] J. S. Sachs, "Recognition memory for syntactic and semantic aspects of connected discourse," *Perception and Psychophysics*, vol. 2, no. 9, pp. 437- 442, 1967.
- [19] W. Kintsch, *The Representation of Meaning in Memory*, Hillsdale: NJ: Erlbaum, 1974.

- [20] R. Ratcliff and G. McKoon, "Priming in item recognition: Evidence for the propositional structure of sentences," *Journal of Verbal Learning and Verbal Behavior*, vol. 17, no. 4, pp. 403- 418, 1978.
- [21] A. G. Huth, S. Nishimoto, A. T. Vu and J. L. Gallant, "A Continuous Semantic Space Describes the Representation of Thousands of Object and Action Categories across the Human Brain," *Neuron*, vol. 76, no. 6, pp. 1210- 1224, 2012.
- [22] R. C. Atkinson and R. M. Shiffrin, "Human memory: A proposed system and its control processes," in *The psychology of learning and motivation: Advances in research and theory*, vol. 2, K. W. Spence and J. T. Spence, Eds., New York, Academic Press, 1968, pp. 89-195.
- [23] A. D. Baddeley and G. J. Hitch, "Working memory," in *The psychology of learning and motivation: Advances in research and theory*, vol. 8, G. H. Bower, Ed., New York, Academic Press, 1974, pp. 47- 89.
- [24] N. Cowan, "Evolving conceptions of memory storage selective attention, and their mutual constraints within the information processing system," *Psychol. Bull.*, vol. 104, no. 2, pp. 163-191, 1988.
- [25] K. Oberauer, "Access to information in working memory: Exploring the focus of attention," *J. Exp. Psychol. Learn. Mem. Cogn.*, vol. 28, no. 3, pp. 411- 421, 2002.
- [26] P. Shah and A. Miyake, "Models of working memory: An introduction," in *Models of working memory: Mechanisms of active maintenance and executive control*, A. Miyake A and P. Shah, Eds., New York, Cambridge University Press, 1999, pp. 1- 26.
- [27] M. D'Esposito, "From cognitive to neural models of working memory," *Phil. Trans. R. Soc. B*, vol. 362, no. 1481, pp. 761- 772, 2007.
- [28] G. A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information," *Psychological Review*, vol. 101, no. 2, pp. 343- 352, 1956.

- [29] R. W. Engle and M. J. Kane, "Executive attention, working memory capacity, and a two-factor theory of cognitive control," in *The psychology of learning and motivation*, vol. 44, B. Ross, Ed., New York, Elsevier, 2004, pp. 145- 199.
- [30] M. J. Kane, A. R. A. Conway, D. Z. Hambrick and W. Engle, "Variation in working memory capacity as variation in executive attention and control," in *Variation in working memory*, A. R. A. Conway, C. Jarrold, M. J. Kane and A. Miyake, Eds., Oxford, Oxford University Press, 2007, pp. 21- 48.
- [31] J. A. Mogle, B. J. Lovett, R. S. Stawski and M. J. Sliwinski, "What's so special about working memory? An examination of the relationship among working memory, secondary memory, and fluid intelligence," *Psychol. Sci.* , vol. 19, no. 11, pp. 1071- 1077, 2008.
- [32] N. Unsworth and R. W. Engle, "The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory," *Psychol. Rev.* , vol. 114, no. 1, pp. 104- 132, 2007.
- [33] N. Unsworth and G. J. Spillers, "Working memory capacity: Attention control secondary memory, or both? A direct test of the dual-component model," *J. Mem. Lang.* , vol. 62, no. 4, pp. 392- 406, 2010.
- [34] J. A. Brown, "Some tests of the decay theory of immediate memory," *Quarterly Journal of Experimental Psychology*, vol. 10, no. 1, pp. 12- 21, 1958.
- [35] L. R. Peterson and M. J. Peterson, "Short-term retention of individual verbal terms," *Journal of Experimental Psychology*, vol. 58, no. 3, pp. 193- 198, 1959.
- [36] B. B. J. Murdock, "The serial position effect of free recall," *Journal of Experimental Psychology*, vol. 64, no. 5, pp. 482- 488, 1962.

