ITHEA

International Journal

INFORMATION CONTENT & PROCESSING

2016 Volume 3 Number 2

International Journal
INFORMATION CONTENT & PROCESSING
Volume 3 / 2016, Number 2

IJ ICP is official publisher of the scientific papers of the members of the ITHEA® International Scientific Society

IJ ICP rules for preparing the manuscripts are compulsory.
The **rules for the papers** for ITHEA International Journals as well as the **subscription fees** are given on www.ithea.org .
The papers should be submitted by ITHEA® Submission system http://ij.ithea.org .
Responsibility for papers published in IJ ICP belongs to authors.

International Journal "INFORMATION CONTENT AND PROCESSING" Volume 3, Number 2, 2016

Edited by the **Institute of Information Theories and Applications FOI ITHEA**, Bulgaria, in collaboration with
Institute of Mathematics and Informatics, BAS, Bulgaria,
V.M.Glushkov Institute of Cybernetics of NAS, Ukraine,
Universidad Politechnika de Madrid, Spain,
Hasselt University, Belgium
Institute of Informatics Problems of the RAS, Russia,
St. Petersburg Institute of Informatics, RAS, Russia
Institute for Informatics and Automation Problems, NAS of the Republic of Armenia.

Publisher: **ITHEA®**
Sofia, 1000, P.O.B. 775, Bulgaria. www.ithea.org, e-mail: office@ithea.org
Technical editor: Ina Markova

Printed in Bulgaria

ISSN 2367-5128 (printed) **ISSN 2367-5152 (online)**

GENDER DIFFERENCES IN THE USE OF NOUN CONCEPT CATEGORIES – A STATISTICAL STUDY BASED ON DATA FROM CHILD LANGUAGE ACQUISITION

Velina Slavova, Dimitar Atanasov, Filip Andonov

***Abstract**: We have analyzed data derived from large corpora of child language acquisition in the attempt to build a model of primary semantic categories of nouns. The statistical result found for noun frequencies, compared with results from brain imaging studies in adults suggests the possible existence of mental representations in early child development that shape the structure of the semantic space. We propose a set of primary categories of noun-concepts and investigate the influence of age and gender on the intensity of use of nouns from these categories. The statistical analysis of the progressive use of nouns from the proposed categories shows a coincidence with a wide range of theories of gender differences. From our data, it is not possible to distinguish between learned gender aptitudes and innate preferences. However, our statistical result supports the idea that gender distinctions may date to early human history.*

***Keywords:** cognitive modeling, mental representations, concept categories, gender differences*

***ACM Classification Keywords**: 1.2.7. Natural language processing, 1.2.0. Cognitive simulation, G.3 Correlation and regression analysis, H.2.8. Database applications.*

Introduction

This paper is based on the idea that the examination of child language acquisition gives scholars the opportunity to investigate the semantics and principles of human language.

One of the earliest scientific explanations of language acquisition was provided by [Skinner, 1957] accounting for language development by means of environmental influence. The behaviorist idea that infants imitate their environment reverses the view of Chomsky claiming that if the language acquisition mechanism was dependent on language input alone, children will never acquire to process an infinite number of sentences. The theory of Universal Grammar states the existence of innate, biological grammatical categories, such as a noun category and a verb category. Later some psycho-linguists argued that innate grammatical categories are biologically, evolutionarily and psychologically implausible. Researchers started to suggest that instead of having a language-specific mechanism for language processing, children utilize general cognitive and learning principles, returning to Piaget's view

that learning language is mainly focused around cognitive development (e.g. [Piaget, 1955]). For example, many works of Michael Tomasello claim the “language instinct” cannot explain how children learn language, but their linguistic ability is interwoven with other cognitive abilities (see for ex. [Tomasello M., 2003]).

Other studies concentrate on the social interaction and demonstrated its importance for the acquisition of language and for the overall cognitive development. For example, the analysis of vast observational data of 1- through 3-year-olds learning to talk during their everyday interactions with their parents, showed how crucial to development is the amount of children's language experience as partners in the social dances of conversation ([Hart & Risley, 1995], [Hart, 2000]).

Specialized studies suggest gender differences in first language acquisition. Fundamental works on verbal ability have shown that girls and women surpass boys and men in verbal fluency, correct language usage, sentence complexity, grammatical structure, spelling, and articulation [Eckert & McConnell-Ginet, 2003]. These general suggestions are regularly confirmed in statistical studies (e.g. [Slik, Hout & Schepens, 2015]). Although these observations are commonly accepted by the scientific community, the reasons for the existence of such differences remain unclear.

Several studies have been conducted in order to explain these gender related differences. For example, fMRI studies have suggested that girls rely on a supramodal language network, whereas boys process visual and auditory words differently [Burman et al., 2008]. Other researchers propose that gender differences are genetically determined. Researchers have found that the amount of FOXP2 protein (a gene related to language) in the brains of 4 to 5 years old boys and girls (in Brodmann area 44) is 30% more in girls’ brains as compared to boys’ [Balter, 2013].

Differences in brain areas related to language faculty were found in studies on gender differences (see [Northwestern University, 2008]). Analyses of these results have led researchers to the assumption that the observed distinctions may date to early human history.

Concerning the importance of social interaction mentioned above, a recent study of 268 children aged between 18 and 35 months has shown that the association between expressive language and social ability is significantly stronger in boys than in girls [Longobardi et al, 2016].

However, finding definitive answers to the questions of language acquisition and gender differences is an unresolved problem.

Our research does not provide an explanation for language-related gender differences. We present the reasoning which has led to the establishment of a set of noun-concept categories, proposed after analysis of brain imaging studies, models of cognition and our data. We performed a statistical analysis of the use of these noun concept categories and found statistically significant gender differences. The

results we have obtained from the statistical analysis of the use of noun-concept categories by children can help the deeper understanding of the problem.

Basic Assumptions and Data Description

The assumptions in this paper follow the model of language faculty proposed in [Slavova & Soschen 2015]. The analysis provided there of contemporary findings in brain imaging, neuroscience, cognitive science, psychology, linguistics etc. lead to the conclusion that the building of a mental representation of the world starts on the basis of genetically determined information treatment processes. Following the proposed *Self-centered model of language faculty*, the primary mental representations are obtained using neuron networks present and functioning at birth - multimodal perception, proprioception, interoception, mirror neuron system and default mode network. The central suggestion derived from the analysis of the specialized sources was that a new-born has a biologically underwritten notion about the existence of his Self as actor in the environment (in compliance with [Barsalou, 2003]). It was suggested that semantic categories are further shaped from the point of view of personal, situational experience. In big, the proposed there model focuses on the creation of Meaning as internal mental representation, which establishment relies initially on inborn mechanisms, on the role of an actor in the environment and on the innate "knowledge" about the existence the Self and the self-similar.

The statistical study presented here concerns the questions of Meaning. We assume that the development of the semantic representation of the world starts on the basis of primary conceptualization mechanisms realized by biologically underwritten processing of "automatic" classification of the information into semantic categories, necessary to guide actions of importance for the individual's survival. We suppose that the initial concepts arise as internal information units which creation relies on inborn brain processing that organizes the information flow coming from the environment and the flow from the "inside" of the biological system, insuring its functioning.

Our study is based on the assumption that language faculty is a *result* of the capacity for conceptualization. We concentrate on the analysis of nouns as the use of nouns starts first in language production, apparently being the most natural vehicle of meaning, and strongly dominates child speech during the period of language acquisition considered in our data.

Data from 30 corpora containing dialogues of child speech in English, annotated with parts of speech and grammar, were extracted from CHILDES [MacWhinney, B. & C. Snow (1985)], [MacWhinney, 2000]) and stored and organized in a relational database [Slavova, 2016]. The obtained data collection contains 125,584 speech utterances of different children aged between 9 and 62 months, produced in free dialogues and collected by researchers in child language acquisition during several decades. Our study is based on the linguistic annotation for parts of speech (POS) taken from CHILDES which

respects the developed MORPH system [Hausser, R. (1989)]. Our previous statistical results based on the ratios of use of POS [Atanasov et al., 2016] confirmed the "special" role of nouns in language acquisition.

Proposed Model of Noun Concept Categories

We try to elaborate a set of noun-concept categories which are "primary" as they emerge using inborn mechanisms for semantic assessment of the perceived Entities. For the purpose, we have looked into the results coming from brain studies in adults, searching for reported "particular features".

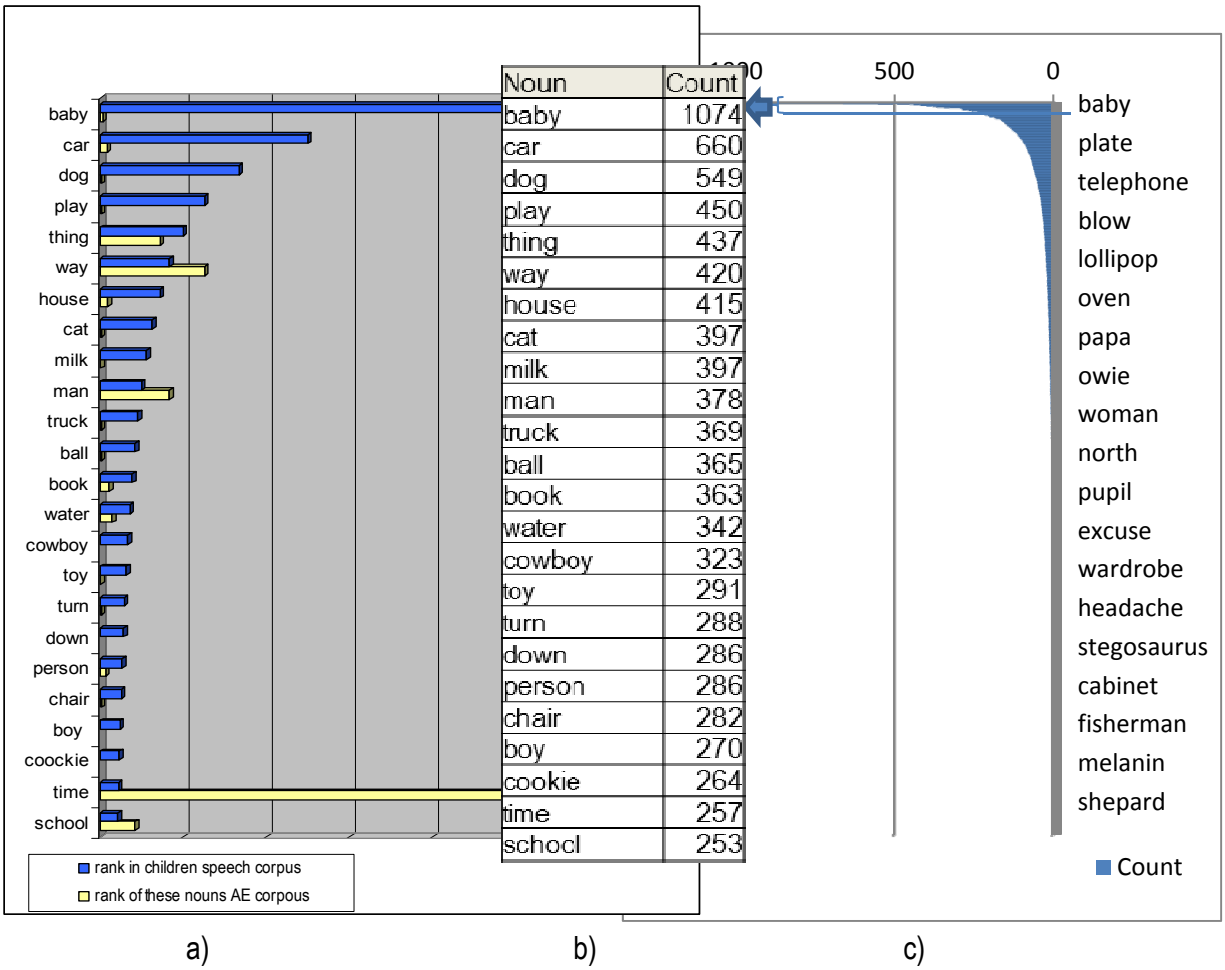
Studies based on application of voxel-wise models and huge fMRI data [Huth et al., 2012, 2016] show that the thousands of distinct object and action categories that humans see and name are represented as locations in a continuous semantic space mapped smoothly across the overall cortical surface. In [Huth et al., 2012] authors used movies to examine the cortical representation of 1,705 object and action categories. The first few dimensions of the underlying semantic space were recovered from the fit models by principal components analysis (PCA). The results suggest that the overall brain activation related to semantics can be represented in a 4-dimensional space which is common across individuals. It is seen from the plots that this fMRI-data derived semantic space is organized across some distant points (meaning that the concepts evoke very specific brain activation) such as "*car*", "*man*", "*face*", "*room*", "*text*" and forms clusters such as "*animals*", "*body parts*", "*humans*", "*communication*", "*structures*", "moving *objects*", "*indoor category*", "*outdoor category*" etc. The finding has been confirmed and unreached recently [Huth et al. 2016]. To model brain responses elicited by *naturally spoken narrative stories* presented to 7 individuals, the authors applied again PCA of voxel-wise model weights. The results for the overall brain activation evoked by words were projected into 985-dimensional word embedding space constructed using word co-occurrence statistics from a large text-corpus. The statistical analyses lead to a space structured around four main axes. After evaluation of the predictive capacity of the proposed model, the authors suggest (again) that most brain areas within the semantic system represent information about specific semantic domains and the obtained brain atlas is common for the individuals (see http://www.nature.com/nature/journal/v532/n7600/full/nature17637.html).

According to our reasoning, if there is a brain atlas common for the individuals, its initial structure would be available at birth and detectable in the first period of language production. Our sample contains speech from the period when children have just started saying recognizable words. In our data from free dialogues, children say a word in order to communicate some idea, so they have a semantic representation for its meaning.

We compared the nouns used often in children speech with the structure of the semantic space derived by Huth and colleagues [Huth et al., 2012]. The examination yielded intriguing results. The Entities

expressed frequently by small children correspond surprisingly well to distant points of the adult fMRI-space (see http://www.cell.com/cms/attachment/2007952467/2030515204/gr5_lrg.jpg).

The nouns observed with significantly higher frequency in child speech are listed in Figure 1.b. A trivial check shows that only five of the first 25 frequently used by children nouns (time, way, man, thing and school) are among the first 25 nouns with high frequencies in the adult's speech corpus. Four of the first 25 "children-preferred" nouns (boy, down, cookie and cowboy) are not among the first 5000 words of the adults' speech (see http://www.wordfrequency.info). The other nouns from the "children list" rank in adults' speech in a quite different way (Figure 1.a.). Obviously there exist differences between children's and adults' vocabulary and experience. Our data suggests that the expressed by children noun-concepts are not a simple consequence of the content of language to which children are exposed in their everyday activities. We assume that the concepts expressed in the children's speech are related to their experience, but through the conceptualization ability available at the corresponding age.



a) b) c)

Figure 1. a) Nouns used the most frequently (from 250 to 1,074 times) observed in our data from child speech, compared with the frequencies in the adult's speech. b) List of the first 25 most frequently used by children nouns, c) general plot of all the frequencies (from 1 to 1,100) of the used by the children nouns

We supposed that the observed coincidence between the Entities expressed frequently by children and the distant points of the adult fMRI-space shows that that the semantic structure is gradually developed around some axes and points, which are available very early (our assumption being that the principles of conceptualization are based on an inborn apparatus for information categorization and insure the basis of the gradual development of the semantic description of the world).

We have discussed the technical aspects related to the data annotation and proposed a model of primary noun concept categories in previous works [Slavova et al., 2016]. Here we give the three main groups of arguments which have led to this set of categories:

1. *Brain studies* [Huth et al., 2012, 2016] - the activation of the overall brain related to semantics has an underlying structure, common across individuals. We have taken into account particularities discussed in this section as well as other suggestions, for example, following [Huth et al., 2016], the perceptual and physical categories (tactile, locational) are separated from human-related categories (social, emotional, violent) by the axis which lies along the first dimension of the common semantic space derived by means of PCA.
2. The *self-centered model of language faculty*, following which the semantic categories are primarily shaped from the point of view of personal proprioceptive, interoceptive and perceptual, situational experience, as explained in the section “basic assumptions”.
3. The observation and the analysis of the child speech collected in our database.

Based on these points, we assembled a set of 14 categories, shown in Figure 2.

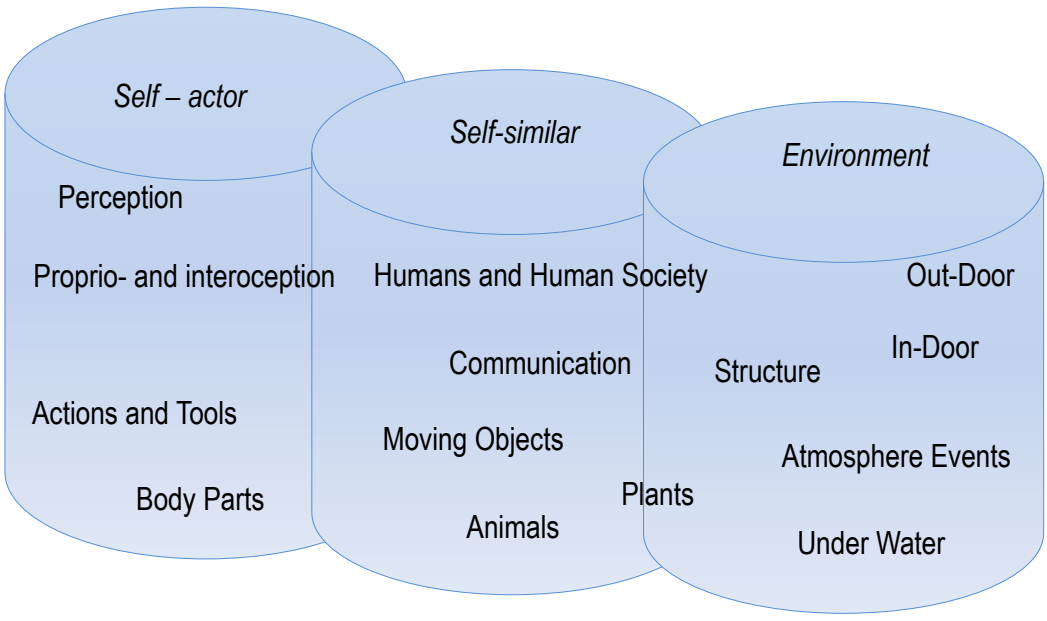


Figure 2. Proposed set of noun concept categories [Slavova et al., 2016]

Investigation of the Use of the Proposed Noun Concept Categories

In order to study the use of nouns from the proposed set, we extracted and treated the nouns from the child speech database, obtaining a big sample of utterances - Nouns (47,777 records) and Proper nouns (12,673 records).

We annotated manually with categories of the proposed semantic set all the common (3026) and proper (1490) nouns extracted from the child speech [Slavova et al. 2016]. Examples for the annotated common nouns for each category are given in table 1.

Table 1. Examples of nouns, extracted from the child speech data

1	Self, proprioception; interoception	baby	hurt	anger	pain	sleep	dress
2	Perception	beauty	light	dark	black	color	voice
3	Actions and tools	action	move	swim	cry	drag	lego
4	Humans and society	friend	guest	family	sport	anger	artist
5	Body parts	arm	beak	beaker	beard	behind	mind
6	Animals	bat	bird	bunny	cat	chicken	mouse
7	Plants	beet	bran	branch	bush	flower	mulberry
8	Moving objects	aeroplane	automobile	baby	barge	beetle	locomotive
9	Indoor	bath	bed	blanket	bowl	candle	room
10	Outdoor	alley	avalanche	avenue	barbecue	forest	river
11	Communication	letter	library	magazine	newspaper	paper	joke
12	Atmospheric	air	tornado	candle	comet	dust	plane
13	Underwater	aquarium	corral	dive	noun	octopus	water
14	Structure	cube	edge	machine	block	count	evening

The "Self-proprioception and interoception" category contains nouns for pain, hunger (including foods), emotional states, feelings (i.e. coldness, warmth) and clothes. The "Perception category" contains nouns for incoming multimodal perceptual information. The "Structure" category contains materials, machines, spatial forms, time and time-periods, quantities, and socially determined structures such as hospitals and cities. The content of the other categories is clear from the given examples.

In order to evaluate how the proposed categories are used by children with relation of their age and gender, we calculated for each dialogue taken from CHILDES the Ratio per Utterance (RU) of the 14 categories:

$$\mathrm{RU}\left(\mathrm{Cat_{ij}}\right) = \frac{\mathrm{NCat_{ij}}}{\mathrm{N_i}} \qquad (1)$$

Where :

RU(Cat$_{ij}$) is the ratio of use in the dialogue *i* the category *j;*

NCat$_{ij}$ is the number of the nouns from category *j* in the dialogue *i;*

N$_i$ is the number of recognizable word-forms pronounced by the child in the dialogue *i*.

Statistical model

Based on the derived Ratio per Utterance we have realized a linear logistic model given with expression (2). For predictor parameters we have used the Age, the Gender and the Type of the noun (proper noun or common noun):

$$logit(\frac{Pr}{1-Pr}) = I + A * age + G * gender + T * type + AG * (age \times gender) + TG * (type \times gender) \quad (2)$$

Where:

A, G, T, AG, TG are the unknown parameters of the model;

the *Cartesian products* note the mixed effect of the two parameters involved;

Pr is the probability for performing (using in the speech) given concept category.

The results we have obtained are shown in Table 2, where with bold we give the statistically significant parameters (for which the estimated *p-value* is less than 0.1). As it is seen from the results, the parameters Gender, Age and Type influence each the categories except one – Plants.