- [37] M. Glanzer, "Storage mechanisms in recall," in *The Psychology of Learning and Motivation*, vol. 5, U. G. H. Bower and J. T. Spence, Eds., New York, Academic Press, 1972, pp. 129- 193.
- [38] M. R. Quillian, "Semantic memory," in *Semantic Information Processing*, M. U. Minsky, Ed., Cambridge, MA: MIT Press, 1968.
- [39] M. R. Quillian, "The teachable language comprehender: A simulation program and theory of language," *Communications of the ACM*, vol. 12, no. 8, pp. 459- 476, 1969.
- [40] A. J. Wilkins, "Conjoint frequency, category size and categorization time," *Journal of Verbal Learning and Verbal Behavior*, vol. 10, no. 4, pp. 382- 385, 1971.
- [41] C. Conrad, "Cognitive economy in semantic memory," *Journal of experimental Psychology*, vol. 92, no. 2, pp. 149- 154, 1972.
- [42] L. J. Rips, E. J. Shoben and E. E. Smith, "Semantic distance and the verification of semantic relations," *Journal of Verbal Learning and Verbal Behavior*, vol. 12, no. 1, pp. 1- 20, 1973.
- [43] A. M. Collins and E. F. Loftus, "A spreading activation theory of semantic memory," *Psychological Review*, vol. 82, no. 6, pp. 47- 428, 1975.
- [44] E. H. Rosch, "On the internal structure of the perceptual and semantic categories," in *Cognitive Development and the Acquisition of Language*, U. T. E. Moore , Ed., New York, Academic Press, 1973, pp. 111- 144.
- [45] E. H. Rosch, "Cognitive representation of semantic categories," *Journal of Experimental Psychology: General*, vol. 104, no. 3, pp. 192-233, 1975.
- [46] E. E. Smith, E. J. Shoben and L. J. Rips, "Structure and process in semantic memory: A featural model for semantic decision," *Psychological Review*, vol. 81, no. 3, pp. 214-241, 1974.

- [47] K. McRae and M. N. Jones, "Semantic memory," in *The Oxford Handbook of Cognitive Psychology*, D. Reisberg, Ed., Oxford University Press, 2013.
- [48] J. R. Anderson, *The Architecture of Cognition*, Cambridge: Harvard University Press, 1983.
- [49] D. Vernon, C. von Hofsten and L. Fadiga, *A Roadmap for Cognitive Development in Humanoid Robots*, Berlin: Springer, 2010.
- [50] P. Langley, J. E. Laird and S. Rogers, "Cognitive architectures: Research issues and challenges," *Cognitive Systems Research*, vol. 10, pp. 141-160, 2009.
- [51] J. R. Anderson and C. Lebiere, *The atomic component of thought*, Mahwah, NJ: Lawrence Erlbaum Associates, 1998.
- [52] J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere and Y. Qin, "An integrated theory of the mind," *Psychological Review*, vol. 111, no. 4, pp. 1036-1060, 2004.
- [53] M. Rutledge-Taylor, C. Lebiere, R. Thomson, J. Staszewski and J. R. Anderson, "A Comparison of Rule-Based versus Exemplar-Based Categorization Using the ACT-R Architecture," in *Proceedings of the 21st Annual Behavior Representation in Modeling and Simulation Conference*, Florida, USA, 2012.
- [54] R. Thomson, A. Pyke, J. G. Trafton and L. M. Hiatt, "An Account of Associative Learning in Memory Recall," in *Proceedings of the 37th Annual Conference of the Cognitive Science Society*, Austin, Texas, 2015.
- [55] G. Trafton, L. Hiatt, A. Harrison, F. Tamborello, S. Khemlani and A. Schultz, "ACT-R/E: An Embodied Cognitive Architecture for Human-Robot Interaction," *Journal of Human-Robot Interaction*, pp. 30-55, 2013.

- [56] D. Choi, M. Kaufman, P. Langley, N. Nejati and D. Shapiro, "An architecture for persistent reactive behavior," in *Third International Joint Conference on Autonomous Agents and Multi-Agent Systems*, New York, 2004.
- [57] P. Langley, "Cognitive architectures and the construction of intelligent agents," in *Proceedings of the AAAI 2004 Workshop on Intelligent Agent Architectures*, Stanford, 2004.
- [58] F. Pratama, F. Mastrogiovanni and N. Y. Chong, "An Integrated Epigenetic Robot Architecture via Context-influenced Long-Term Memory," in *2014 Joint IEEE International Conferences on Development and Learning and Epigenetic Robotics (ICDL-Epirob)*, 2014.
- [59] A. D. Baddeley, "The episodic buffer: A new component of working memory?," *Trends in Cognitive Science*, vol. 4, no. 11, pp. 417- 423, 2000.
- [60] J. Winkler, M. Tenorth, A. Kaan Bozcuoğlu and M. Beetz, "CRAMm – Memories for Robots Performing Everyday Manipulation Activities," in *Proc. of the Second Annual Conference on Advances in Cognitive Systems*, 2013.
- [61] M. Tenorth and M. Beetz, "KnowRob – A Knowledge Processing Infrastructure for Cognition-enabled Robots. Part 1: The KnowRob System," *International Journal of Robotics Research*, no. 32, pp. 566-590, 2013.
- [62] J. L. Phillips and D. Noelle, "Working Memory for Robots: Inspirations from Computational Neuroscience," in *5th International Conference on Development and Learning*, 2006.
- [63] R. C. O'Reilly, D. C. Noelle, T. S. Braver and J. D. Cohen, "Prefrontal cortex and dynamic categorization tasks: Representational organization and neuromodulatory control," *Cerebral Cortex*, vol. 12, no. 3, pp. 246-257, 2002.

- [64] J. L. Phillips and D. C. Noelle, "A biologically inspired working memory framework for robots," in *IEEE International Workshop on Robot and Human Interactive Communication*, 2005.
- [65] M. Tugcu, X. Wang, J. E. Hunter, J. Phillips, D. Noelle and D. M. Wilkes, "A computational neuroscience model of working memory with application to robot perceptual learning," in *3rd IASTED International Conference on Computational Intelligence*, 2003.
- [66] M. Skubic, D. Noelle, M. Wilkes, K. Kawamura and J. M. Keller, "A Biologically Inspired Adaptive Working Memory for Robots," in *2004 AAAI Symposium Series*, Washington D.C., 2004.
- [67] B. Goertzel, R. Lian, I. Arel, H. de Garis and S. Chen, "A world survey of artificial brain projects, Part II: Biologically inspired cognitive architectures," *Neurocomputing*, vol. 74, no. 1-3, pp. 30-49, 2010.
- [68] B. Goertzel, "Opencogprime: a cognitive synergy based architecture for embodied artificial general intelligence," in *8th IEEE International Conference on Cognitive Informatics*, Hong Kong, 2009.
- [69] D. Danks, *Unifying the Mind, Cognitive Representations as Graphical Models*, MIT Press, 2014.
- [70] N. Cowan, *Attention and memory: an integrated framework*, Oxford: Oxford University Press, 1995.
- [71] L. Wittgenstein, *Philosophical Investigations*, Blackwell, 1953.
- [72] J. R. Firth, "A synopsis of linguistic theory, 1930-1955.," in *Studies in linguistic analysis*, In Philological Society (Great Britain), Ed., Oxford, Blackwell, 1957, pp. 1-32.
- [73] Z. Harris, "Distributional structure," in *Papers in Structural and Transformational Linguistics*, D. Reidel, Ed., Dordrecht, 1970, pp. 775-794.
- [74] T. Landauer and S. Dumais, "A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and