The observations of use of nouns in the categories "Atmospheric events" and "Underwater" are further discarded because of their insufficient number in our sample. They are well-pronounced as distant "special points" in the adult's fMRI semantic space obtained by Huth and colleagues [Huth et al. 2012], but according to our data they are not used early in childhood.

Table 2. Estimation of the model parameters

model categories	A	G	A*G	T*G	T
01 Self, proprioeption, interoception	**0.0309**	**3.4932**	***-0.0079***	***-2.7225***	**6.4785**
02 Perception	***-0.0306***	1.2405	**0.0114**	***-1.8828***	**6.3894**
03 Actions and tools	***-0.0156***	0.8979	**0.0175**	***-1.8882***	**7.3912**
04 Humans and society	**0.0090**	0.0109	***-0.0159***	**0.8203**	***-5.4844***
05 Body parts	**0.0260**	0.9772	***-0.0192***	*-0.3763*	25.1459
06 Animals	**0.0248**	**0.9663**	***-0.0090***	***-0.3648***	**2.9320**
07 Plants	***-0.0096***	18.0716	*-0.0040*	18.4748	13.9519
08 Moving objects	**0.0169**	*-0.7577*	**0.0071**	0.4570	**5.0394**
09 Indoor	**0.0196**	*-0.8133*	***-0.0085***	1.4555	22.0610
10 Outdoor	***-0.0303***	*-1.8982*	**0.0165**	0.9104	23.0149
11 Communication	***-0.0136***	*-3.6582*	**0.0111**	3.2120	20.7665
12 Atmospheric	***-0.0345***	*-0.5586*	**0.0229**	*-0.6173*	24.6671
13 Underwater	***-0.0447***	*-2.0171*	**0.0271**	0.5412	21.9022
14 Structure	***-0.0138***	***-0.5373***	0.0013	**0.3736**	**2.0220**

As the ratios of use all the 14 categories are not independent in the speech, the results in table 2 cannot be seen as a hall, but has to be analyzed category by category. Examples of plots are shown in figure 3.

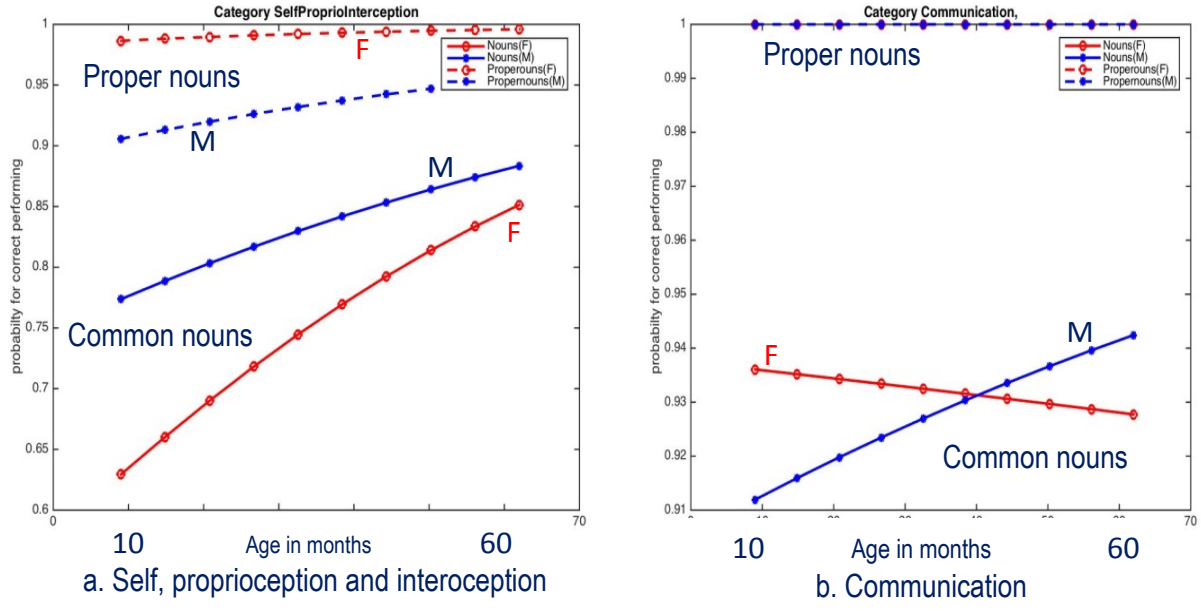


Figure 3. Examples of use of nouns and proper nouns in two of the categories.

Figure 3.a illustrates how develops with the age the probability for performance of gills (red) and boys (blue) of use of the category “Self-proprioception and interoception”, for the two types of nouns. As seen, for both genders this probability of use increases. The same dependencies are shown in figure 3.b. for the category Communication, where Proper nouns are not observed and are displayed as a constant. As it is seen, the performance of boys gradually increases, whereas girls tend to decrease the relative use of nouns from this category.

The results for all the other categories are obtained in the same way.

Analysis of the Results with regard of Gender Differences

The graphics in figure 4 expresses the overall result, where the categories are grouped depending on the influencing parameters. As shown, the gender influences the use of three of the categories – Self - proprioception and interoception, Structure and Animals (their plots are given in figure 3.a, figure 5.a and figure 5.b respectively).

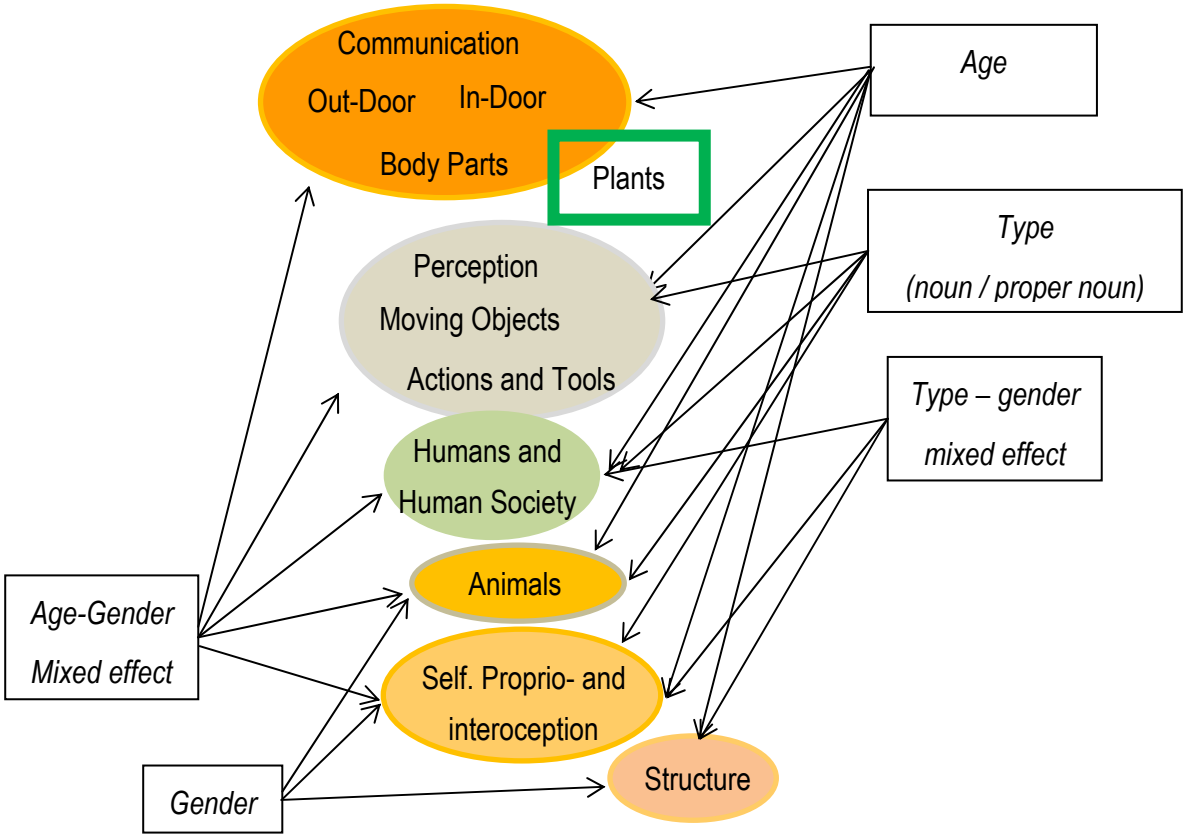
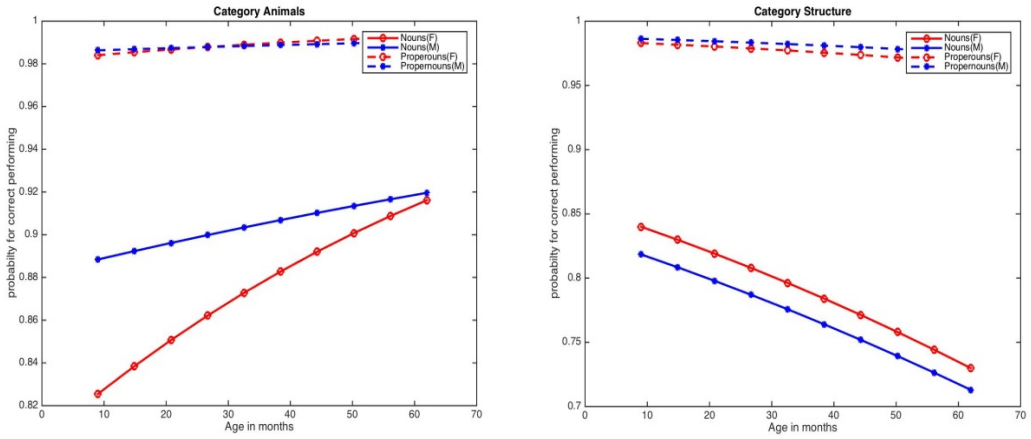


Figure 4. Influence of the parameters Age, Gender, Type of noun and their mixed effect on the proposed categories.

The relative use of the category “Structure” decreases with age for both genders, in the same way (figure 5.b). We have determined this category as a semantically large set (machines, spatial forms, time and time-periods, quantities, materials, and socially determined structures), so its split into sub-categories could give a more precise picture of its decreasing tendency of use.



a. Animals b. Structure

Figure 5. Use of nouns and proper nouns in two of the “gender dependent” categories by boys and girls.

Next, we examine how develops over the age the categorical content of boys’ and girls’ vocabulary. The value found for the mixed effect of Age and Gender (A*G) given in Table 2 shows how progresses the use of noun categories by the two genders.

Table 3. Gender differences in the use of noun concepts over the course of language acquisition

Categories in which ***girls*** show relatively increasing use of nouns		Categories in which ***boys*** show relatively increasing use of nouns	
model categories	**A*G**	**model categories**	**A*G**
Body parts	***-0,019***	**Actions and tools**	**0,018**
Humans and society	***-0,016***	**Outdoor**	**0,017**
Animals	*-0,009*	**Perception**	**0,011**
Indoor	*-0,009*	**Communication**	**0,011**
Self, proprioeption, interoception	*-0,008*	Moving objects	0,007

We have annotated the two genders with 1 for the girls and with 2 for the boys, so the positive value of A*G means that the use of the corresponding category by boys increases at a faster rate than that of girls. (Or, that the use of the category by boys decreases at a slower rate than that of girls.) A negative value obtained for A*G means that girls develop more intensively the use of nouns from this category that boys do.

After having taken into account only the statistically significant A*G (for *p-values* less than 0.1), the reorganization of the results from Table 2 provides the two lists given in Table 3. The lists show the concept-related differences in the "gender dependent development of the vocabulary" (in bold are the categories for which the differences are considerable, threshold taken at 0.01).

These "gender-distinctive" lists of categories showed up after the concluding analysis of the statistical result obtained from the vast corpus of child speech. As it can be seen, the result corresponds to the classical view about the social function of "men-hunters" and "women-gatherers".

It is not possible from our data to distinguish between abilities which are learned and abilities which are innate. However, our statistical finding supports in general the supposal that the observable gender distinctions may date to early human history.

Conclusion

We proposed a set of "primary" noun-concept categories, supposing that the conceptualization mechanisms are functional at birth and taking into account results obtained in studies of the semantic brain activation in adults.

We classified the nouns from a huge database of child speech to the proposed set. The linear logistic model we have applied gave statistically reliable results concerning the influence of Age, Gender and Type of noun on the use of these categories.

The obtained statistical result is in compliance with suggestions and general theories concerning social, cognitive, anthropological etc. gender differences.

As we don't treat purely language-related parameters such as richness of vocabulary, correctness, grammatical parameters etc., we accord the result to the establishment of initial mental representations obtained in interaction with the environment.

At this point, the proposed set of noun-concept categories gives a meaningful picture and the statistical result indicates directions for further investigation and adjustment of the noun-concept's model.

Acknowledgement

This paper is published with partial support by the ITHEA ISS (www.ithea.org) and the Central Fund for Strategic development, New Bulgarian University.

References

[Atanasov et al, 2016] Atanasov, D., Slavova V. & Andonov F. A statistical study of first language acquisition: no gender differences in the use of parts of speech, in proc. of the 12th Annual International Conference on CSECS 2016, Germany, in print

[Balter, 2013] Balter, M. Language Gene More Active in Young Girls than Boys. ScienceMag.org, Balter, M. (2013, February 19). http://www.sciencemag.org/news/2013/02/language-gene-more-active-young-girls-boys

[Barsalou, 2003] Barsalou, L. W. Situated simulation of human cognitive processes, Language and cognitive processes, 2003, 18 (516), 513-562

[Burman el al, 2008] Burman, D. D., Bitan, T., & Booth, J. R. Sex differences in neural processing of language among children. Neuropsychologia, 46(5), 2008. 1349-1362.

[Eckert & McConnell-Ginet, 2003] Eckert, P., & McConnell-Ginet, S. Language and gender. Cambridge University Press, 2003.

[Hart & Risley, 1995] Hart, B., & Risley, T. R. Meaningful differences in the everyday experience of young American children. Paul H Brookes Publishing, 1995.

[Hart, 2000] Hart, B. A natural history of early language experience. Topics in Early Childhood Special Education, 20(1), 2000, 28-32

[Huth et al, 2012] Huth, A. G., Nishimoto, S., Vu, A. T., & Gallant, J. L. A continuous semantic space describes the representation of thousands of object and action categories across the human brain. Neuron, 76(6), 2012, 1210-1224.

[Huth et al, 2016] Huth A.G., de Heer W.A., Griffiths T.L., Theunissen F. E. & Gallant J. L. Natural speech reveals the semantic maps that tile human cerebral cortex, Nature, 2016 .

[Longobardi et al, 2016] Longobardi, E., Spataro, P., Frigerio, A., & Rescorla, L. Gender differences in the relationship between language and social competence in preschool children. Infant Behavior and Development, 43, 2016, 1-4.

[Northwestern University, 2008] Northwestern University. Boys' And Girls' Brains Are Different: Gender Differences In Language Appear Biological. ScienceDaily. 2008, March 5. www.sciencedaily.com/releases/2008/03/080303120346.htm

[Piaget, 1955] Piaget, Jean. The Language and Thought of the Child, New York: The World Publishing, 1955, 1924/55

[Skinner, 1953] Skinner, B. F. Science and human behavior. Simon and Schuster. 1953

[Slavova & Soschen, 2015] Slavova, V., A. Soschen, On mental representations: Language structure and meaning revised, International Journal Information theories & applications. 2 (4), 2015, 316-325.

[Slavova, 2016] Slavova V. Data collection for studying language acquisition, in proc. of the 12th Annual International Conference on CSECS 2016, Germany, in print

[Slavova et al, 2016] Slavova V., Andonov F. & D. Atanasov (2016), A study of noun concept categories using data from child language acquisition – an outline, in proc. of the 12th Annual International Conference on CSECS 2016, Germany, in print

[Slik, Hout & Schepens, 2015] van der Slik FWP, van Hout RWNM & JJ Schepens (2015) The Gender Gap in Second Language Acquisition: Gender Differences in the Acquisition of Dutch among Immigrants from 88 Countries with 49 Mother Tongues. PLoS ONE 10(11): e0142056. doi:10.1371/journal.pone.0142056

[Tomasello, 2003] Tomasello, M. Constructing a Language: A Usage-Based Theory of Language Acquisition. Cambridge, MA: Harvard University Press, 2003.

Authors' Information

***Velina SLAVOVA**, New Bulgarian University, department of Computer Science, vslavova@nbu.bg*

***Major Fields of Scientific Research**: AI, Cognitive Science*

***Dimitar ATANASOV**, New Bulgarian University, department of Computer Science, datanasov@nbu.bg*

***Major Fields of Scientific Research**: Probability, Statistics and related fields, Psychometrics.*

***Filip ANDONOV**, New Bulgarian University, department of Computer Science, fandonov@nbu.bg.*

***Major Fields of Scientific Research**: multicriteria optimization, data mining, text processing, Python language*

COGNITIVE AGENT BASED SIMULATION PLATFORM FOR MODELING LARGE-SCALE MULTI-LEVEL SOCIAL INTERACTIONS WITH EXPERIMENTAL GAMES

Maurice Grinberg, Emiliyan Todorov

Abstract: *The paper introduces a new multi-agent architecture for simulations of complex social interactions with cognitive agents of different sophistication using experimental games. Such games are the Prisoner's dilemma, Chicken, Battle of the Sexes, Ultimatum, Trust, and Dictator games, Public goods game, and other important games that have been used to investigate social dilemmas. The main principles and innovative components are presented together with the design and first implemented components of a computer architecture aimed at large scale distributed calculations allowing for simulations involving millions of interacting agent living in environments built on games. The main goal of the ABM architecture is to build and explore societies of artificial cognitive agents based on rich social interactions modeled by games with agent implementing various computational and cognitive models. The platforms introduces several innovative tools like agent and environment spaces of features and information, interfaces for cognitive models incorporation, possibility to build social environments on experimental game-theoretic games.*

Keywords: *agent-based modeling, cognitive modeling, social interactions, experimental game theory, parallel, distributed systems*

ACM Classification Keywords: *I.6 Modeling and Simulations*

Introduction

Social interactions have been subject of intense interest in many scientific fields like biology, psychology, social psychology, and philosophy. The main focus of this research was on cooperation among various agents like bacteria, animals, humans, etc. The importance of the understanding of the mechanisms behind cooperative behavior, i.e. going far beyond self-interest, is related to the understanding of the main driving force of evolution and human society and culture.

Any exploration of the mechanisms of cooperation in strategic interactions requires appropriate approaches and models. One of the most influential approaches is Game Theory (GT) proposed by Morgenstern and von Neumann [Morgenstern and von Neumann, 1947]. Although GT is a normative and not a descriptive theory, it allowed building a formal theory based on preference for outcomes and

expected utility maximization. GT makes several strong assumptions which make possible the derivation of the expected utility maximization principle. For instance GT assumes that the definition of the game by its payoff matrix (strategies and outcomes) and everything that can be deduced from it is common knowledge, i.e. players know it, and they know that the other players know it, and that they know that the other players know that they know, etc. Additionally, players are assumed to be instrumentally rational which means that they always chose strategies which maximize their own payoff, based on their knowledge and beliefs, the latter assumed to be also common knowledge.

Although many discrepancies between extensive empirical data from experimental and behavioral game theory [Camerer, 2003; Colman, 1995; Fehr & Fischbacher, 2004; Van Lange et al. 2013] and GT have been established, the main idea of GT to model social interactions and more specifically social dilemmas with games has been widely accepted.

Some recent theoretical approaches assume people maximize some utility function as in GT but propose new decision making mechanisms like Cognitive hierarchy theory [Camerer, Ho, & Chong, 2004], Stackelberg reasoning, and team reasoning [Colman, Pulford, & Lawrence, 2014] which involve taking into account a model of the opponent and her payoffs. Related models are Social projection theory, e.g. [Acevedo & Krueger, 2005] and other-regarding social values and preferences [Fehr & Schmidt, 2006; Van Lange & Rusbult, 2011].

Some more minimalistic models of decision making are based on reinforcement learning [Camerer, Ho, & Chong, 2002; Erev & Roth, 1999; Grinberg, Hristova, & Lalev, 2010; Macy & Flache, 2002] or on heuristics (e.g. the 'avoid the worst' heuristics) [Gigerenzer & Goldstein, 1996; Krueger, 2014]. An interesting perspective in experimental games is the concept of psychological games [Geanakoplos, Pearce, & Stacchetti, 1989] in which the choices of the agents depend not only on the payoffs of the game and the choices made but also on the beliefs about the other agents, and their beliefs and intentions. This approach allows accounting for the emotions of other players and their influence on game outcomes and equilibria.

All these approaches have their strengths and weaknesses and are complementary in the sense that none of them can account fully for the data and there are many evidences showing that one and the same player can use more than one of them depending on the characteristics of the games [Camerer, 2003; Colman et al., 2014; Fehr & Schmidt, 2006] The brief discussion of the various approaches to investigate the mechanisms of decision making in experimental games shows the complexity of the factors that have been considered. In most models of decision making, the influence of GT is quite strong and they try to modify the way players reason about the game in order to account for experimental results (e.g. by using other-regarding strategies). Some of the models start by defining a

utility function or a quantity which reflects the attractiveness of a strategy which then allow to determine the best move or the probability of making such a move [Flache & Macy, 2002].

One possible development [Chater, 2015; Sun, 2006] is to apply the progress made in the field of cognitive modeling for agents in simulation of decision making in experimental games. Such attempts are not very common in the literature so far (e.g. [Grinberg & Lalev, 2008; Taiji & Ikegami, 1999; West, Lebiere, & Bothell, 2006]). In general, players are not explicitly considered as cognitive agents with specified perceptual capabilities, attention mechanisms, memory, etc.