- representation of knowledge," *Psychological Review*, vol. 104, no. 2, p. 211–240, 1997.
- [75] C. Shaoul and C. Westbury, "Word frequency effects in highdimensional co-occurrence models: A new approach," *Behavior Research Methods*, vol. 38, no. 2, p. 190–195, 2006.
- [76] T. L. Griffiths, M. Steyvers and J. B. Tenenbaum, "Topics in semantic representations," *Psychol. Rev.*, vol. 114, pp. 211–244, 2007.
- [77] K. Lund and C. Burgess, "Producing high-dimensional semantic spaces from lexical co-occurrence," *Behavioral Research Methods, Instrumentation, and Computers*, vol. 28, no. 2, p. 203–208, 1996.
- [78] M. N. Jones, J. Willits and S. Dennis, "Models of Semantic memory," in *The Oxford Handbook of Computational and Mathematical Psychology*, J. T. Townsend and J. R. Busemeyer, Eds., New York, Oxford University Press, 2015.
- [79] M. N. Jones and D. J. K. Mewhort, "Representing word meaning and order information in a composite holographic lexicon," *Psychological Review*, vol. 114, no. 1, p. 1–37, 2007.
- [80] M. W. Howard and M. J. Kahana, "A distributed representation of temporal context," *Journal of Mathematical Psychology*, vol. 46, no. 3, pp. 269–299, 2002.
- [81] M. W. Howard, M. J. Kahana and A. Wingfield, "Aging and contextual binding: Modeling recency and lag-recency effects with the temporal context model," *Psychonomic Bulletin & Review*, vol. 13, pp. 439–445, 2008.
- [82] M. W. Howard, K. H. Shankar and U. K. K. Jagadisan, "Constructing semantic representations from a gradually-changing representation of temporal context," *Topics Cogn. Sci.*, vol. 3, no. 1, pp. 48–73, 2011.

- [83] W. Wenwen, T. Ah-Hwee and T. Loo-Nin, "Semantic Memory Modeling and Memory in Learning Agents," *IEEE Transactions on System, Man and Cybernetics*, vol. 47, no. 11, pp. 2882-2895, 2017.
- [84] D. E. Rumelhart, J. L. McClelland and PDP research group, *Parallel distributed processing: Explorations in the microstructure of cognition. Volume I.*, Cambridge: MIT Press, 1986.
- [85] T. T. Rogers and J. L. McClelland, "Parallel Distributed Processing at 25: Further Explorations in the Microstructure of Cognition," *Cognitive Science*, vol. 38, no. 6, pp. 1024-1077, 2014.
- [86] D. E. Rumelhart and P. M. Todd, "Learning and connectionist representations," in *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience*, D. E. Meyer and S. Kornblum, Eds., Cambridge, MIT Press, 1993, pp. 3-30.
- [87] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proceedings of the National Academy of Sciences of the USA*, vol. 79, no. 8, pp. 2554-2558, 1982.
- [88] D. C. Plaut and J. R. Booth, "Individual and developmental differences in semantic priming: Empirical and computational support for a single-mechanism account of lexical processing," *Psychological Review*, vol. 107, no. 4, p. 786–823, 2000.
- [89] G. S. Cree, K. McRae and C. McNorgan, "An attractor model of lexical conceptual processing: Simulating semantic priming," *Cognitive Science*, vol. 23, no. 3, p. 371–414, 1999.
- [90] J. L. McClelland and R. M. Thompson, "Using domain-general principles to explain children's causal reasoning abilities," *Developmental Science*, vol. 10, no. 3, pp. 333-356, 2007.
- [91] D. E. Rumelhart, G. E. Hinton and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, no. 323, p. 533–536, 1986.