Another relatively distinct, but very influential approach for exploring social interactions and their evolution is Evolutionary game theory [Gintis, 2009; Nowak, 2006; Maynard Smith, 1982]. One of the most influential example of this approach is the study of cooperation in iterated Prisoner's dilemma game tournaments [Axelrod, 1984; Axelrod, 1997; Axelrod & Hamilton, 1981] and the demonstrated advantages and power of agent-based model simulations.

The experience of this line of research showed that, in some situations, simulations provide solutions and evolutionary stable strategies [Adami et al. 2015] which are beyond the ones expected based on GT. These and many other results demonstrate the power of multi-agent simulations, based on rigorous mathematical treatments [Shoham & Leyton-Brown, 2008; Epstein, 2006, 2014].

The goal of the present paper is to introduce the idea and the first implementation steps of a novel approach which aims at building a flexible large scale distributed multi-agent platform entirely based on experimental game-theoretic interactions, agents implementing various computational and cognitive models of decision making, and various learning and evolutionary approaches. The platform focuses on the multiple roles an agent can have depending on the group, environment, and the specific interactions (family, company, country, etc.) they imply. A central question of interest is the interactions among these social roles related to cooperation and competition. Part of the social theoretical background of the platform was broadly inspired by the social relational model theory of Alan Fiske [Fiske, 1992; Fiske & Haslam, 1996; Fiske, 2012]. It posits four main relationships which according to this theory underlie any more complex relations: communal sharing, authority ranking, equality matching, and market pricing. According to Fiske [Fiske, 1992], these four relationships can be regarded as psychological models that underlie human sociality. This theory has a considerable empirical support but what is more important here is its amenability to game theoretic terms one version of which has been explored in [Grinberg, Hristova, & Borisova, 2012; Hristova, Grinberg, Georgieva, & Borisova, 2013].

In our agent-based platform, we want to provide the possibility to build social environments consisting only of games that model various social interactions (for a formal multi-agent approach in game theory see [Shoham & Leyton-Brown, 2008]). For instance, some of these games or combinations of them can model the four relationships in Fiske's relational model theory but could also stand for interactions

among artificial agents or artificial agent-human interactions [Grinberg, 2011; West et al., 2006]. This essential capability of the platform will allow for the construction of complex environments in which the dynamics of cooperation among agents of various sophistication can be explored.

The agents in our simulation environment will be modelled after the general model of a cognitive architecture (see e.g. [Grinberg, 2011; West et al., 2006] providing an interface appropriate for the implementation of agent models from the main modern approaches in cognitive modelling (computational models, connectionist models, dynamic system models, Bayesian models, etc.).

A third innovative aspect of the multi-agent platform is the conceptualization of agents as multi-dimensional vectors in a space spanned by the various characteristics of the agents which can include not only spatial and temporal localization but also specific parameters (with a predefined distribution or learned), history of strategy choices, accumulated payoffs in various games, etc. Such multi-dimensional agent space would allow finding closeness and similarity among agents and providing information for in-depth analysis of the structure of the artificial agent societies. This information also would allow for the analysis of the complex network structure that emerges out of the multi-level interactions and apply social network analysis to understand better the results of the simulations.

A fourth important feature of the environment is the possibility to run simulations involving large number of agents (up to billions) in order to be able to explore large scale phenomena of multi-agent interaction in a multidimensional space.

While the theoretical background behind the multi-agent platform with characteristic examples will be presented elsewhere, in this paper, we want to report the progress made so far in its implementation.

Agent Based Modeling

Agent Based Modeling (ABM) is an approach for modeling and simulating complex systems composed of autonomous agents interacting with one another [Macal & North, 2010; Epstein, 2006]. The basic idea is to represent inhabitants and artifacts of the real world as mock-up agents inside an artificial environment. Then let them communicate and act with one another inside it, according to specific rules and observe their evolution. This way agents can influence each other, learn from their experience and modify their behavior to survive in an ever changing environment. The environment modifies the behavior of the agents and the agents modify the environment, which leads to a system which is "…therefore, emergent on the interaction of the individual parts." [Barnes & Chu, 2015]. The application of ABM spreads through a wide range of areas and domains – social sciences [Axelrod, 1997], bioinformatics [Barnes & Chu, 2015], epidemics [Parker & Epstein, 2011], etc.

The typical structure of an ABM platform consists of three main elements [Macal & North, 2010]:

1. Autonomous agents with attributes and behavior.
2. Relationships and methods of interaction and communication.
3. Environment to interact with and within in addition to other agents.

Agents are supposed to be heterogeneous and active entities in pursuit of their internal goals, rather than returning passive responses (Figure 1). Relationships between agents are generally defined by distances and connectedness in an underlying space where agents are situated. The most used topologies are grids, Euclidean space, graph, Geographic Information System (GIS) and a spatial "Soup" model. Many ABMs include agents interacting in multiple topologies. Additionally the environment may contain constraints for the evolution of the agents like restricted amount of resources available to them, or restrains over the exact form and implementation of the topology (e.g. infrastructure, capacities of nodes, number of links allowed).

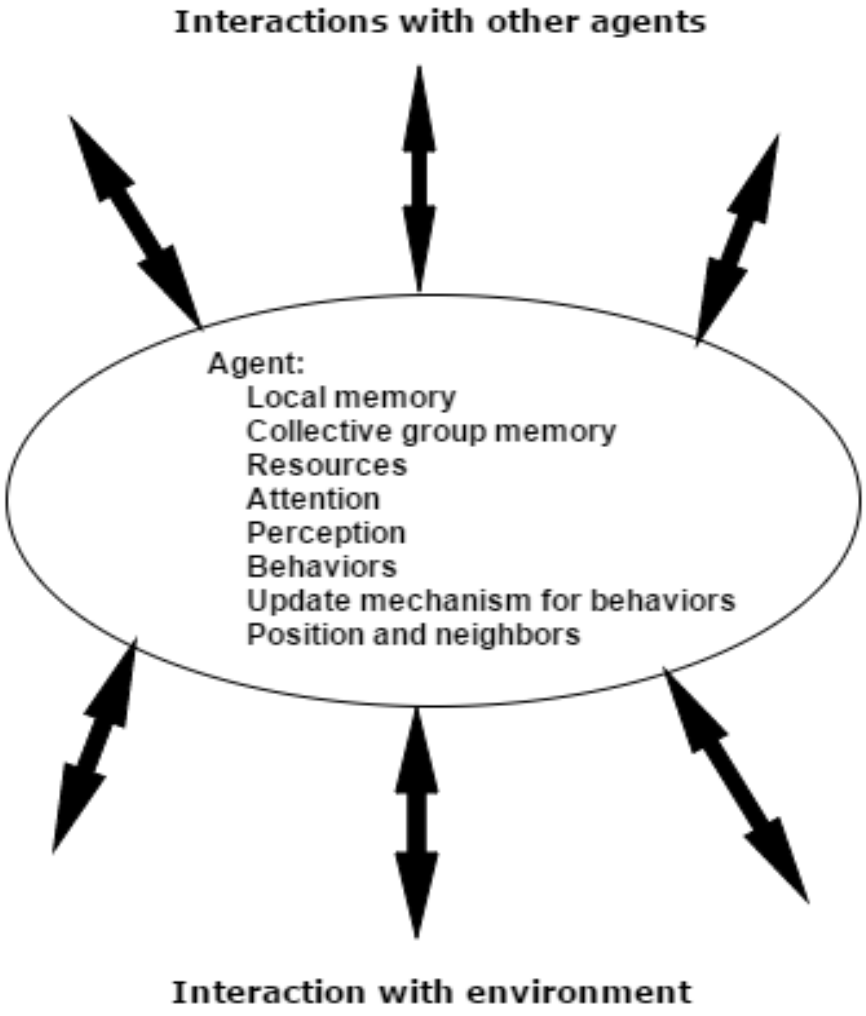


Figure 1: A general representation of an agent.

Optionally, agents can evolve within the environment via the standard evolutionary mechanisms like cross-over and mutation. This leads to additional complexity of the ABM and augment it with Evolutionary Computations (EC). This requires the need to additionally define rules for parent selection, mating, and mutation. These rules are not part of the internal agent model, but are imposed from outside. Therefore the system dynamics will not be the result only of the interaction rules defined by through the agents, which makes the models go outside the pure ABM.

To perform a simulation with ABM means to let agents repeatedly interact with their neighbors and behave within the environment. For more details and brief introduction see [Macal & North, 2010].

ABM Drawbacks

ABM simulations are a valuable research tool and add value and insights to the purely formal theoretical models of the real world. Several software toolkits have made the usage ABM relatively easy to be attractive for the scientific researchers. The usage of these toolkits allows researchers to concentrate on the modeling, as the infrastructure and reliability of the tools are created and optimized by specialists in the respective fields, mainly software developers. However they come with their drawbacks and limitations.

On one hand, the demands and formalism for creating such toolkits are not well established and can vary in their demands and complexity. Because of this, designers usually target specific use cases and problems, although trying to keep their approach as general and conceptual as possible, which leads to some limitations. The creation of new types of models that require different agent communication and interaction topologies, together with more sophisticated techniques for behavior modeling (e.g. artificial intelligence, deep learning), proved it to be not so simple and often require better programming skills and deep knowledge of the framework in order to be done. Often it turns out that a simulation can be done more easily from scratch rather than using an existing ABM platform. Another main issue is the computational cost in terms of time and memory needed to run large scale simulations. Moreover, if the behavior of the agents or/and the environment are probabilistic, several runs have to be carried out in order to have sufficient statistical confidence in the results [Barnes & Chu, 2015].

Additionally, more complicated and intricate agents and scenarios require more complex simulation models. The presence of multiple factors that need to be taken into account and the complexity of the resulting behavior to be analyzed afterwards increases the computational resource needed. Similar problems arise when the number of agents are very large, i.e. millions and billions [Parker & Epstein, 2011].

ABM Frameworks and Technologies

In terms of implementation the ABM can be regarded as a concept similar to the object-oriented programming (OOP). Similar to an object an agent can have internal state or memory and behavior. An agent, however, can undertake actions or learn and change its own state or the state of the environment while an object is more constraint.

Although related the Multi-Agent Simulations or Systems (MAS) and ABM are not the same. ABM tries to understand the emerging phenomena as a result of the behaviors of individual agents. Taking the examples with GT, given in the preceding sections, most of the game-theoretic models consider decision making at or near equilibrium, while ABM can consider complex and interesting behavior far from equilibrium. MAS focuses much more on solving practical and engineering problems [Parker & Epstein, 2011]. Another subset of objects that exhibits similarities to ABM is the actor-based modeling that is on top of the idea of messaging systems where interactions are based on immutable messages sent between actors. However this paradigm is motivated by multi-threading engineering problems and not by ABM although it seems very close to and useful for MAS [Suereth, 2012].

Some of the most used platforms are JADE[1], Repast[2] and Mason[3]. JADE is a Java-based framework designed to simplify the implementation of multi-agent systems. Although very powerful and well-constructed, it is designed according to the FIPA standards and does not allow too much freedom when designing a model. Because of this, it can be regarded not so much of an ABM framework but more like MAS software. However, if the model is in accord with FIPA the framework gives a stable and distributable environment that can handle use cases requiring excessive needs of computational power.

In terms of model definition Repast and Mason provide almost overlapping functionalities. Both allow the definition of multi-dimensional space where agents can live and interact. Cases where agents can be positioned into a subset of dimensions or topologies within an environment, and where agents can have different types of communications according to concrete dimension, may be tricky or even impossible to implement. Moreover, complicated scenarios require specific knowledge of the frameworks and software development skills. Evolutionary computations are supported mainly on intra-agent level through JGAP or ECJ respectively.

1:JADE: http://jade.tilab.com

2:Repast: http://repast.sourceforge.net

3:Mason: http://cs.gmu.edu/~eclab/projects/mason

In terms of large-scale simulation support there are two main solutions: parallelization and distribution of processes. Repast supports the first through another framework - Repast HPC, but this requires rewriting the model from Java to C++, and a good understanding of parallel operations' implementation. The distribution is offered through additional frameworks for distributed data like GridGain [1]that have a certain level of integration within Repast [Repast Symphony, 2006]. However their usage requires a good knowledge of them. Mason on its term provides multi-threading, but it is defined for advanced users [Luke, 2015] and require additional knowledge for multithreading and concurrency within the Java framework and more specifically the MASON implementation. MASON states [Luke, 2015] states that it can support millions of agents when no user interface is defined, but there are no marks on the memory consumption. The distribution of the simulation on cluster of servers is achievable through additional framework D-MASON [2]that may still require an adaptation from the user.

All three of JADE, MASON and Repast, as well as many other ABM frameworks, show a great level of advancement and provide a lot of functionalities, often overlapping, and allowing researchers to develop sophisticated simulations. Although they are relatively easy to use and therefore attractive for the scientific researchers, more complicated cases executed with parallelized runs on distributed machines, are difficult if not impossible to implement. They can require both integrated usage of several toolkits and additional software development knowledge and skills, shifting time and attention of researchers away from their domain problems.

CASPer Framework Overview

The above mentioned uses cases are not supported out-of-the-box from the currently most used platforms. Future research may also lead to the need to define new uses case that may also not be supported. Therefore a new architecture for ABM is proposed – Cognitive Agent-based Simulation Platform, briefly “CASPer”. In what follows an overview of functionalities is given. Then a general design and implementation details are presented. An example of two agents that are positioned in three types of environments is presented on Figure 2.

[1] GridGain: http://www.gridgain.com

[2] D-Mason: https://sites.google.com/site/distributedmason

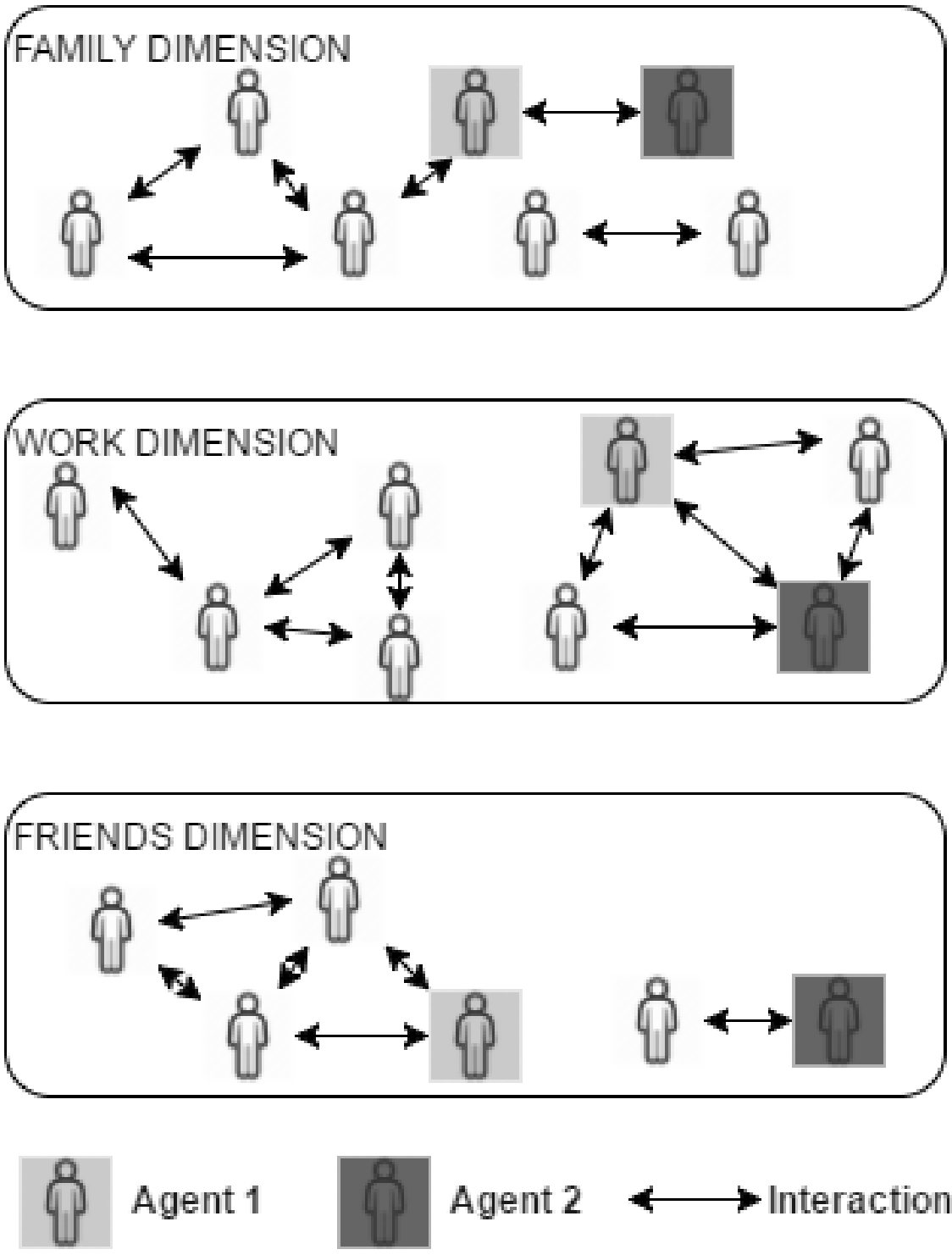


Figure 2: In the given example 2 agents that are positioned in 3 types of environments. In the FAMILY both agents are part of the same group. In the WORK they also make part of the same group, but in the FRIENDS dimension they are in separated groups. Each interaction represents a game played between agents, and in each environment there can be different type of games played.

On top of the basic ABM functionality presented above the CASPer framework aims to address the following use cases:

1. Agents can have access for read/write operations over a collective memory that is shared with other agents according to a grouping criteria.
2. Each agent is allowed to inhabit a subset or all of the dimensions available within an environment. Different agents may inhabit different subsets of dimensions.
3. Agents can be in multiple groups that are formed according to different criteria in one or more dimension. A simple example is to have an agent that is a member of a family (family group). In the same time he/she is an employee (work group), and keep close contacts with childhood friend (friend group). This agent can work with his sister, therefore in some of the groups he belongs to, he/she may see the already familiar agents.
4. Agents are supposed to have different types of communication within different groups. This means that the agent can play Prisoners' Dilemma at work group and Battle of the sexes in the family group.
5. Agents can have complex decision making models motivated by deep learning, neural nets, machine learning or artificial intelligence.
6. Also the framework is meant to support large-scale simulations. The aim here is to support millions and even billions of interacting agents in order to simulate the whole Earth human population or even more.

CASPer Implementation

In terms of usability recent trends in software development [Tulach, 2008; Martin, 2009] are oriented towards hiding implementation details from users and letting them add only their domain specific customization. This means hiding the multi-threading parallel execution on distributed machines from users and letting them implement only the logic of their domain specific model.

Following software design best practices [Reinhold, 2016] separation and encapsulation of different functionalities into modules, so that they can be easily removed or replaced, is favored instead of having one major component that rules everything.

The basic idea behind the framework consist of making a few interfaces available to users, where they can customize the simulation steps according to their specific needs. All the rest – parallel execution, distribution over machines and search algorithms for spaces is provided on the back side. Users should have no need to perform specific installations. To realize this functionalities are hidden behind interfaces. Moreover they are separated into different modules allowing an easy mechanism for

replacement or updating these functionalities, as only a new implementation of the corresponding interface is required.

The basic outline of the framework consist of an engine of type ISimualtionEngine that takes care of the algorithm's execution steps. The other parts of the public API are:

An IStateSpace instance take care of the environment dimension, placing the agents inside, and perform the search and group actions. There are two implementation - DefaultStateSpace and HazelcatsStateSpace, but the interface is extensible and allows implementations with other dimensions and search engines (e.g. ElasticSearch, graph data bases etc.).

IEnvironmetDimension describes a particular dimension where agents can be positioned. The default implementations are for discrete, continuous and string representation. The interface implementation requires also a meaningful way to search within it, and a may require new IStateSpace and IRepositioning interfaces.

ICollectiveMemory is a data holder object that can be positioned inside one or more IEnvironmetDimensions. It allows read and write by IAgent that are in his range (i.e. belongs to the same group, society or island).

An IGame instance take care about the communication between agents. A default implementation gives the possible moves and the payoff matrix for them. The interface can be implemented to allow other types of communication not based on game theory.

IAgentMemory instance represents the memory of the agent. The instance can hold other IAgentMemory instances dedicated to concrete tasks. A subtype is IGameStrategyMemory that can hold the memory for specific type of IGame. It can be used in ICognitiveModel. IAgent instances represents an agent in the environment. It has an IAgentMemory, ICognitiveModel and access to ICollectiveMemory.

ICognitiveModel implementations provide the IAgent's behavior logic. It has access to IAgentMemory, ICollectiveMemorys that are visible to the IAgent. The implementation allows to have different behavior depending on who started the communication game. The implementation of ICognitiveModel can be easily made to access 3rd party libraries for machine learning or AI (Deep learning, neural networks, recurrent learning, etc.).

IFitnessFunction implementations gives the possibility to make an evaluation of the IAgent according to criteria. The function will be called when the IAgent has played all games within a particular cycle or epoch. The interface has access to IAgentMemory, ICollectiveMemorys and IEnvironemntalArtifact that are visible to the IAgent. The result is written in the memory of the agent and in the IStateSpaceContext and can be used for repositioning or the agents within the space later on.

IRepositioning implementation allows the optional repositioning of an IAgent instance. It has access to the IStateSpaceContext and the IAgent's memory. ITerminationCondition implementation states when the simulation can terminate. Examples are number of cycles or fitness gained by an IAgent. It has access to the IStateSpaceContext. Asynchrounous events can be sent on each step in order to monitor and save the simulation data for further analysis. Advanced users are allowed to modify easily all hidden parts of the implementation.