- [92] A. H. Tan, G. A. Carpenter and S. Grossberg, "Intelligence Through Interaction: Towards a Unified Theory for Learning," in *Advances in Neural Networks. Lecture Notes in Computer Science*, vol. 4491, D. Liu, S. Fei, Z. G. Hou, H. Zhang and C. Sun, Eds., Berlin, Springer, 2007.
- [93] C. Cuppini, E. Magosso and M. Ursino, "A neural network model of semantic memory linking feature-based object representation and words," *Biosystems*, vol. 96, no. 3, pp. 195-205, 2009.
- [94] S. Recanatesi, M. Katkov, S. Romani and M. Tsodyks, "Neural Network Model of Memory Retrieval," *Frontiers in Computational Neuroscience*, vol. 9, pp. 1-11, 2015.
- [95] K. Poettrich, P. H. Weiss, A. Werner, S. Lux, M. Donix, J. Gerber, R. von Kummer, G. R. Fink and V. A. Holthoff, "Altered neural network supporting declarative long-term memory in mild cognitive impairment," *Neurobiology of Aging*, vol. 30, no. 2, pp. 284-298, 2009.
- [96] D. Horn, E. Ruppin, M. Usher and M. Herrmann, "Neural Network Modeling of Memory Deterioration in Alzheimer's Disease," *Neural Computation*, vol. 5, no. 5, pp. 736-749, 1993.
- [97] S. Savić, M. Gnjatović, D. Mišković, J. Tasevski and N. Maček, "Cognitively-Inspired Symbolic Framework for Knowledge Representation," in *8th IEEE International Conference on Cognitive InfoCommunications (CogInfoCom 2017)*, Debrecen, Hungary, 2017.
- [98] M. Gnjatović and D. Rösner, "Inducing genuine emotions in simulated speech-based human-machine interaction: the NIMITEK corpus," *IEEE Trans. Affect. Comput.*, pp. 132-144, 2010.
- [99] M. Gnjatović, *Adaptive Dialogue Management in Human-Machine Interaction*, Munich: Verlag Dr. Hut, 2009.
- [100] M. Gnjatović, "The Child, the Therapist and the Robot: Adaptive Dialogue Management in Three-Party Interaction," in *COST Workshop on Social Robotics, The Future Concept and Reality of*

Social Robotics: Challenges, Perception and Applications - The Role of Social Robots in Current and Future Society, Brussels, 2013.

- [101] B. Borovac, M. Gnjatović, S. Savić, M. Raković and M. Nikolić, "Human-Like Robot MARKO in the Rehabilitation of Children with Cerebral Palsy," in *New Trends in Medical and Service Robots. Mechanisms and Machine Science*, vol. 38, H. Bleuler, M. Bourri, F. Mondada, D. Pisla and A. Rodić, Eds., Springer, Cham, 2016.
- [102] M. Gnjatović, J. Tasevski, M. Nikolić, D. Mišković, B. Borovac and V. Delić, "Adaptive multimodal interaction with industrial robot," in *Intelligent Systems and Informatics (SISY), 2012 IEEE 10th Jubilee International Symposium on*, Subotica, Serbia, 2012.
- [103] D. Mišković, M. Gnjatović, P. Štrbac, B. Trenkić, N. Jakovljević and V. Delić, "Hybrid methodological approach to context-dependent speech recognition," *International Journal of Advanced Robotic Systems*, vol. 14, no. 1, pp. 1-12, 2017.
- [104] M. Gnjatovic, J. Tasevski, D. Mišković, S. Savić, B. Borovac, A. Mikov and R. Krasnik, "Pilot Corpus of Child-Robot Interaction in Therapeutic Settings," in *Proc. of the 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, Debrecen, Hungary, 2017.
- [105] М. Ѓњатовић, Увод у проналажење информација на вебу, Београд: Висока школа електротехнике и рачунарства струковних студија, 2017.
- [106] R. Björn and J. Born, "About Sleep's Role in Memory," *Psychological Reviews*, vol. 93, no. 2, pp. 681-766, 2013.
- [107] C. D. Manning, P. Raghavan and H. Schütze, *Introduction to Information Retrieval*, New York: Cambridge University Press, 2008.
- [108] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed., Prentice Hall, 2008.