The parallel run is supported by an ISimulatoinEngine. Whenever a different approach is requested (e.g. using sequential single-thread implementation), the only things that need to provided is the respective implementation of the interface.

The distribution over a cluster of machines is provided by the Hazelcast [1]framework. It provides the support for the clustering, distributed data structures and locking mechanism for accessing and modifying objects. Similar to the parallelization there is an interface to provide access and locking to the data structure that can be re-implemented with other frameworks.

The different types of dimensions or topologies currently follow custom implementation, but can be extended to use sophisticated engines that provide better support for specific topologies like ElasticSearch[2], Neo4j[3], OrientDB[4], etc. The only thing that needs to be done is implementing the IEnvironmentDimension.

Conclusion

In the paper a novel framework for modeling social interaction has been presented with emphasis on the general approach and the first step of the implementation of a multi-agent environment. The general approach consists in combining artificial agents that instantiate cognitive models and a simplified social environment of experimental games like Prisoner's dilemma, Stag Hunt, Chicken, Ultimatum game, and other games used to model interesting social interactions. This simplification is regarded as a reasonable trade-off between the complexity of the expected emergent phenomena and the possibility to perform analyses using game theoretic approaches and social network theory.

[1] Hazelcast: http://hazelcast.com

[2] ElasticSearch: https://www.elastic.co

[3] Neo4j: https://neo4j.com

[4] OrientDB: http://orientdb.com/orientdb

The presented platform – CASPer – combines the complexity of the cognitive agents with the concept of a multi-dimensional agent space specified by the characteristics and history of the agents.

The incorporated constructs and mechanisms allow the agents to be efficiently monitored as part of more than one group with specific behavior and communication for each the interaction within each of the groups they may belong to permanently or temporarily. This is achieved by implementing an agent related and an environment related memories that can be used to model dynamically the state of the agents and the environment.

The framework takes care of the usage of multi-core machines and clusters of servers when available and needed. The cluster support will make possible large scale simulations with possibly billions of agents. The software design of the framework allows and promotes the usage of third party tools and modules for enhancement and customization of different parts of the simulations or adding new functionalities.

Acknowledgments

The paper is published with partial support by the ITHEA ISS (www.ithea.org), the ADUIS (www.aduis.com.ua), and the Central Fund for Strategic Development of the New Bulgarian University.

References

[Acevedo & Krueger, 2005] Acevedo, M., & Krueger, J. I. Evidential reasoning in the prisoner's dilemma. The American Journal of Psychology, 431–457., 2005.

[Adami et al., 2015] Adami, C., Schossau, J., & Hintze, A. Evolutionary Game Theory using Agent-based Methods. arXiv Preprint arXiv:1404.0994., 2015.

[Axelrod, 1984] Axelrod, R. The Evolution of Cooperation. New York: Basic Books, 1984.

[Axelrod & Hamilton, 1981] Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. Science, 211(4489), 1390–1396., 1981.

[Axelrod, 1997] Axelrod, R. The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration. Princeton University Press, 1997.

[Barnes, 2015] Barnes, Chu. Guide to simulation and modeling for biosciences, 2nd edition, Springer-Verlag London 2010, 2015.

[Binmore, 2007] Binmore, K. G. Playing for real: a text on game theory. Oxford ; New York: Oxford University Press, 2007.

[Camerer, 2003] Camerer, C. Behavioral game theory: experiments on strategic interaction, 2003.

[Camerer, Ho, & Chong, 2002] Camerer, C. F., Ho, T.-H., & Chong, J.-K. Sophisticated Experience-Weighted Attraction Learning and Strategic Teaching in Repeated Games. Journal of Economic Theory, 104(1), 137–188, 2002.

[Camerer, Ho & Chong, 2004] Camerer, C. F., Ho, T.-H., & Chong, J.-K. A cognitive hierarchy model of games. The Quarterly Journal of Economics, 861–898., 2004.

[Chater, 2015] Chater, N. Can cognitive science create a cognitive economics? Cognition, 135, 52–55., 2015.

[Colman, 1995] Colman, A. M. Game theory and its applications in the social and biological sciences (2nd edition). Oxford: Butterworth-Heinemann Ltd., 1995.

[Colman et al., 2014] Colman, A. M., Pulford, B. D., & Lawrence, C. L. Explaining strategic coordination: Cognitive hierarchy theory, strong Stackelberg reasoning, and team reasoning. Decision, 1(1), 35–58., 2014.

[Erev & Roth, 1999] Erev, I., & Roth, A. On the role of reinforcement learning in experimental games: the cognitive game-theoretic approach. In D. Budescue, I. I. Erev, & R. Zwick (Eds.), Games and human behavior: Essays in honor of Amnon Rapoport. Mahwah, NJ: Lawrence Erlbaum., 1999.

[Epstein, 2006] Epstein, J. M. Generative Social Science: Studies in Agent-Based Computational Modeling. Princeton University Press, 2006.

[Epstein, 2014] Epstein, J. M. Agent_Zero: Toward Neurocognitive Foundations for Generative Social Science. Princeton University Press, 2014.

[Fehr & Fischbacher, 2004] Fehr, E., & Fischbacher, U. Social norms and human cooperation. Trends in Cognitive Sciences, 8(4), 185–190., 2004.

[Fehr & Schmidt, 2006] Fehr, E., & Schmidt, K. M. The economics of fairness, reciprocity and altruism–experimental evidence and new theories. Handbook of the Economics of Giving, Altruism and Reciprocity, 1, 615–691, 2006.

[Fiske, 1992] Fiske, A. The Four Elementary Forms of Sociality: Framework for a Unified Theory of Social Relations. Psychological Review, 99(4), 689–723., 1992.

[Fiske & Haslam, 1996] Fiske, A., & Haslam, N. Social Cognition is Thinking about Relationships. Current Directions in Psychological Science, 5(5), 143–148, 1996.

[Fiske, 2012] Fiske, A. P. Metarelational models: Configurations of social relationships. European Journal of Social Psychology, 42(1), 2–18., 2012.

[Geanakoplos, Pearce & Stacchetti, 1989] Geanakoplos, J., Pearce, D., & Stacchetti, E. Psychological games and sequential rationality. Games and Economic Behavior, 1(1), 60–79., 1989.

[Gigerenzer & Goldstein, 1996] Gigerenzer, G., & Goldstein, D. G. Reasoning the fast and frugal way: models of bounded rationality. Psychological Review, 103(4), 650, 1996.

[Gintis, 2009] Gintis, H. Game theory evolving: a problem-centered introduction to modeling strategic interaction (2nd ed.). Princeton: Princeton University Press, 2009.

[Grinberg, 2011] Grinberg, M. A cognitive approach to web-based intelligent agents: The TRIPLE architecture. Web Intelligence and Agent Systems, 9(1), 69–79, 2011.

[Grinberg, Hristova & Borisova, 2012] Grinberg, M., Hristova, E., & Borisova, M. Cooperation in Prisoner's Dilemma Game: Influence of Social Relations. Proceedings of CogSci, 2012.

[Grinberg, Hristova & Lalev, 2010] Grinberg, M., Hristova, E., & Lalev, E. Models for cooperative decisions in prisoner's dilemma. Advances in Cognitive Systems. IET: London, 2010.

[Grinberg & Lalev, 2008] Grinberg, M., & Lalev, E. Anticipation in Coordination. In The Challenge of Anticipation (pp. 215–236). Springer, 2008.

[Hristova et al., (2013)] Hristova, E., Grinberg, M., Georgieva, I., & Borisova, M. Cooperation in Prisoner's Dilemma Game: Influence of Players' Social Roles. In Proceedings of the 35th Annual Meeting of the Cognitive Science Society (pp. 2578–2583), 2013.

[Krueger, 2014] Krueger, J. I. Heuristic game theory. Decision, 1(1), 59–61, 2014.

[Luke, 2015] Luke, S. Multiagent Simulation and the MASON Library, 2015.

[Macal & North, 2010] Macal, C.M., North, M.J. Tutorial on agent-based modelling and simulation. Journal of Simulation 4, 151–162, 2010.

[Macy & Flache, 2002] Macy, M. W., & Flache, A. Learning Dynamics in Social Dilemmas. PNAS, (99), 7229–7236, 2002.

[Nowak, 2006] Nowak, M.A. Evolutionary dynamics _ exploring the equations of life. Belknap Press of Harvard University Press, 2006.

[Martin, 2009] Martin, R. Clean Code. Pearson Editions, Boston, 2009.

[Maynard Smith, 1982] Maynard Smith, J. Evolution and the theory of games. Cambridge ; New York: Cambridge University Press, 1982.

[Niazi, & Hussain, 2011] Niazi, M., & Hussain, A. Agent-based Computing from Multi-agent Systems to Agent-Based Models: A Visual Survey. Scientometrics (Springer) 89 (2): 479–499, 2011.

[Parker & Epstein, 2011] Parker, J., & Epstein. J.M. A Distributed Platform for Global-Scale Agent-Based Models of Disease Transmission. ACM Transactions on Modeling and Computer Simulation. Vol. 22, Issue 1, 2011.

[Reinhold, 2016] Reinhold, M. The State of the Module System, 2016.

[Repast Symphony, 2006] Repast Symphony. Repast Reference, 2006. Retrieved from: http://repast.sourceforge.net/docs/RepastReference.pdf.

[Shoham & Leyton-Brown, 2008] Shoham, Y., & Leyton-Brown, K. Multiagent systems: Algorithmic, game-theoretic, and logical foundations. Cambridge University Press, 2008.

[Suereth, 2012] Suereth, J. Scala in Depth. Mining publications, New York, 2012.

[Sun, 2006] Sun, R. (Ed.). Cognition and Multi-agent Interaction: From Cognitive Modelling to Social Simulation. Cambridge University Press, 2006.

[Taiji & Ikegami, 1999] Taiji, M., & Ikegami, T. Dynamics of internal models in game players. Physica D: Nonlinear Phenomena, 134(2), 253–266, 1999.

[Tulach, 2008] Tulach, J. Practical API Design. Apress, New York, 2008.

[Van Lange & Rusbult, 2011] Van Lange, P. A., & Rusbult, C. E. Interdependence theory. Handbook of Theories of Social Psychology, 2, 251–272, 2011.

[Van Lange et al., 2013] Van Lange, P., Balliet, D. P., Parks, C. D., & Van Vugt, M. Social dilemmas: Understanding human cooperation. Oxford University Press, 2013.

[Von Neumann & Morgenstern, 2007] Von Neumann, J., & Morgenstern, O. Theory of Games and Economic Behaviour (60th Anniversary Commemorative edition). Princeton, NJ: Princeton University Press, 2007.

[West, Lebiere & Bothell, 2006] West, R. L., Lebiere, C., & Bothell, D. J. Cognitive architectures, game playing, and human evolution. Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation, 103–123, 2006.

Authors' Information

Maurice Grinberg *– Research Center for Cognitive Science, Department of Cognitive Science and Psychology, New Bulgarian University, 1618 Sofia, Bulgaria; e-mail: mgrinberg@nbu.bg*

Major Fields of Scientific Research*: decision making, cognitive modeling, experimental game theory.*

Emiliyan Todorov, e-mail: *emo_todorov@mail.bg*

Major Fields of Scientific Research*: data analysis, decision making.*

PROCESSING DIFFERENCES BETWEEN NEAR AND FAR ANALOGIES

Alexandra Alexieva, Penka Hristova

Abstract: *It has been shown, that far analogies do engage more cognitive resources than near analogies. Word couples of the A goes to B as C goes to D (A:B :: C:D) kind are mapped slower and more mistakes appear if comprised of far (Tongue:Taste :: Antenna:Signal), rather than near analogies (Tounge:Taste :: Nose:Smell). According to Adam Green, that is because in the Working Memory (WM) there will be too much information when it comes to comparing the far analogies.*

*That explanation was tested in the current paper. The presentation of elements, comprising simple verbal analogies were varied at four levels: no information is shown to the entry; element A is only visible at the entrance; A:B are shown and the first three elements are visible A:B :: C***:.** *If mapping of far analogies truly requires more cognitive resources to be available as the results of Green's study indicate, the time for the encoding of the three elements at the entrance, shall be longer for far compared to near analogies, since the mapping between the pairs can be initiated under that condition. Such distinction is not to be expected when it comes to the other three (0; A:; A:B) conditions.*

The collected results show that near analogies are addressed more accurate and faster that the far ones. The way of presentation affects the participant's time for response, but not how accurate, they might be. Results for the encoding type reveal a main effect for way of presentation, but none for type of analogies and insignificant interactions between them. The mapping of far compared to near analogies was found to be not significantly slower, although this was suggested to be the cognitively demanding part for the far compared to near analogies.

Keywords: *analogies, near, far, encoding, mapping, evaluation*

Introduction

Analogy allows comparison between featurally unlike structures to be performed on the basis of relational similarity (Holyoak, 2012). Analogical mapping allows finding of systematic correspondences between the source and target situation based on shared relations (Tohill & Holyoak, 2000). People seek to put the objects of the source in one-to-one correspondence with the objects of the target, so as to obtain the highest structural match. The corresponding objects in the base and target need not resemble each other, the importance here is that, they hold alike roles in the matching relational

structures. A well-understood base or source situation is mapped to a target situation that is less familiar and/or less concrete. Then new inferences are derived by importing connected information from the source to the target. A good example would be the following: in the analogy between blood circulation and plumbing, people might first align the known facts that the pump *causes* water *to flow through* the pipes with the fact that the heart (being a unique, alive pump) is the prime *cause* which moves the blood to *pour through* the veins. When comparing the heart with a pump, we are practically saying: *"Heart is to blood, as pump is to water."* Afterwards blood returns to the heart to be re-cycled and re-poured again and again, creating an ideal, closed system cycle (Gentner, 1982).

According to the *Structure Mapping Theory,* analogy-making requires first a selection of the primary candidate number of predicates to map attends only to the *structure* of the knowledge representations for the two analogs, and not the content. The mapping takes place not only between objects, but also between relations among objects (i.e. first-order predicates) and between relations of relations (i.e. second-order predicates). The whole mapping yields the assignment of a predicate or a relation to the target.

Gentner (1982), proposed that in order to facilitate the making of clear alignments and reasonable inferences, an analogy must be structurally consistent, meaning: it should have one-to-one *correspondences*, and the relations in the two domains should have a parallel structure. With respects to the circulation/plumbing system analogy, the pump cannot correspond at the same time to the veins and the heart. Another factor influencing the quality of it, is *systematicity*: Analogies that convey an interconnected system of relations, such as the circulation/pumping analogy, are more useful than those that convey only a single isolated fact, such as ''*The brain looks like a walnut*.''

Near and Far Analogies

When a new highway system is about to be designed, the devisors could draw a near analogy – another highway system from another city, which in its essence is deriving an analogy from a closely related base domain. On the other hand, a more distant domain that can act as a base, say the human blood circulatory system. This distinction is important, since mapping and transferring elements (objects, attributes, relations) are different with respects to both analogy types (Dahl & Moreau, 2002). When near analogies are drawn, both surface-level attributes (roads) and the relations between the attributes (the course of automobiles through the highway) are mapped and transferred. When it comes to far analogies, however, much less surface-level attributes can be mapped, and thus leaving the mappings to occur between common relations. When the base and target realms pose few such similarities, the process itself becomes limited and uneasy to deal with (Dahl & Moreau, 2002).

The difference between near and far analogies can be defined, as analogies being derived from semantically afar fields, and are also referred to as within-domain (near) and cross-domain (far) analogies (Green, Fugelsang, Kraemer & Dunbar 2008; Green, Kraemer, Fungelsan, Grey & Dunbar, 2010). Green et al (2006a; 2008; 2010), show an apparent relationship between semantic distance of analogical mapping and proportional recruitment of the frontopolar cortex. The fMRI data has shown that the frontopolar cortex is recruited more strongly for the mapping of cross-domain (i.e. Nose:Scent :: Antenna:Signal) analogies than within-domain (i.e. Nose:Scent :: Tongue:Taste). Moreover, frontopolar activity largely reflects a taxonomic (etymologically: "law of order of arrangements", nowadays known simply as a science of classification of things or concepts) distinction, between these classes of analogy (Green et al, 2010). Semantic distance values were used as parametric regressor, which allowed regions of the brain, where semantic distance collaborated with stimulus-related activity in a relationship to be identified. Furthermore, even after these measures of task difficulty were excluded, the semantic distance still covaried positively with activity in the frontopolar ROI. These results highly suggest that difficulty-related factors cannot explain the affinity between semantic distance and frontopolar activity and are also consistent with previous work, indicating that specific task arrogations of multiplex reasoning, rather than time-on-task or difficulty by its very nature, accounts for frontopolar enrolment (Green et al, 2010). Further, the results implicate a degree of frontopolar activation as a marker of semantic distance in analogical reasoning. When the cortical activity is to be increased, proportionally it could reflect in increasing the fixed calculational demand on the neuronal integration network (Green et al, 2010). Relations between AB and CD should be represented in much abstract form during analogical mapping of far analogies but not necessarily during near analogies. Also mapping between far compared to near analogies is more likely to be ambiguous (i.e. enabling several alternative mappings between AB and CD pair), since cross-domain analogies rely on mapping of more abstract relation(s) than within-domain analogies. Henceforth, as Green et al (2010), discussed cross-domain analogies may require harder evaluation, which was also supported from the involvement of left-sided inferior frontal gyrus.

Sub-processes of Analogies

The process of analogy making can be broken into five, slightly different sub-processes (Kokinov, & French, 2003). They are *recognition* (building a representation), *retrieval* (re-captures the 'base' for the analogy), *mapping* (the 'base' to the target), *transferring* (re-location of the unmapped elements from the 'base' to the 'target') and *evaluating* (certifying if the inferences are suitable).

Recognition is about how context-appropriate and adjustable representations are build, i.e. when describing a target of a source. The *retrieval* process concerns recapturing of similar features, structures

or objects from the base to the target domain. The retrieval concerns also relations. *Mapping* is about mapping mechanisms, when, exploring and finding which elementals of the base do correspond to those of the target (Kokinov & French, 2003). This one is definitely a critical sub-process for the analogy-making (Gentner, 1983). *Transferring* shows how the new knowledge can be in truth implemented into the target. According to some researchers, this sub-process is a part of the mapping henceforth, it is not recognized as an independent one. The *evaluation* sub-process, on the other hand, is a one which has been least investigated upon. It regards the confidence strength, meaning, when one has more or less second thoughts or doubts, establishing the possibility of the transferred knowledge to be suitable for the target domain. Not-rarely at all, this process is assigned to either transfer or mapping. The latter is compatible with the Green's et al (2010) suggestion that cross-domain analogies require more relational intergration, eanbling the mapping and evaluation than the within domain analogies.

One way to differentiate mapping from evaluation is provided by the the 4-level model of A:B::C:D analogy presentation of Sternberg (1977; 1996):

a) Zero information at the entrance (called C-0 cue phase);

b) 1 element is shown **A:**; encoding of 1 element (cue phase C-1);

c) 2 elements are visible **A:B**; encoding of 2 elements and 1 relation (cue phase C-2);

d) 3 are visible **A:B :: C**; 3 elements, and 2 relations (C-3 cue phase);

Sternberg believed that during each cue phase the participant would be able to complete a bit of the component processes, i.e. in C-1 one encodes one element; C-2 one encodes 2 elements and could make a relational inference. The last condition (i.e. C-3), is actually, where the mapping starts, i.e., presentation of C**:** element marks the beginning of the mapping. Sternberg was confident that the components are "*independent and additive*": presentation of **A:B** requires encoding and presentation of A:B::C: need encoding *and* mapping. The C-3 cue phase may enable mapping between the domains, but not evaluation, since the D term is not present and analogical mapping cannot be acompliched, hence, evaluated.

Experiment

Earlier in the text it is shown that Green et al, (2010) go into research of the verbal pair word analogies of the type A goes to B as C goes to D (A:B :: C:D). In the research, cross-domain (far) and within-domain (near) analogies are the subject of interest with an important difference being added. That is the introduction of neuroimaging, regarding the brain areas, that could be affiliated with both analogical types. An example for far analogies, would be Tongue:Taste :: Antenna:Signal. Taste is an abstract,

specific function of the tongue, and antenna is a receiver-giver of signals of any kind. The relations between each of the two couples are specific, and are from different fields. On the other hand, an example for near analogies would be “Kitten:Cat :: Puppy:Dog”, where the relation is a close one. As the fact that they are derived from the same area (cat and a dog are both animals, which can and are, domesticated; and kitten and puppy are their small ones, baby versions), that is why they are called within-domain.

Green et al, (2010) used both types of analogies to try to find or allocate which areas of the brain are more sensitively activated. That is when participants are facing these types of analogies, and which areas are more dormant. Cross-domain (far ones) analogies could require more attention and cognitive reserves, than the within domain ones, they are going to be recognized slower and with much more incorrect answers. That is because the Working Memory (WM) will be more on high alert, needing to call, remember and compare the amount of information that came at the surface, and all needs be simultaneous. That is why, usually, near analogies are easier to deal with, as in to be recognized and pointed out, than far ones (Green et al, 2006a; 2008, 2010). Henceforth, mapping but not recognition of relations is difficult, meaning that if participants start to map relations across domains they should dwell more.

Therefore, in this case, a difference can be anticipated concerning the encoding time for condition four (A:B :: C:) as per Sternberg’s method, compared to the first three conditions (O:; A:; A:B;). This is anticipated, because in this condition, the coding of the second word couples begins and apposition as a process is starting. This is the condition where all that is necessary for the task to be more challenging and burdensome for the participants is visible. Logically, a longer encoding time for this condition should be registered, when it is compared to the other three conditions, but this time should be even longer for the far compared to near analogies. While Green et al (2010) study is of neuroimaging nature, this is a simple behavioral one and uses the 4-level method of presentation of Sternberg (1977), whereas Green shows to the participants, the entire analogical pair at once.

Goals

To check whether analogical mapping poses difficulties for far compared to near analogies. The aim of the experiment is to test if the encoding time for the first three elements, i.e. A:B :: C: will be longer for far analogies than near ones. Since, there are all necessary prerequisites for the mapping process to begin, such result is anticipated.

Hypothesises

Encoding time for coding the three elements from the far analogies (i.e. A:B :: C:;) will be longer than the same operation with near analogies. Such difference is not to be expected when it comes to the other, first three conditions, where 1. Zero, 2. A:, 3. A:B; are being displayed.

Method

Design

The design was a 2x4, within-subject one.

The Independent Variables are two: type of analogy and way of presentation, with different sub-levels.

The First Independent Variable has two levels, according to which the analogies are either near (within-domains) or far (cross-domains), being derived from different semantic areas.

The Second Independent Variable is the type of presentation, of the analogical word-pair couples, presented through the method of Sternberg (1977). The method was already discussed in the beginning, but shortly again, it is: a) Zero information at the entrance; b) 1 element is shown A:; c) 2 elements are visible A:B; d) 3 are visible A:B :: C:. After that the participants clicks “next” so the whole analogical word-couple to be shown.

The Dependent Variables are Encoding time, which marks the time from which the specific elemental is shown (0; A:; A:B; A:B :: C:;) at the entry, until the very instant when the button “next” is being pressed. Accuracy (includes correct answers) and Response time (how fast the participant indicates if the word-pair couples are analogical or not, measured from presing next button, after the initial presentation of 0/A/AB/ABC elements and until response) were also measured.

Participants

All together 95 participants (22 males and 73 females) with average age for males is M = 24, and for females is M = 23.5) took part in the experiment. Subsequently, almost all of them were students from the New Bulgarian University and participated either for credits or just on good will. All of them were right handed and Bulgarian native speakers.

Stimuli

Two pre-tests with word-pair couples were done and were later used in the main experiment. The original used stimuli, in Green et al's (2010), research, were translated into Bulgarian, and modified where it was necessary, so that participants could understand and make their decisions upon them. The stimuli word-couples were then divided into two separated files and a pre-test in semantical diversity and analogical difference, which the participants had to evaluate, were conducted.

The word pairs were a total of 160 as follows: 80 of them *cross* domain and 80 for *within* domain. They were divided into two different files, which the participants had to evaluate. A 7-point scale was used for that purpose, as 1 being "not semantically identical at all" to 7: "entirely semantically identical". Between some of the word-pairs, the difference between Mean and Standart Deviation has variated (from 1 to 3 points). They were excluded from the study. This difference could be explained because of the big spatial difference between the words themselves, as in a complementary word from the within or cross domain. The Mean for the far analogies is M = 4.10 and SD = 1.73, and for the within-domains is M = 5.38 with a SD = 1.48. The difference in the Mean is: 1.28.

An additional questionnaire was made which included only 94 word pairs (47 near and 47 far) of the 160. The question posed this time was about "How analogically to each other are the word pairs bellow*?*". The "*semantic"* was replaced with "*analogical".* The participants were less this time, only 24.

The mean for the analogical test for cross-domains is M = 4.27, with a SD = 1.64; and differed again from the mean for within-domains: M = 5.75, with a SD = 1.33, and this time it was a little higher than the within ones. The difference between the cross-domains Mean for semantic and analogous was small, as was the one between the within-domains. In the end, the final word-pairs were chosen from the analogical pre-test.

An example for near analogy from the experiment is: "Table:Cover :: Floor:Carpet", and for far analogy "Table:Cover :: Nobleman:Mantle".

Each base pair has two targets – one close and one far as in the Green's study (Green et al, 2010). That allows counterbalancing of the near-far analogies across participants, so that all participants could see the same base A:B pairs, but with different near and far targets.

Procedure

Finally, 32 analogical pairs (16 far and 16 near ones), were chosen. Also, 32 false (fillers, non-analogical) pairs were created and added to the experiment (i.e. "Volcano:Lava :: Sleep:Awakening").

Each participant was invited separately into the laboratory, and then inside a small booth with a computer and two chairs inside. Before going indoors, each participant had to fill a paper with consent

which stated that he/she agrees to take part in it and therefore receives one credit point for his/hers seminar's requirements. No phones or devices were allowed inside.

The analogical task itself, was the following: word-pair analogical couples from *the A:B :: C:D* type, in which those pairs were either far or near analogies, and false as a balancing feature were included. They were shown by Sternberg's (1977) 4-level method of presentation. The experimenter gave each participant a quick explanation in addition with a short example, regarding what could they expect from the task itself. The example was the following: "*The cat goes with meowing just as the dog goes with barking*". The incomplete analogical pairs would appear differently, such as, first they might see just element *A***:** which is the "*Cat*", and have to press "next" so that the last part of the word couple is disclosed. It could be that they are presented with the first three elements *A:B :: C:;* which corresponds to "*Cat:Meowing :: Dog:*", etc. After when they have clicked "next" button and the whole analogy is visible, only then could they decide on pushing either "Yes" or "No", evaluating the couple to be true or false. Before starting, the example with the *cat* and the *meowing,* was repeated again to make sure they have understood.

There was a training session before the veritable word-pairs were presented. The analogical task itself was no longer than seven-eight minutes, depending entirely on the participant, which had to do a speeded button press (via an E-Prime button box), to continue forward with the task. Participants were asked to respond as fast and accurate as they could.

The four types, of stimulus presentations can be seen in *Figure 1*, below. Stimulus presentation was randomized for each participant. The stimuli were, additionally counterbalanced for their type of presentation across participants.

Results

Green expects far analogies to be mapped slower with numerous mistakes than the near ones, because they are taken from different semantical fields. His stimuli were used and translated; and worse accuracy and response time is expected to be registered for the far analogies. When it comes to the *encoding* time of 4th condition (A:B :: C:) it should be slower compared to the first three conditions (0:; A:B; A:B :: C), because then the mapping process should start.

When the data was gathered, the formula "M+/- 2SD", was used, and 247 trials out of 3040 for the Encoding time were cut (8.1 %). The afterwards data that remained – correct answers (Accuracy), Response time and Encoding time - was analyzed by mean of condition.

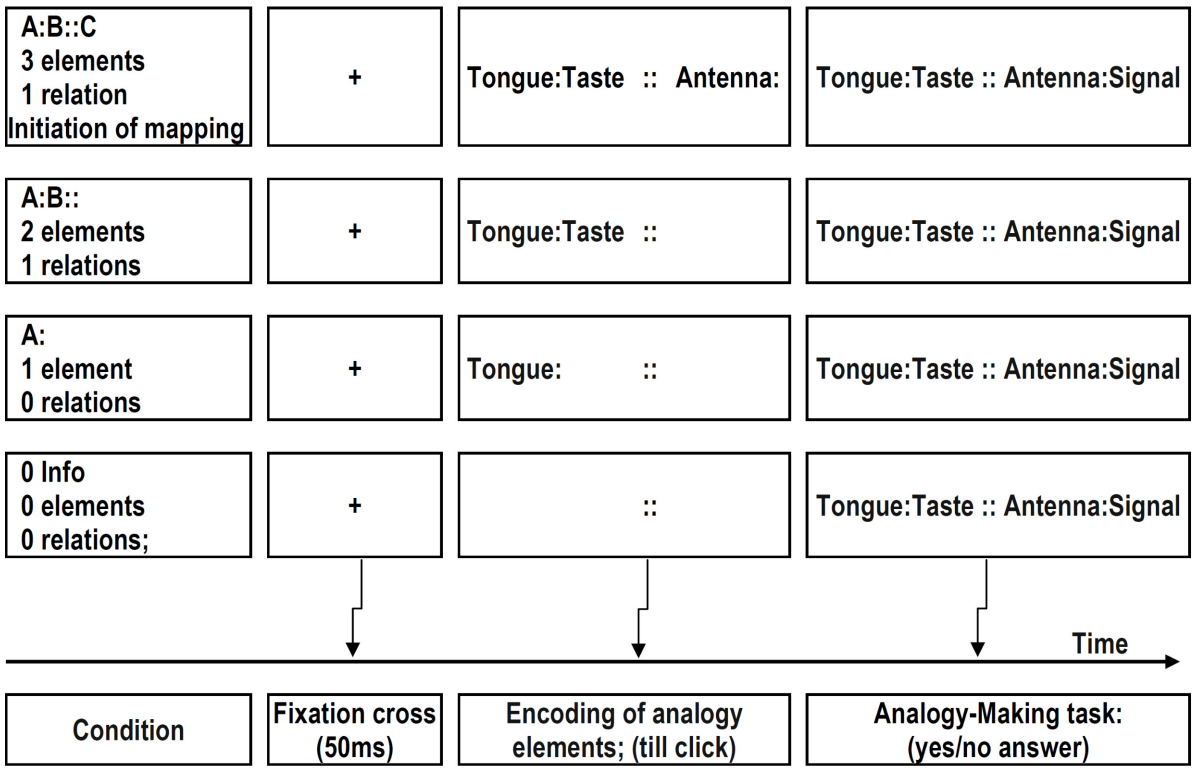


Figure1. Timeline of the analogical task performance

Manipulation check for differences between far and near analogies

As Green et al (2010), hypothesis states, far analogies should be more difficult than near ones, because they are derived from divergent semantic areas, they would require more cognitive reserves, since higher attention is to be required and for the Working Memory's function, too. The below collected results support that accuracy and response time for far analogies do worsens than for near ones, and supports Green's findings.

A 2 (near and far analogies) x 4 (presentation of *0; A; AB; AB:C:*) Repeated Measure ANOVA analyzed the results for accuracy and response time.

The table below shows the Means, Standart Deviations for accuracy for each condition for far and near analogies.

Table1. Mean and Standart Deviation for accuracy of near and far analogies by condition:

Accuracy by condition	Mean	Std. Deviation	Accuracy by condition	Mean	Std. Deviation
Far_0	.85	.204	Near_0	.97	.089
Far_1	.82	.225	Near_1	.96	.127
Far_2	.82	.211	Near_2	.97	.088
Far_3	.86	.211	Near_3	.95	.121

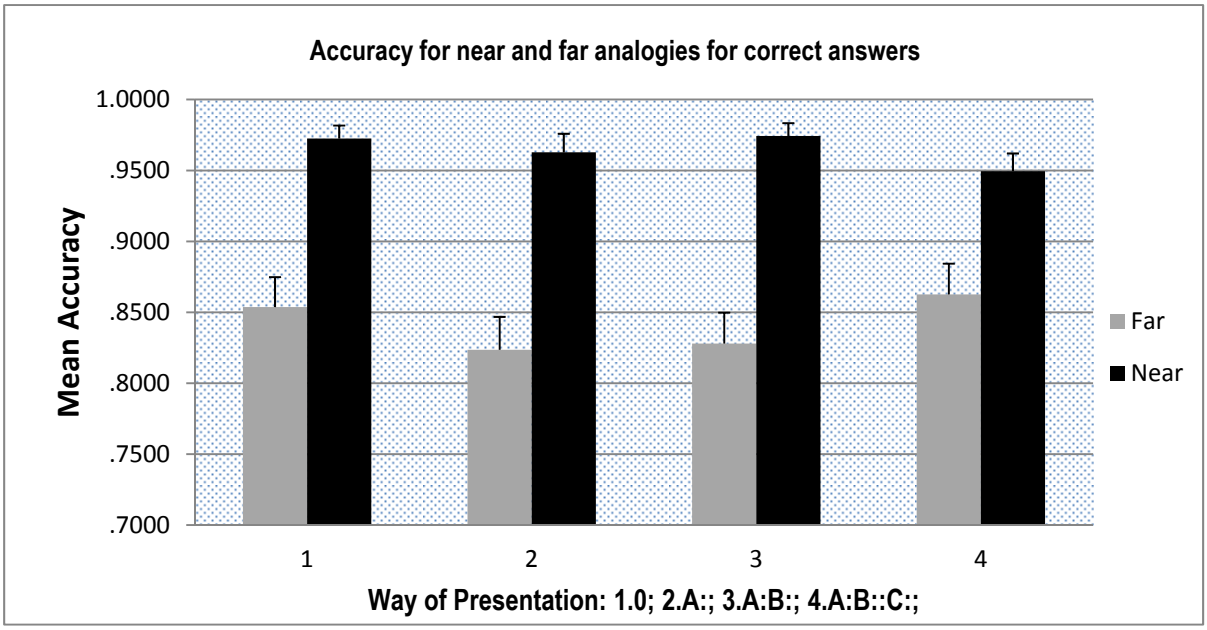


Figure I. It shows the main effects on the type of analogies and the way of presenting them upom accuracy.

Overall, participants were more accurate in identifying the near (Mean accuracy, M = .97) than the far analogies (Mean accuracy, M = .84). Main effect in the analogy type was discovered, being the following: F (1, 93) = 100.383; p = 0.001; and there was no main effect for the way of presentation, F (1, 91) = .441; p = 0.724 (Figure I). Lastly, no interaction between both factors was detected: F (1, 93) = 1.545, p = 0.208.

The Mean and Standart Deviation for Response time for near and far analogies, by condition, is shown, in Table2.

Table2. Mean and Standart Deviation for Response time by condition, for far and near analogies.

RT by condition	Mean (ms)	Std. Deviation	RT by condition	Mean(ms)	Std. Deviation
Far_0	4059.41	1263.29	Near_0	3345.62	1151.14
Far_1	3876.59	1959.08	Near_1	2853.12	988.75
Far_2	3617.20	1394.81	Near_2	2674.85	969.20
Far_3	3073.59	1754.02	Near_3	2300.21	1264.90

As was expected Response time for the correct identification of the near analogies (M = 2793.45) was significantly lower (participants were faster), than for the far analogies (M = 3656.70). The results show: F (1, 93) = 111.553, p = 0.001. Further, for the way of presentation of the analogy couples, also a main effect was found (i.e., F (3, 91) = 19.250, p = 0.001), which indicates a gradual slow dawn in response, when follows A:B::C, A:B, A, and 0 elements presentation. Consitent with the Stenbergs original findings, people made their decisions slower if they were given less information during encoding phase (Figure II), meaning that they most probably have processed the encoding information as was instructed instead of waiting the whole analogy to be visible. This important observation gives as the opportunity to consider encoding time as a meaningful dependent variable for further analyses.

As it was expected, with respect to Sternberg's findings, the amount of information for encoding increases and respectively the amount of information remaining for final processing decreases, as the Response time for correct answers also decreases. Participants were able find the analogies faster if they received almost all information beforehand, i.e. they were shown the base A:B pair and the C element of the target before AB::CD presentation. On the contrary, they were slower if they did not received anything at the beginning, i.e. just a blank screen and then were shown the AB::CD analogy.

Which is in consistent with Sternberg (1977), he also receives a gradual decrease in the Response Time. There is no significant interaction between the two factors (i.e., type of analogy and the way of presentation): F (1, 91) = 1.149; p = .334 (Figure II).

These results are consistent with findings of Green et al (2010), which state that slower response time can be expected from participants, when far analogies are concerned. So we can safely proceed with testing the main hypothesis, namely that the difficulty of the far compared to near analogies is exactly in the mapping phase, rather than in their encoding or evaluation.

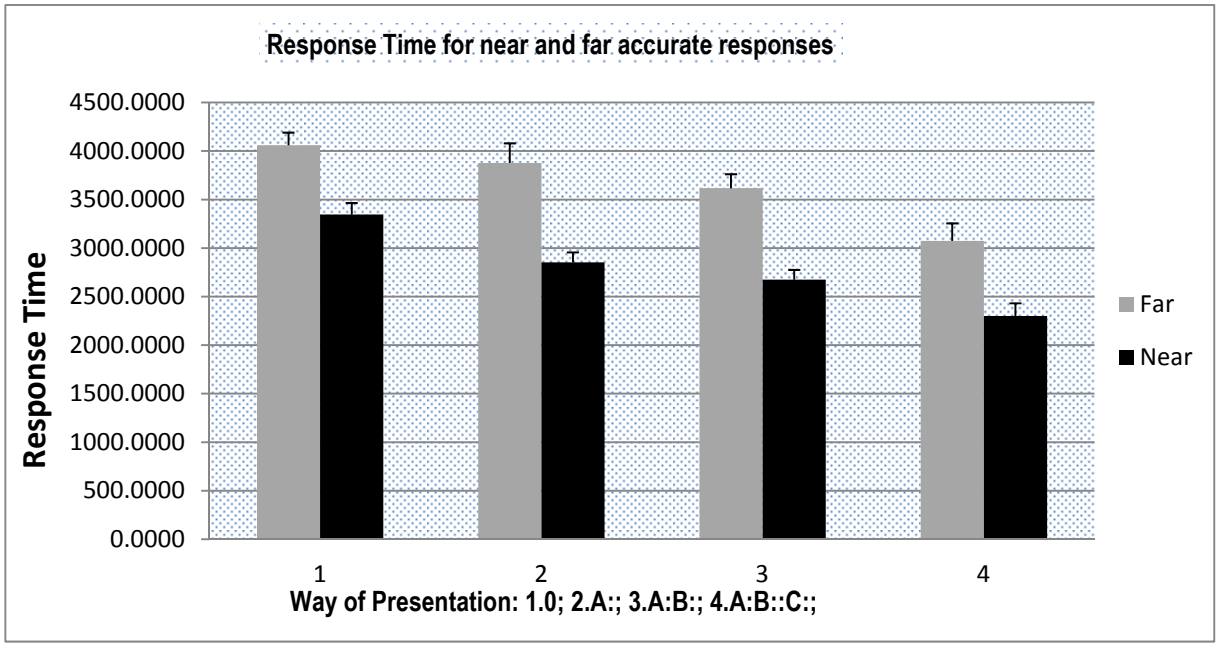


Figure II. It shows the Response time for near and far accurate responses, until participants had to click the "Yes" button.

Results obtained for the Encoding Time

The next results concern the main hypothesis and show the *Encoding Time*. This is the measurement time from the beginning of the trial until the participant clicks the "Next" button and the whole analogy becomes visible. According to this Hypothesis, a difference is anticipated to appear for the *encoding* time, when mapping starts in the 4th condition (A:B :: C:) between the far and near analogies, but not in the other three conditions (0; A:; A:B). A 2 (near and far analogies) x4 (presentation of 0; A; AB: AB:C) Repeated Measures Anova was again used for the data analysis. In Table 3 the Mean, Standart Deviation for Encoding time, for each condition for far and near analogies, are shown:

Table3. Mean and Standard Deviation for Encoding time for near and far analogies, by condition

ET by condition	Mean (ms)	Std. Deviation	Accuracy by condition	Mean (ms)	Std. Deviation
Far_0	736.97	291.46	Near_0	742.11	337.74
Far_1	834.77	349.25	Near_1	834.89	327.59
Far_2	1141.20	559.10	Near_2	1099.41	483.39
Far_3	1449.30	552.29	Near_3	1416.35	507.03

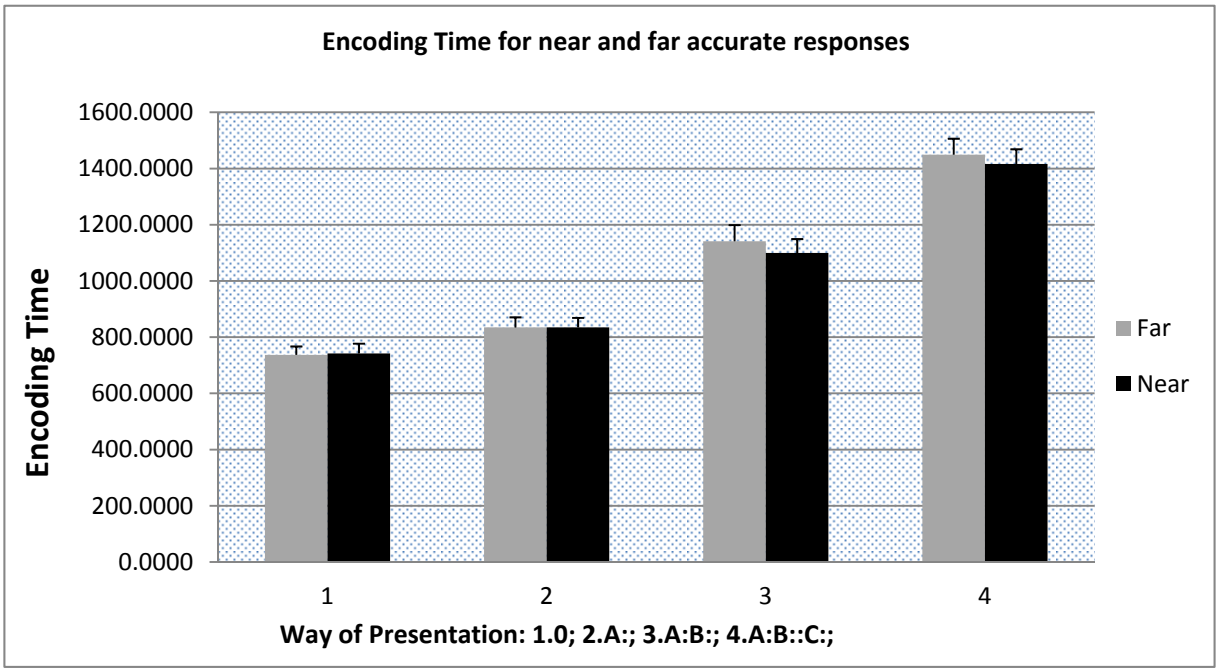


Figure III. Shows the Encoding time for type of analogy with the Standart error values.

The registered mean for Encoding time for near analogies (mean Encoding time, M = 1023.19) is not statistically significant, when compared to far analogies (mean Encoding time, M = 1040.56).

The results show that no main effect was registered for the type of analogy, F (1, 93) = .950; p = .332. Nonetheless, for the way of presentation of the word-pairs there is a main effect: F (3, 91) = 92.497, p =

0.001. Likewise, there is no significant interaction recorded between both factors: F (1, 91) = .429, p = .732 (Figure III).

The hypothesis regarding the *Encoding* time is unconfirmed, due to the lack of any significant interaction between the two factors. Furthermore, the Encoding time for *A:B::C* for near and far analogies was not statistically different from Encoding time of the other conditions. Hence, mapping that may be initiated because of the presentation of the *C* element of the analogy was not statistically more difficult than encoding of *A:B* elements that composed far, compared to near analogies.

Discussion and Conclusions

This study was set out to explore, additionally, the process of analogy making, when one is faced with distant (cross-domain) and close (within-domains) analogies. Green et al (2010), explained the processing differences between near and far analogies with the cognitive demands of far analogy mapping compared to near ones. For that reason the analogy presentation was conducted via the 4-term pairs type: *A goes to B as C goes to D (A:B :: C:D),* which were presented with the 4-level method of Sternberg (1970). If the mapping sub-process of analogies was more cognitively demanding than mapping of the near analogies (Green et al., 2010) than *encoding* time for the 4th condition *(A:B::C:)* should be slower for the far compared to near analogies. However, such differences between A:B::C encoding time of far and near analogies were not obtained, although the far analogies were identified slower and with more mistakes than the near ones, as was expected. Therefore, since far analogies were more difficult than near ones, but no differences for the type of analogies were found, with respect to the condition where analogical mapping was enabled. The evaluation, rather than mapping of analogous of far compared to near pairs may be considered to be responsible for the delayed answers reported in the Green et al's study, (2010).

Evaluation, in analogy making is the final sub-process, where one has to make a decision, whether one is confident in the validity of analogy. The bigger the confidence the stronger the evaluation, and no additional time is wasted on wondering on "what if". In the current experiment there was no condition which showed the whole analogical pair and then to have to click "yes" or "no". So it is difficult to say that evaluation, rather than mapping of the far compared to near analogies slows down the response, but at least the mapping seems to be less probable candidate for explaining the difficulties with far analogies and the heightened recruitment of the prefrontal brain areas, reported in Green et al, (2010).

Acknowledgements

The paper is published with partial support by the ITHEA ISS (www.ithea.org), the ADUIS (www.aduis.com.ua) and the Central Fund for Strategic Development, New Bulgarian University.

Bibliography

[Dahl and Moreau, 2002] Dahl, D.W and Moreau, P. (2002). *The influence and value of analogical thinking during new production ideation.* Journal of Marketing Research, (February), 47-60;

[Gentner, 1982] Gentner, D. (1982). *Are scientific analogies metaphors?* In D.S Miall (Ed.), Metaphor: Problems and Perspectives (pp.106-132). Brighton, England: Harvester Press Ltd.

[Gentner, 1989] Gentner, D. (1989). *The mechanisms of analogical learning.* In S. Vosniadou & A. Ortony (Eds.), Similarity and analogical reasoning (pp. 199-241). London: Cambridge University Press (Reprinted in Knowledge acquisistion and learning, 1993, 673-694).

[Green et al, 2006] Green, A. E., Fugelsang, J., & Dunbar, K. (2006). Automatic activation of categorical and abstract analogical relations in analogical reasoning. Memory & Cognition, 7, 1414-1421.

[Green et al, 2008] Green, A. E., Fugelsang, J. A., Kraemer, D. J. M. & Dunbar. K. (2008). *The Micro-Category account of analogy.* Cognition, 106, 1004 – 1016.

[Green et al, 2010] Green, A. E., Fugelsang, J. A., Kraemer, D. J. M., Grey. J, & Dunbar. K. (2010). *Connecting long distances: distance in analogical reasoning modulates frontopolar activities.* Cerebral Cortex January 2010; 20:70-76.

[Holyoak, 2012] Holyoak, K.J. (2012). *Analogy and relational reasoning*. In K. J. Holyoak & R. G. Morrison (Eds.), The Oxford handbook of thinking and reasoning (p. 234-259). New York: Oxford University Press.

[Kokinov and French,2003] Kokinov, B. N. & French, R. M**.** (2003) *Computational models of Analogy-making*. In Nadel, L. (Ed.) Encyclopedia of Cognitive Science. Vol. 1, pp.113 - 118. London:Nature Publishing Group.

[Sternberg,1977] Sternberg, R. J., (1977). *Component processes in analogical reasoning.* Psychological review, *Vol. 84, No. 4, 353-378.*

[Sternberg,1996] Sternberg, R. J. (1996). *Intelligence: knowns and unknowns*. American Psychological Association. Vol. 51, No. 2, 77-101.

Authors' Information

Alexandra Alexieva *–Master student in Cognitive Science, NBU, Sofia-1618, Bulgaria, Montevideo 21; e-mail: sandra_al@abv.bg*

Penka Hristova *- Assist. Prof., NBU, Sofia-1618, Bulgaria, Montevideo 21, e-mail: phristova@cogs.nbu.bg;*

Major fields of scientific research (keywords) *- thinking, memory, emotions, unconsciousness.*

EMPATHY AND MORAL JUDGMENT IN TROLLEY-LIKE DILEMMAS

Veselina Kadreva, Evgeniya Hristova

Abstract*: A large body of recent research in the field of moral psychology has established the role of emotional processing in judgment. Previous findings support the dual-process model of moral judgment [Greene et. al., 2001] suggesting that it is driven both by controlled cognitive processes and automatic emotional processes. The present study provides data in support of this model exploring emotional empathy and its association with moral judgment of emotionally salient (personal) moral dilemmas. Strictly controlled stimuli are used in order to explore this correlation. Overall, personal dilemmas are judged as less permissible and receive lower permissibility ratings than impersonal dilemmas, which confirms the importance of the previously established in the literature personal-impersonal distinction. Results also show that individuals who score low on a measure of trait emotional empathy towards people and animals give higher permissibility ratings for the utilitarian action in personal trolley-like moral dilemmas, e.g. lower scores on emotional empathy scale is correlated with higher permissibility ratings for highly emotional personal moral dilemmas. No such correlation is observed for impersonal dilemmas. This pattern of results adds to previous findings revealing the importance of the emotional component in moral judgment and underscores the specific role of empathy in judgment of personal moral dilemmas.*

Keywords*: moral judgment, empathy, moral dilemmas, emotions*

ACM Classification Keywords*: A.0 General Literature - Conference proceedings*

Introduction

Moral judgments are often studied using moral dilemmas - descriptions of hypothetical situations that pose a conflict between moral rules. E.g., in the classical Trolley problem [Foot, 1978], one needs to decide whether it is morally permissible to pull a switch so that a trolley about to kill five people on its way is redirected to another track, killing one person only. An alternative situation is represented by the 'Footbridge dilemma' [Thomson, 1985]. Here, only pushing a stranger in front of it could stop the trolley headed towards the five people.

A lot of studies have established that people judge it morally permissible to pull the switch in the Trolley problem, but not to push the stranger in the Footbridge dilemma (e.g., [Greene et al., 2001, 2008];

[Cushman et al., 2006]; [Hauser et al., 2007]). This behavioral dissociation is interesting as, from utilitarian point of view, both situations are equal: the sacrifice of one person leads to saving five other. It appears that moral judgment cannot be explained only by the application of certain principles either complying with utilitarian or consequentialists ethical norms.

As trolley-like problems successfully capture this tension, representing a moral dilemma, they have been extensively used as stimuli in a large body of research in the field of moral psychology. What is more, this type of moral dilemmas allow for a lot of factors (e.g. severity of harm, personal interest, intentionality, etc.) to be manipulated so that their influence on judgment could be studied in a detailed manner.

Emotions, empathy, and moral judgment

Actually, the most prominent theory to explain this dissociation is the Dual-process theory introduced by [Greene et.al, 2001]. It states that moral judgment is led both by controlled cognitive processes and automatic emotional processes. Controlled cognitive processes are in favor of utilitarian judgments: 'as five people are more than one, doing harm for the greater good is permissible'. On the other hand, harm aversion produces an emotional response that, if strong enough, might interfere with rational utilitarian calculations and restrain from action. The authors claim that 'personal' infliction of harm in the 'Footbridge dilemma' produces a salient emotional response compared to the one in the 'Trolley problem' where harm is inflicted impersonally and this distinction in emotional reaction is responsible for the observed differences in judgment. Using a set of personal and impersonal moral dilemmas, they provide both response time and neuroimaging evidence in support of the theory.

A lot of further studies on emotion and moral judgment have also provided evidence in support of the dual-process model of moral judgment. E.g, [Valdesolo & Desteno, 2006] have established that preliminary induced positive mood might diminish the negative response produced by personal scenarios resulting in a greater proportion of utilitarian judgments. For impersonal dilemmas, no similar effect was observed. On the contrary, when negative affect was induced preceding judgment, participants judged moral transgressions as less permissible [Wheatley & Haidt, 2005]; [Eskine et al., 2011]. What is more, neurophysiological studies [Moretto et al., 2010] have established that judgment of personal dilemmas in neurotypical populations is accompanied by greater arousal that is related to the smaller proportion of utilitarian judgments. For patients with damaged ventromedial prefrontal cortex who failed to exhibit a visceral reaction to personal dilemmas, no similar behavioral pattern was observed.

Other studies on clinical groups with disrupted emotional processing have also demonstrated that diminished emotional response could explain a greater probability for endorsement of utilitarian judgments, specifically for personal dilemmas. E.g. patients with frontotemporal dementia, frontal

traumatic brain injury, clinical psychopaths and individuals with psychopathic tendencies who are characterized by deficits in emotionality, social affect and inability to produce an empathetic response, all exhibit abnormal patterns of moral judgment [Patil & SIlani, 2014].

Further studies aiming to identify the key components of emotional processing have actually established that specificities in judgment are observed also for healthy adults, but could be explained by individual differences in predispositions to experience empathic response. In a sequence of three experiments, [Gleichgerrcht & Young 2013] provided consistent evidence that people who score low on empathic concern, as measured by IRI scale [Davis, 1983] are prone to judge in a utilitarian manner for personal dilemmas, but there is no specific difference for impersonal ones. What is more, controlling for gender, the authors established that even previously documented gender differences in judgment, might actually be attributed to gender differences in empathic response. These results are in line with dual-process theory: people low on empathic concern fail to produce an emotional response, strong enough to prevent them from utilitarian judgment for personal dilemmas which are considered emotionally salient and there is no difference for less emotional impersonal dilemmas.

Factors Affecting Moral Judgment

Apart from the well-documented distinction between personal and impersonal dilemmas, in the literature have been established three other important factors which need to be controlled for: Benefit Recipient (self vs. other benefit), Inevitability of death (avoidable vs. inevitable), and Instrumentality of harm (instrumental vs. incidental) [Christensen & Gomila, 2012].

Benefit Recipient: this factor reflects whether the suggested moral violation would lead to consequences beneficial for the protagonist or not. Research consistently show that people are more likely to approve of moral violations when consequences would bring benefit for themselves [Bloomfield, 2007]. What is more, self-beneficial dilemmas are rated as both more arousing and negative compared to other-beneficial dilemmas [Christensen et al., 2014].

Inevitability of death: the factor is also well-recognized in the literature: when the person to be sacrificed is going to die regardless of intervention (inevitable death), dilemmas are judged as morally permissible compared to dilemmas where the victim is not endangered (avoidable death) [Moore et al., 2008]. On the other hand, according to [Christensen et al., 2014] findings, avoidable and inevitable dilemmas are not discriminative in terms of arousal and valence.

Instrumentality of harm: Instrumental dilemmas are considered those in which the action is intentional, aiming to harm one person in order to save five other. In incidental dilemmas, harm is inflicted as a collateral damage, indirectly resulting from the intervention. When harm is inflicted as a side effect (incidental harm), it is judged as more permissible compared to intentionally induced harm (instrumental

dilemmas), (e.g. [Hauser, 2006]; [Mikhail, 2007], [Moore et al., 2008]) and no differences in valence or arousal are established [Christensen et al., 2014].

The results reviewed above demonstrate that these factors need to be accounted for and need to be strictly controlled in further studies on moral judgment.

Goals and Hypotheses

The present study aimed to study the on the association between moral judgment and trait empathy using the Bulgarian adaptation of the Emotional Empathy Tendency Scale (EETS) [Mehrabian & Epstein, 1972]. Based on previous research, personal and impersonal moral dilemmas were used, as this distinction is found to be important as moral judgments in personal moral dilemmas are considered to involve stronger emotions. Strictly controlled stimuli were used in order to study this relationship, exploring several other factors established as significant in judgment (inevitability of death and instrumentality of harm). In line with previous research, participants low on trait empathy were expected to be more prone to endorse utilitarian actions for personal dilemmas. No similar effect was expected for impersonal dilemmas.

Method

Stimuli and Design

In the current experiment, moral dilemmas in which three factors are varied, are used:

• Physical directness of harm - harm is inflicted by physical contact (personal harm) or is mediated through mechanical means (impersonal harm);

• Instrumentality of harm - harm is inflicted intentionally as an instrument to save other endangered people (instrumental harm) or is a byproduct of engagement in another activity, aiming to save more people threatened (incidental harm);

• Inevitability of death - deathful harm needs to be inflicted either to a person that is going to die anyway (inevitable death), or to a person that is not endangered by the situation described in the scenario (avoidable death).

Physical directness of harm (personal vs. impersonal), instrumentality of harm (instrumental vs. incidental) and inevitability of death (avoidable vs. inevitable) factors are manipulated in a within-subjects design. Each participant is presented with a total of 12 dilemmas.

Stimuli are designed based on 4 scenarios (2 avoidable scenarios and 2 inevitable scenarios) each with 3 possible resolutions - personal instrumental, impersonal instrumental and impersonal incidental resolutions. However, all personal dilemmas are designed so that harm is always inflicted as an

instrumental harm, so that did not allow for full-factorial design. As a result, 8 of the 12 dilemmas are instrumental (4 personal and 4 impersonal) and the other 4 dilemmas are impersonal incidental (see also Table 1).

Table 1. An example of avoidable dilemma with three possible resolutions.

Scenario	You are in a factory. You are standing on a platform above a railway track. Some loaded trolleys are moving along the rails. One heavy loaded trolley is speeding towards five workers as its breaks had suddenly stopped working. There is no time for them to run away and they are going to die. The trolley could be stopped only if a heavy object is set on its way.		
Resolution	**Personal/Instrumental**	**Impersonal/Instrumental**	**Impersonal/Incidental**
	The only thing that you can do is to push the worker standing next to you on the platform. He is going to fall down on the rails. Together with the tools that he is equipped with, the worker is heavy enough to stop the moving trolley. He is going to die but the other five workers will be saved.	The only thing that you can do is to activate a control button and to release the safety belt of a worker hanging from a platform above the rails. The worker will fall onto the rails of the trolley. Together with the tools that he is equipped with, the worker is heavy enough to stop the moving trolley. He is going to die but the other five workers will be saved.	The only thing that you can do is to activate a control button and to release a large container hanging from a platform. It will fall onto the rails of the trolley. The container is heavy enough to stop the moving trolley. On the top of the container there is a worker who will also fall on the rails. He is going to die but the other five workers will be saved.
Question	**Is it permissible to act as described?**		
Rating Scale	**To what extent is it permissible to act as described?** 1 2 3 4 5 6 7 **Forbidden Permissible Obligatory**		

All of the stimuli are constructed with the aim to strictly control for possible confounding factors identified in the previous research:

1) All of the stimuli are homogenously structured: introductory paragraph describes the situation, followed by one sentence that introduces the one and only means of escape; finally, a resolution is suggested in a third paragraph.

2) In all dilemmas, only two avoidable and two inevitable scenarios (introductory paragraph) are used. In order to manipulate physical directness of harm and instrumentality of harm only the resolution paragraphs are modified. Three versions of resolutions for each of the situations are designed: personal/instrumental, impersonal instrumental, and personal incidental. In such a way, we control for possible differences arising from the specific situations described.

3) In all situations a constant tradeoff between killing one person and saving five other persons is described.

4) In all dilemmas participants are assigned the role of the protagonist.

5) In all of the dilemmas there is no self-risk for the protagonist.

6) The introductory paragraph describes simply a presence of the protagonist in a certain working environment without explicitly assigning a specific role or any responsibilities to them.

7) All of the six endangered persons are identified with equal roles in the described working environment /one and the same for all six persons- workmen, miners, crew members/ thus suggesting equal responsibilities.

8) The endangered and potentially sacrificed persons are adults only.

9) All situations are designed to illustrate artificial scenarios in order to avoid potential confounding effects, e.g. familiarity with a certain situation or readily available personal opinion on resolutions.

Each situation is followed by one and the same question: 'Is it permissible to act as described?' with two possible responses –'Yes' and 'No'. After that, the described action is evaluated on a 7-point Likert scale (where '1' stands for forbidden, '4' stands for permissible, and '7' stands for obligatory) answering the question 'To what extent is it permissible to act as described?'.

For each dilemma the following measures were recorded:

- Number of responses 'Permissible': in further analyses, calculated as the average ratio of 'Yes' responses given to the question 'Is it permissible to act as described?'.
- Permissibility ratings: participants indicated on a seven-point Likert scale, with anchors: 1 – forbidden, 4 – permissible, 7 – obligatory, whether the suggested action was morally permissible answering the question 'To what extent is it morally permissible to act as described?'.

- Response time: time needed to give a response to the question 'Is it permissible to act as described' following a key press confirmation that they have read and understood the dilemma.
- Skin-conductance response (SCR) during the response period was recorded as a measure of autonomic arousal.

At the end of the experiment, participants filled out a paper-and-pencil version of EETS questionnaire [Stoyanova, 2011]. The EETS questionnaire is a self-report measurement used to evaluate empathy as a construct which reflects specifically the emotional response to perceived emotional experience of others, rather than cognitive components of empathy (e.g. perspective taking and theory of mind). High scores on the questionnaire indicate a high responsiveness to other peoples' emotional reactions.

The adapted version of the instrument [Stoyanova, 2011] consists of 25 items assessed on a 4-point response scale. Participants are presented with 25 statements describing certain behaviors and need to indicate on a 4 point response scale (from '1 – No, Never' to '4 – Yes, Always') to what extent each of the statements apply for themselves. Although, in the literature, total scores are reported, scores on 2 factors with high reliability are also subjected to further analyses. The two factors are 'empathy towards people and animals' ($\alpha = 0.76$) and 'empathy towards fictitious characters' ($\alpha = 0.64$). Sample items are presented at Table 2.

Table 2 Sample items of EETS.

Item N	Sample Items
1	It makes me sad to see a lonely stranger in a group.
5	I tend to get emotionally involved with a friend's problems.
7	I tend to lose control when I am bringing bad news to people.
6	Sometimes the words of a love song can move me deeply.
10	I like to watch people open presents.
20	I am very upset when I see an animal in pain.

Participants and Procedure

A total of 34 participants (11 males, 23 females) took part in the experiment. The age range was from 16 to 40 (M = 23). The participants took part in the experiment in exchange for partial credit toward an undergraduate course requirement or voluntarily.

Each participant was presented with 12 dilemmas – 4 situations (2 avoidable and 2 inevitable) each with 3 possible resolutions (personal instrumental, impersonal instrumental, and impersonal incidental). For each participant, the dilemmas were presented in a pseudo-randomized order ensuring that the same situation never appears in two consecutive dilemmas. In total, twelve pseudo-randomized orderings were used.

Participants were tested individually. First, the electrodes for recording skin conductance were put to provide enough time for the gel used to be absorbed. Next, the experimenter read the instructions. In the instructions, it was emphasized that participants had to imagine that the action described was the only action possible; that they had to disregard legality and had to consider only moral appropriateness of judgment. Each participant was asked to remain relatively still in order to avoid artifacts in the recordings.

First, four practice dilemmas were shown. Next, the twelve stimuli were presented using E-Prime 1.2 software. Each of the stimuli was presented on two consecutive screens: on screen 1 appeared only the first paragraph of the dilemma – the scenario. Participants indicated that they have read and understood the scenario by pressing the spacebar and advanced to the second screen which presented both the scenario and the resolution (TFigure 1). Following a keypress confirmation for reading completion and understanding, participants advanced to the next screen, the dilemma text disappeared and the question: 'Is it permissible to act as described?' appeared on a new screen. Participants indicated either 'Yes' or 'No' using the computer keyboard and advanced to the final screen presenting a question: 'To what extent is it permissible to act as described?' and rating scale with anchors '1 – forbidden', '4 – permissible', '7 – obligatory'. Participants entered a number from 1-7 using the computer keyboard. Response was followed by 700 ms inter-trial interval (Figure 1).

After the participants completed the first part of the experiment, each of them filled out a paper-and-pencil version of EETS questionnaire adapted for Bulgarian population [Stoyanova, 2011].

Results

In the analyses, we used the responses for 8 of the 12 dilemmas. In those 8 dilemmas physical directness of harm and inevitability of death factors were varied resulting in 2 dilemmas for each combination of the factors' levels. All of those 8 dilemmas were instrumental.

Responses Permissible

Number of responses 'permissible' was analyzed in a repeated measures ANOVA with physical directness of harm (personal vs. impersonal) and inevitability of death (avoidable vs. inevitable) as within-subjects factors. Analysis revealed main effect of physical directness of harm ($F(1, 33 = 9.72)$, $p = .004$): impersonal harm was judged as more permissible than personal harm (57 % vs. 41 %) 'permissible' responses (Figure 2).

The interaction was not significant and there was no main effect of inevitability of death.

Permissibility ratings

Permissibility ratings (on a scale from '1 – forbidden' to '7 – obligatory') were analyzed in a repeated measures ANOVA with physical directness of harm (personal vs. impersonal) and inevitability of death (avoidable vs. inevitable) as within-subjects factors. Analysis revealed main effect of physical directness of harm ($F(1, 33 = 8.42$, $p = .007$): impersonal harm was judged as more permissible than personal harm (3 vs. 2.6) (Figure 3). The interaction was not significant and there was no main effect of inevitability of death.

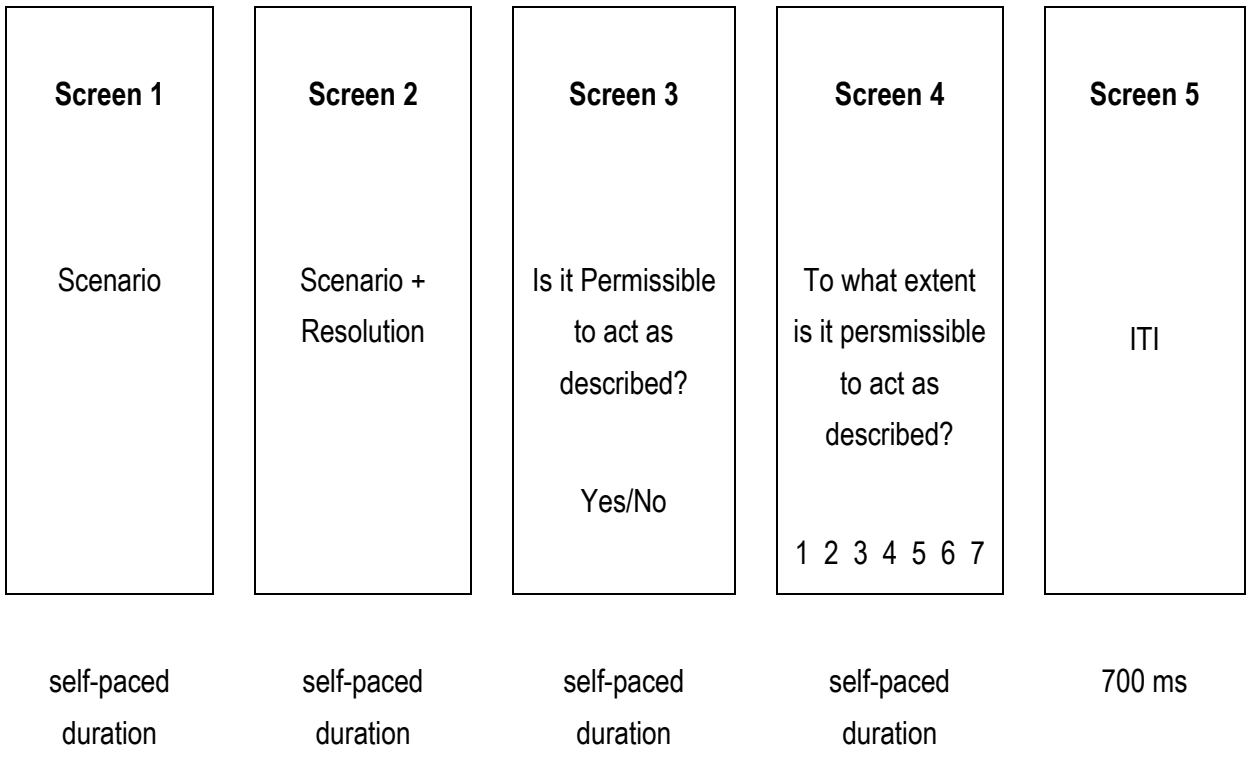


Figure 1. A schematic representation of a trial.

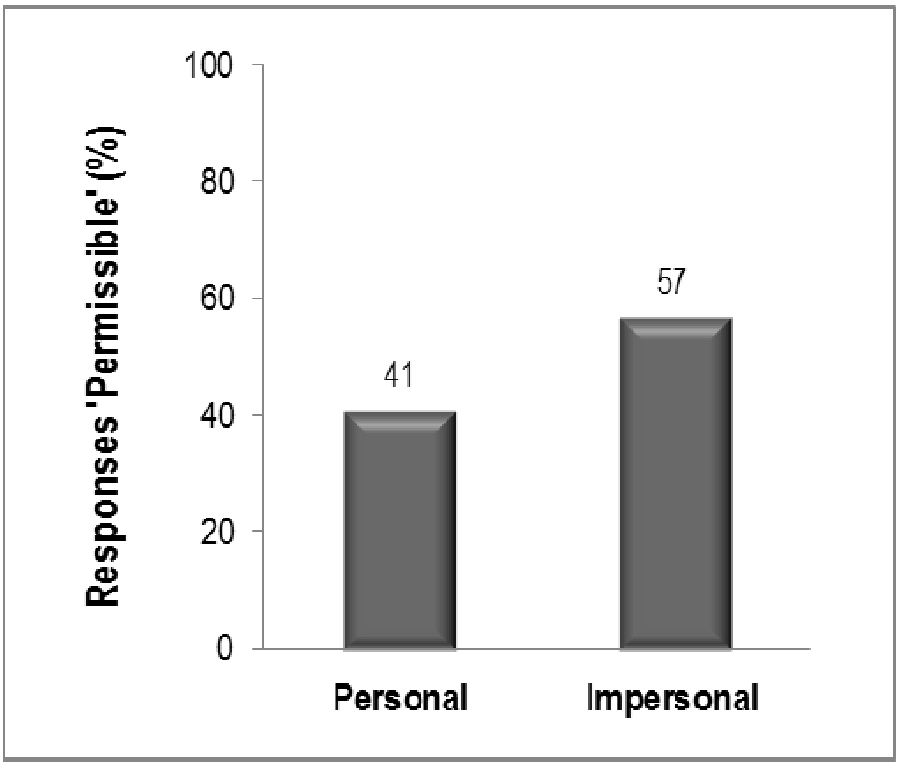


Figure 2 Percentage of responses 'permissible' for each dilemma type.

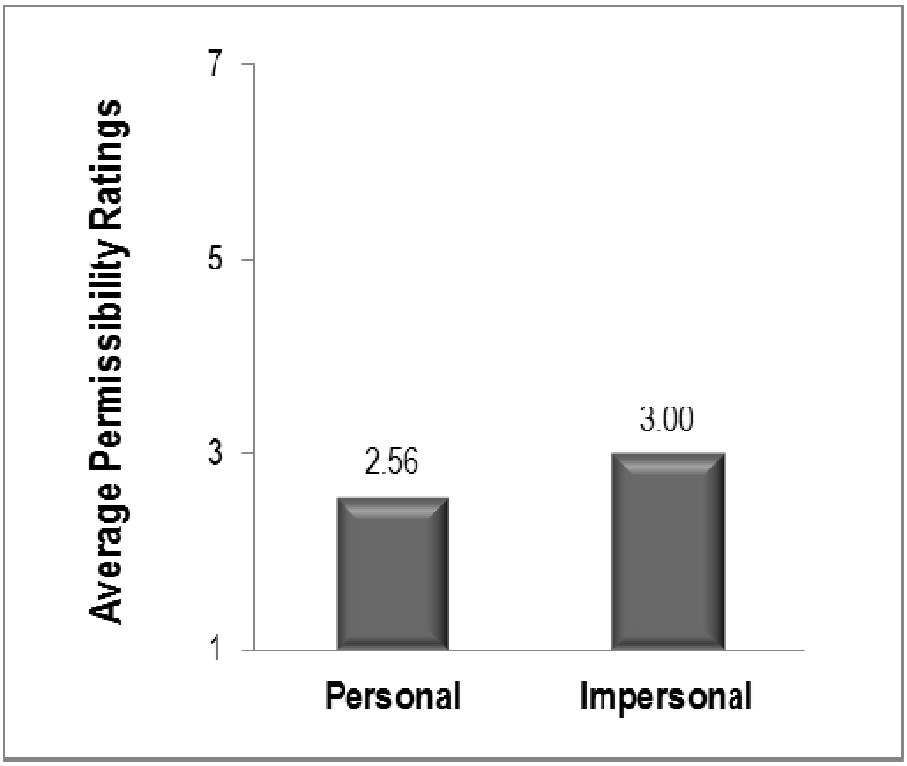


Figure 3 Average permissibility ratings for each dilemma type.

Empathy (EETS) Total Scores and Moral Judgment

First, we explored the association between average scores on the EETS scale and average permissibility ratings for each of the participants using a Pearson correlation analysis. Higher scores on EETS questionnaire indicate a predisposition for high responsiveness to others' emotional reactions and a tendency to experience a strong empathic response. No significant correlation was found between EETS scores and average permissibility ratings.

Then, average permissibility ratings for personal and impersonal dilemmas were entered in separate correlation analyses. For personal dilemmas, the correlation between EETS scores and average permissibility ratings was not statistically significant. For impersonal dilemmas, the correlation between EETS scores and average permissibility ratings was, also, not statistically significant.

Empathy (EETS Factor 1) scores and Moral Judgment

Next, we explored the association between average scores on the first factor of the EETS scale ('Empathy towards people and animals' subscale) and permissibility ratings. Higher scores on EETS factor 1 scale indicate a predisposition for high responsiveness to others' emotional reactions, a tendency to experience a strong empathic response towards people and animals. First, average scores on EETS factor 1 ('Empathy towards people and animals') and average 'permissibility' ratings for all dilemmas for each of the participants were subjected to a Pearson correlation analysis. Analysis revealed that 'permissibility' ratings were negatively correlated with average factor 1 scores ($r(29) = -.37$, $p = .046$). Lower empathy scores on that factor are correlated with higher ratings about the permissibility of the utilitarian actions.

Then, the same correlation (between empathy EETS factor 1 scores and permissibility ratings) was studied in separate correlation analyses: one for personal dilemmas and one for the impersonal dilemmas.

For personal dilemmas, there was a significant negative correlation ($r(29)= -.40$, $p = .031$). For personal dilemmas, those who scored lower on factor 1, on average, gave higher permissibility ratings (Figure 4).

For impersonal dilemmas, the correlation between EETS factor 1 scores and average permissibility ratings was not statistically significant.

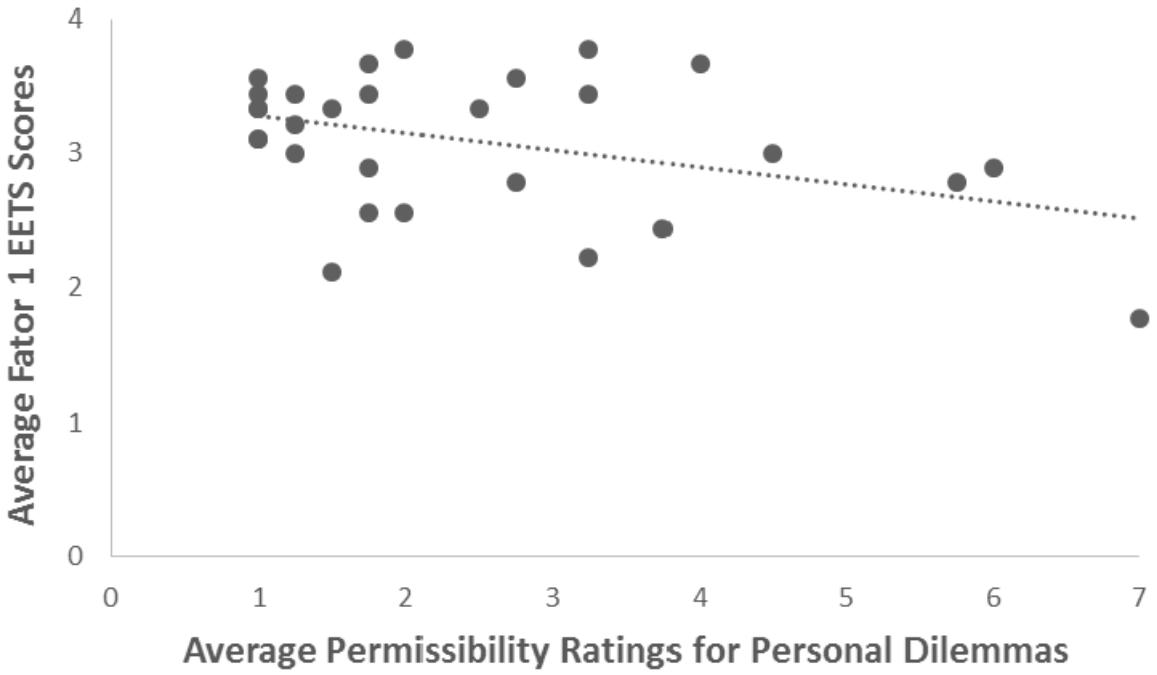


Figure 4 Negative correlation between average permissibility ratings for personal dilemmas, ranging 1 to 7 and average scores on EETS factor 1 (Empathy towards people and animals), ranging 1 to 4.

EETS Factor 2 scores

Further, the association between factor 2 scores (empathy towards fictitious characters) and permissibility ratings was explored. Higher scores on this factor indicate a tendency for high emotional responsiveness to the perceived emotional reactions of fictitious characters.

The association between average factor 2 scores and permissibility ratings for each of the participants was explored using a Pearson correlation analysis. No significant correlation was found.

Then, average permissibility ratings for personal and impersonal dilemmas were entered in separate correlation analyses. For personal dilemmas, the correlation between EETS factor 2 scores and average permissibility ratings was not statistically significant. For impersonal dilemmas, the correlation between EETS factor 2 scores and average permissibility ratings was, also, not statistically significant.

Discussion

A large body of recent research have established that emotional processing is relevant in moral judgment. Prior studies on clinical groups with disrupted emotionality have revealed that these deficiencies might be related to and explain specific patterns in moral judgment. The present study adds to previous findings demonstrating that trait predispositions to exhibit emotional empathic response in

healthy participants are related to moral judgment as well: participants low on trait emotional empathy are characterized by patterns of judgment typical for clinical populations: they judge personal dilemmas as more permissible and there is no association between emotional empathy scores and permissibility ratings on impersonal dilemmas. The present findings are compliant with a lot of studies demonstrating that individuals inclined toward utilitarianism exhibit personality traits related to diminished emotional reactivity. These results are in line with Dual-process theory showing that diminished utilitarian judgment is related to susceptibility to dilemmas of high emotional salience.

Future work needs to explore in a more detailed manner associations with empathy as a multidimensional construct. Relationships between moral judgment and tendencies to experience personal distress, empathic concern, empathy towards fictitious characters and perspective taking (as a measure of cognitive empathy) need to be studied as well. Further research will also benefit from the inclusion of a self-benefit condition in order to establish more firmly the specific role of empathy as an 'other-oriented' emotion.

Acknowledgements

The paper is published with partial support by the ITHEA ISS (www.ithea.org), the ADUIS (www.aduis.com.ua). and the Central Fund for Strategic Development, New Bulgarian University.

Bibliography

[Bloomfield, 2007] Bloomfield, P. (2007). Morality and Self-Interest. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780195305845.001.0001

[Christensen & Gomila, 2012]]Christensen, J. F., & Gomila, A. (2012). Moral dilemmas in cognitive neuroscience of moral decision-making: A principled review. *Neuroscience & Biobehavioral Reviews, 36*(4), 1249-1264.

[Christensen et al., 2014] Christensen, J. F., Flexas, A., Calabrese, M., Gut, N. K., & Gomila, A. (2014). Moral judgment reloaded: a moral dilemma validation study. *Frontiers in psychology*, 5, 607.

[Cushman et al, 2006] Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment testing three principles of harm. *Psychological science, 17*(12), 1082-1089.

[Davis, 1983] Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of personality and social psychology, 44*(1), 113.

[Eskine et al, 2011] Eskine, K. J., Kacinik, N. A., & Prinz, J. J. (2011). A bad taste in the mouth gustatory disgust influences moral judgment. *Psychological Science, 22*(3), 295-299.

[Foot, 1967] Foot, P. (1967). The Problem of Abortion and the Doctrine of Double Effect. *Oxford Review, 5*, 5–15.

[Gleichgerrcht & Young, 2013] Gleichgerrcht, E., & Young, L. (2013). Low levels of empathic concern predict utilitarian moral judgment. PloS one, 8(4), e60418.al damage. *Journal of Cognitive Neuroscience, 22*(8), 1888-1899.

[Greene et al, 2001] Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., Cohen, J. D. (2001). An fMRI investigation of Emotional Engagement in Moral Judgment. *Science, 293*, 2105–2108

[Greene et al, 2008] Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition, 107*(3), 1144-1154.

[Hauser et al, 2007] Hauser, M., Cushman, F., Young, L., Kang‐Xing Jin, R., & Mikhail, J. (2007). A dissociation between moral judgments and justifications. *Mind & language,22*(1), 1-21.

[Mehrabian & Epstein, 1972] Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of personality, 40*(4), 525-543.

[Mikhail, 2007] Mikhail, J. (2007). Universal moral grammar: Theory, evidence and the future. *Trends in cognitive sciences, 11*(4), 143-152.

[Moore et al, 2008] Moore, A. B., Clark, B. A., & Kane, M. J. (2008). Who shalt not kill? Individual differences in working memory capacity, executive control, and moral judgment. *Psychological science, 19*(6), 549-557.

[Moretto et al, 2010] Moretto, G., Làdavas, E., Mattioli, F., & Di Pellegrino, G. (2010). A psychophysiological investigation of moral judgment after ventromedial prefrontal damage. *Journal of Cognitive Neuroscience, 22*(8), 1888-1899.

[Patil & Silani, 2014] Patil, I., & Silani, G. (2014). Reduced empathic concern leads to utilitarian moral judgments in trait alexithymia. *Frontiers in psychology*, 5.

[Stoyanova, 2011] Stoyanova S. (2011), Adaptation of an empathy test among students. *Bulgarian Journal of Psychology, 1*, 181-188.

[Thomson, 1985] Thomson, J. J. (1985). The Trolley Problem. *The Yale Law Journal, 94*(6), 1395–1415.

[Valdesolo & DeSteno, 2006] Valdesolo, P., & DeSteno, D. (2006). Manipulations of emotional context shape moral judgment. *Psychological science, 17*(6), 476-477.

[Wheatley & Haidt, 2005] Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological science, 16*(10), 780-784.

Authors' Information

Veselina Kadreva *– a PhD candidate, Cognitive Science and Psychology Department, New Bulgarian University, New Bulgarian University, 21 Montevideo str., Sofia, Bulgaria; e-mail: vkadreva@nbu.bg*

Major Fields of Scientific Research*: Moral judgment, Biosignal-based research*

Evgeniya Hristova, PhD *– Research Center for Cognitive Science, Cognitive Science and Psychology Department, New Bulgarian University, 21 Montevideo str., Sofia, Bulgaria; e-mail: ehristova@cogs.nbu.bg*

Major Fields of Scientific Research*: Moral judgment, Social Dilemmas, Eye-tracking, Biosignal-based research*

THE FACTORS OF EFFECTIVENESS AND THE ORGANIZATION OF THE SEARCH OF ELEMENTS WITHIN A GRAPHIC INTERFACE

Irina Blinnikova, Anna Izmalkova, Maria Semenova

Abstract. *In our experimental study we analyzed the impacts of the following groups of factors (1 – low-level factors: the physical components (i.e. color grade) of the explored context; 2 – high-level factors: the format of the stimulus presented, which determines the accuracy of the search query; 3 – intermediary factors: subjective estimates of the familiarity and attractiveness of the stimuli) and their impacts on the search rate as well as the organization in a simulated interface that contains graphic elements. The obtained results demonstrated multiple effects of the high-level factors and intermediary subjective rating on search time and the organization of eye-movement activity. When the searched stimulus was presented in the form of a word, the search for graphic elements (icons) on the interface happened to be slower and was characterized by longer fixations in combination with short-amplitude and low-speed saccades. When the target stimulus was set as an image, the search was quick, the fixations were shorter and the saccades had a longer amplitude and higher speed. A stronger familiarity with an element had a positive effect on the rate and scope of the search. However, a higher attractiveness led to a slowing down of the search process. Although no particular effects of the low-level factors were observed in this experiment, it was demonstrated that the color grade of the stimulus pattern (matrix) might have an impact in situations when the stimulus is presented as an image and its subjective attractiveness increases. This kind of evidence creates new possibilities in the area of interactive interface development.*

Keywords*: visual search, eye movement, graphic interfaces, websites*

ACM Classification Keywords*: Human centered computing, HCI, user studies, interaction paradigms, graphical user interfaces*

1. INTRODUCTION

Visual search is one of the key fields of study in cognitive psychology. Treisman and her colleagues laid the foundation for the experimental approach in this area in the early 1980s [Treisman, Gelade, 1980]. The goal of visual search implies detecting a specific target in a display filled with distracters. The targets can be presented to subjects in various templates, i.e. imagery or verbal. In classical

experiments on visual search each stimulus was introduced as a clear and simple verbal description of the target and in later experiments stimuli were presented as images [Malcolm,Henderson, 2009]. The study of visual search remains relevant up to this day. However, we can witness a significant shift in this area. Researchers analyze the process of searching for realistic, semantically rich objects in complex visual contexts [*see* Eckstein, 2011].

The development of internet-based technologies has created new contexts and goals for visual search studies. Web pages differ from traditional stimulus materials and represent a complex combination of textual, imagery, and multimedia elements [Pogue, 2013]. This raises new questions concerning the way visual search is organized in a virtual environment and the factors that underlie its effectiveness [Hall-Phillips et al., 2013]. Our research is focused on the mechanisms of detecting graphic elements among similar distracter-stimuli. Although searching for icons on a screen is a familiar task for every computer and mobile device user, this process is the least studied in this area of research. Symbolic images (icons) have long been used in the development of tools for human-computer interaction [Garcia et al., 1994]. At first, the pictographic dictionary was relatively small, easy to memorize and use. However, with the development of internet-based technologies, it has been growing exponentially and is becoming virtually infinite. Anyone can create his or her own web page and fill it with various graphic content. It increases the interest in studying the mechanisms of visual search, perception, memorizing, and identifying pictograms [Goonetilleke et al., 2001]. For web users the pictograms can either merely exist as elements of visual design or play the role of defining points or indicators of a particular block of information. In some cases, they can even appear as the means of transition from one web page to another or from one block of information to another. Anyway, they can be viewed as the representation of access to a source of information [Hout, Goldinge, 2015].

Icons, as the representation of access, possess physical and semantic characteristics. Physical characteristics include movement, orientation, shape, size, color, etc. In his work, J. Wolfe describes dozens of such categories. Those characteristics are typically called low-level characteristics of cognitive processing [Wolfe, 1998]. In the 1980s, the icons used to be black and white and always had a frame. Later on, different variations of icons would appear and very little was done to analyze the purpose of those innovations. In our research we attempted to verify how the effectiveness of visual search depends on whether the graphic elements are presented in black and white or in color and whether they have round or square-shaped frames.

Semantic characteristics are connected with the ability to determine or complete the meaning of the searched element and are often described as the higher-level characteristics of cognitive processing. One of the possible ways to manipulate semantic parameters is with a search query or, in particular, the format and target template of the searched element. When analyzing the mechanisms of visual search,

the target stimulus can be set as an image, a word, or even as a text. In the first case, we model a strictly defined task for visual search. For instance, a subject can search for the icon that represents a folder with deleted items, commonly known as Trash. All users have seen this icon multiple times, remember it well, and possess an accurate mental representation of it. They know what they search for. However, even in this case, the search takes some time because the icon can be located in various parts of the screen or appear in a cluttered context (when there are too many icons on the Desktop).

If the target stimulus is set verbally, a less defined task for visual search is modeled. A search of this kind occurs when we do not know precisely what the target icon looks like but can predict it by constructing mental images. As an example, we used a website that sells tickets for different kinds of transportation. The task was to find the icon that directs the user to the 'air travel' category. If the web site is being used for the first time, it is impossible to say with certainty what the icon will look like. However, the user can presume that it might look like a plane, wings, or anything that can be associated with aviation. A mental activity of this kind is equivalent to the one that is described by specialists in the field of information search as *search query reformulation* [Sutcliffe, Ennis, 1998]. One of the goals of our research is to compare the different types of search processes.

The low and high-level factors can be viewed as the objective search conditions. However, one should not ignore the subjective aspect, such as a user's evaluation of familiarity and attractiveness of the searched element. These factors can be described as intermediary variables. In addition, the familiarity depends on the user’s experience interacting with a similar stimulus in the past, whereas the attractiveness can be related to the user’s emotional reaction.

The familiarity influence of target stimuli and distracters on the search rate and strategy has been the subject of previous research. The topic of familiarity is related to the impact of the user's experience interacting with the searched element or context. In one of Treisman's studies [Treisman, Vieira, Hayes, 1992], it was determined that prior encounters with the stimulus patterns sped up the search process. Moreover, the correlation between the familiarity of the stimuli and distracters also plays a significant role [Malinowski, Hübner, 2001]. Nevertheless, the effects of familiarity are not as clear when a search for a real object is involved, e.g. a company’s logo or interface icon [Qin, Koutstaal, Engel, 2014]. In this case, it is important to consider the correlation between the familiarity and both low and high-level factors [Shen, Reingold, 2001].

The attractiveness effects of target stimuli on the search process have been studied to a lesser degree. This might be explained by its higher level of subjectivity in comparison to familiarity [Salimun et al, 2012]. The primary interest here is focused on the perception, search, and detection of faces [Rhodes at al, 2003]. However, the attractiveness and esthetic value of realistic and symbolic images also has critical importance, especially when interface elements are involved [Hou, Lu ,Ho, 2011]. The most

important concern is the relation between the attractiveness and certain esthetic, particularly the color aspects on the web site as well as the characteristics of the search query.

Traditional studies on visual search normally attempted to minimize the number of eye movements in the process. However, the size and complexity of natural scenes have created the need for the analysis of eye movement activity and other behavioural components of the search process [Atkins et al.,2006]. The term *active vision* has been introduced [Halverson, Hornof, 2011]. Faraday [2000] has proposed the term *web page viewing behavior* and defined it as a series of eye movement acts. From our point of view, visual search can be defined as a consistently organized activity that carries a goal, set of conditions, achievement strategy, performance control, and so on. Nowadays, researchers in the field of eye movement analysis take into account a considerable number of factors, the main ones being fixation time, the amplitude and pace of saccades[1]. Furthermore, in the experiments on target search the so called *dwell time* is described as the time spent in the zone of interest, a part of the stimulus pattern where the searched element is located.

As far as visual search on web pages is concerned, prominent results have been achieved that connect the effectiveness of visual search and the way the virtual environment is organized with the parameters of eye movements. In one study [Pan et al., 2004], the oculomotor criteria of visual search were registered on 22 web pages. It was found that they depended on the web site's style, information organization, and the user's gender. In another study [Burmistrov et al., 2015], it was illustrated that the icon search was connected with the general design of a web page. The researchers studied the effectiveness of visual search and the patterns of the eye movement activity in relation to different visual styles of user interfaces: traditional or flat. The results showed that the participants required double the time to spot flat icons in comparison to more realistic, three-dimensional ones. The studies of oculomotor parameters have demonstrated that recognizing flat, two-dimensional pictograms in comparison to three-dimensional ones is usually characterized by lower levels of amplitudes and peak saccade rates. These results suggest the following conclusions. The first and more obvious conclusion explains the correlation between visual search and the web page's esthetic appearance. Secondly, the effectiveness of a search is related to the specific organization of oculomotor activity. Specific eye

[1] A saccade is a rapid movement of the eye that adjusts the vision system towards information analysis within a certain segment of the surroundings. A fixation is a period of motionless gazing which is traditionally considered as the point of information processing. Modern researchers seek to establish and substantiate the integrative criteria of the eye movement activity.

movement patterns can be singled out that may be able to predict a more or less effective search process.

In this study, we focused on the roles of low and high-level cognitive processes, subjective evaluation of familiarity and attractiveness of the stimuli, and the way these factors affect the search rate and organization of eye movements. For the low-level factors, we decided to consider a physical characteristic, such as color grade. For the high-level factor, we agreed to consider the format the stimulus was presented in: a picture or a word. If the target was set as a picture, the subject had to store it in the working memory and to compare it with those he or she saw on the screen. If the goal was set as a word, the subject was required to mentally produce possible images of objects they were looking for. The second situation was less certain and more complicated, and we assumed that it would increase the time and affect the strategy of eye movement activity. The effects of attractiveness and familiarity of target stimuli were assessed separately as well as in correlation with the low and high-level factors.

2. METHOD

Sixty-two healthy volunteers took part in this experiment; 41 females and 22 males ranging in age from 18 to 48, the mean age being 22 years 3 months. Prior to the experimental sessions, an orientation and a series of trials were performed. During the experiment the participants were seated 0.65 m away from a 19-inch computer screen. They were instructed to find symbolic images of real-life objects (such as a butterfly, a cactus, a book, etc) among a variety of other objects. The pictograms were taken from the Internet and modified into the stimulus material. A total of 5128 pictures were collected, among which 32 were selected at random. Prior to the main experiment, the subjects were required to rate the familiarity and attractiveness of the symbolic images from 1 to 4. Afterwards, the average indexes of attractiveness and familiarity were estimated for each stimulus, which were divided into two groups of more and less attractive stimuli (the Attractiveness Factor) and two groups of more and less familiar ones (the Familiarity Factor).

The stimuli were arranged in rectangular full screen stimulus patterns (matrix) 9x9. It was a 1466 x 954 pixel rectangle with the resolution of 300 pixels per square inch. Each matrix contained 81 objects, one of which was the target. The target stimulus was situated in one of 8 quadrants (the central quadrant was not used). Thirty-two matrices were shown, half of which were black and white, the other half were colored (the Color Factor). The stimuli consisted of black circles or squares, in the center of which either a white or colored image of a symbolic object was placed.

In half of the cases, the target stimulus was set as a text. In the other cases, it was presented in the form of an accurate copy of the searched image but in gray (the Format of the Target Stimulus Factor). The order of the matrices presented varied from subject to subject.

We recorded the search rate and eye movement data. Eye movements were sampled monocularly at 250 Hz using SMI iView X RED 4 (FireWire) tracking system with on-line detection of saccades and fixations and a spatial accuracy < 0.5°. There were 32 trials recorded for each subject, which counted for 2016 trials altogether. The trials were calculated and subjected to factorial ANOVA (two-way ANOVA) using IBM SPSS Statistics 19.

3. RESULTS AND DISCUSSION

3.1 The effects of low and high-level factors. The results demonstrated that the target template affects the search rate and the scanning process (Table 1). If the target stimulus is set as a word, it requires a longer time period on average in comparison to a picture. This outcome was, in fact, expected. A more surprising result was the fact that the format of the stimulus could determine the search strategy, i.e. a specific eye movement pattern. In other words, the stimulus' format – a high-level factor – regulates eye movement activity. If the stimulus is presented as a word, the search query turns out to be less determined. A subject would use as a target sample not just a mental representation stored in the working memory but rather its semantic field, which could include a number of visual representations. The search in this case requires a deeper and more detailed processing of information. It leads to a longer average fixation duration, shorter-amplitude and slower saccades. This type of eye movement pattern is characteristic of focal vision. If the stimulus is set as a picture, the search query is highly accurate. The stimulus sample is a particular mental representation stored in the working memory. The subject's task is to compare it with the images he or she comes across. This does not require deep semantic processing and can be accomplished with shorter fixations and quicker saccades with longer amplitudes. This eye movement pattern is typical for ambient vision.

The target stimulus was outlined in the matrix as the area of interest (AOI) and certain eye movement data was registered around that segment. The time spent within the zone of interest (AOI dwell time) not only includes the time spent on the stimulus perception and its comparison to the sample image but also on making the final decision as to whether the searched element was indeed found. As we can see in Table 1, the average time spent in the interest zone is longer when the stimulus was set as a text, although the differences are not too profound. The more apparent differences can be observed when counting the number of eye fixations in the area of interest, which could be connected with the performed cycles of cognitive processing. This outcome can possibly imply the necessity for the subject

to verify the correct stimulus in cases when the format of the search query was incongruent with the element (word vs. picture).

The assumption that the search effectiveness will differ whether the stimulus matrix is colored or black and white did not prove valid. The Color Factor turned out to be insignificant for all the measured criteria. However, some interesting results were discovered regarding the correlation between the color and frame factors with the format of stimulus. These results suggest that the stimulus format not only determines the search strategy but also affects the significance of low-level factors.

Table 1. The time of search and eye movement parameters in different formats of the target stimulus presented – as a word or a picture

	All trials	The target template		F (1, 1983)	Sig
		word	picture		
Search Time [ms]	**12119,808**	**13,142**	**11,091**	**11,972**	$p < 0,01$
Fixation Duration Average [ms]	**226,518**	**230,909**	**222,101**	**7,575**	$p < 0,01$
Saccade Amplitude Average [°]	**3,901**	**3,736**	**4,068**	**6,575**	$p < 0,01$
Velocity Average [°/s]	**108,152**	**105,107**	**111,218**	**18,008**	$p < 0,01$
Fixation Count in AOI	**3,480**	**3,667**	**3,287**	**7,013**	$p < 0,01$
DwellTime in AOI [ms]	**1745,093**	**1821,405**	**1666,006**	**4,079**	$p < 0,05$

The Color Factor in itself did not appear to be important. However, its correlation with the stimulus format had an effect on the search rate. It appears that color might actually interfere when searching for the target stimulus if it is set in the form of a word and, on the contrary, serve as an aid when the stimulus is presented as an image (Fig.1).

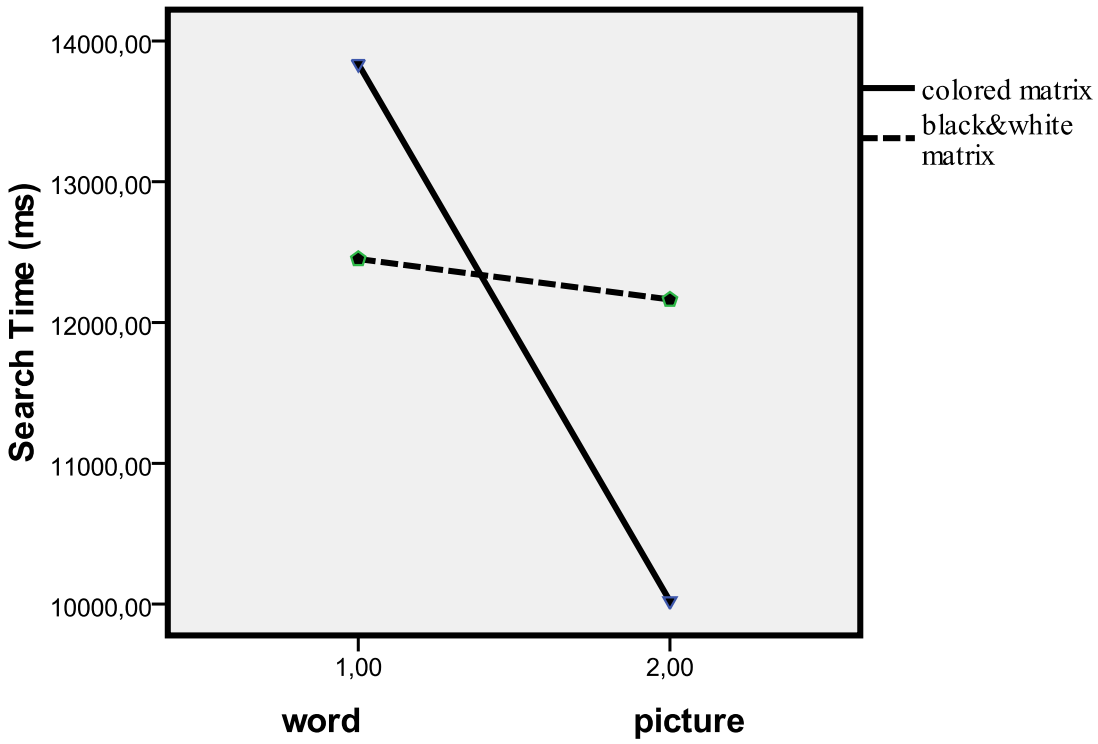


Fig.1. Search time as a function of the format and color of the stimulus matrix

While interpreting the data, it is important to consider that even when the stimulus was set as a picture it was black and white, so the color did not play a key role in detecting the searched element but rather affected the mechanisms of cognitive processing. The color grade of the stimulus provided faster cognitive processing on lower levels that are related to a mere comparison between the perceived image with the one stored in the working memory. On higher levels, when semantic analysis is involved, the color can in fact present itself as a limitation. No significant interaction was found between the two factors in relation to eye movement activity.

3.2 The influence of the familiarity of the target stimulus on the search rate and characteristics of eye movement activity. The results of our study have confirmed that the elements perceived as more familiar are usually found significantly more quickly. Furthermore, it was illustrated that the increase of familiarity changes the characteristics of the search process: the number of fixations and saccades as well as the length of distance covered per time unit increase (Table 2).

At the same time, the duration of fixations and saccade amplitudes do not vary for either levels of familiarity. The search process has a wider scope than with the less familiar stimuli. These conclusions are in line with the earlier assumptions of Q. Wang, P. Cavanagh, M. Green [1994]. They supposed that

more familiar stimuli allowed subjects to use intuition in searching and decision-making processes. In these cases subjects rely on a sudden (pop out) spotting of a searched stimulus.

Table 2. The time of search and eye movement parameters in different levels of familiarity

	All trials	Low-level Familiarity	High –level Familiarity	F (1, 1983)	Sig
Search Time [ms]	12119,808	13493,057	10551,226	11,972	$p< 0,01$
Fixation Count	18,439	17,828	19,589	3,455	$p= 0,05$
Scanpath Length [px]	3883,5	3686,07	4255,08	7,003	$p< 0,01$
Fixation Count in AOI	3,480	3,662	3,274	7,311	$p< 0,01$
DwellTime in AOI [ms]	1745,093	1768,011	1719,154	0,401	$p> 0,5$

A prominent correlation was observed between the level of the target stimulus' familiarity and its presentation format in regards to the search rate index (Fig. 2).

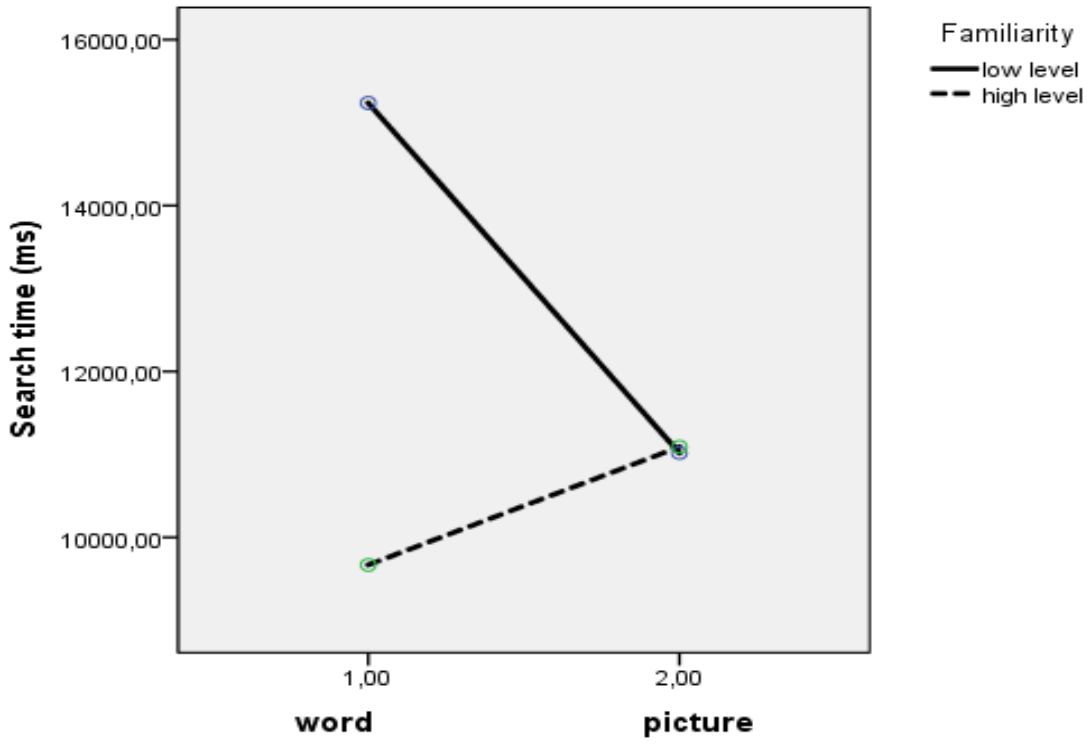


Fig. 2. Search time as a function of familiarity and target presentation format

As we presumed, the familiarity was more important in the situations when the target stimulus was set as a text. When it is set as a picture, its representation is stored in the working memory. In this case, its general long-term familiarity does not play an important role. However, when searching for a stimulus presented as a word, the subject has to compare new elements with the information stored in the long-term memory. In such instances, familiarity (i.e. the past experience of interaction with the stimulus) can play a crucial role. In addition, no interaction between color and familiarity was observed.

3.3. The effects of stimulus attractiveness on the search time and eye movement activity. Somewhat weaker effects were registered for the factor of attractiveness. It is not particularly odd since the estimates of attractiveness are highly individual and can be characterized by a wide range of results. In general, it appeared that the attractiveness can lengthen the search rate and time spent in the zone of interest (Table 3).

Table 3. The time of search and eye movement parameters in different levels of attractiveness

	All trials	Low-level Attractiveness	High –level Attractiveness	F (1, 1983)	Sig
Search Time [ms]	**12119,808**	**11227,557**	**13047,990**	**6,016**	***p*==0,01**
Fixation Duration Average [ms]	**226,518**	**223,450**	**229,698**	**6,988**	***p*< 0,01**
DwellTime [ms]	**1745,093**	**1566,793**	**1937,741**	**13,317**	***p*< 0,01**

The attractiveness of the target stimulus can make subjects fixate their attention on the target and become distracted. It also corresponds with the lack of data that can be found in scientific literature. Particularly, in the study by Salimun, Purchase, and Simmons [2012], it was demonstrated that subjective attractiveness can lead to longer search rates and increase the number of errors.

As we predicted, a relevant correlation between the factor of attractiveness and the color factor was found in terms of search rate (Fig. 3). In colored stimulus matrices, searching for attractive stimuli takes longer than in black and white matrix. It gives evidence to the fact that the color factor is related to and contributes to attractiveness. We did, however, assume that the attractiveness of the target stimulus would have a stronger effect for the stimuli represented as images, but that assumption was not verified. Likewise, we did not see a particular correlation between the factors of attractiveness and familiarity.

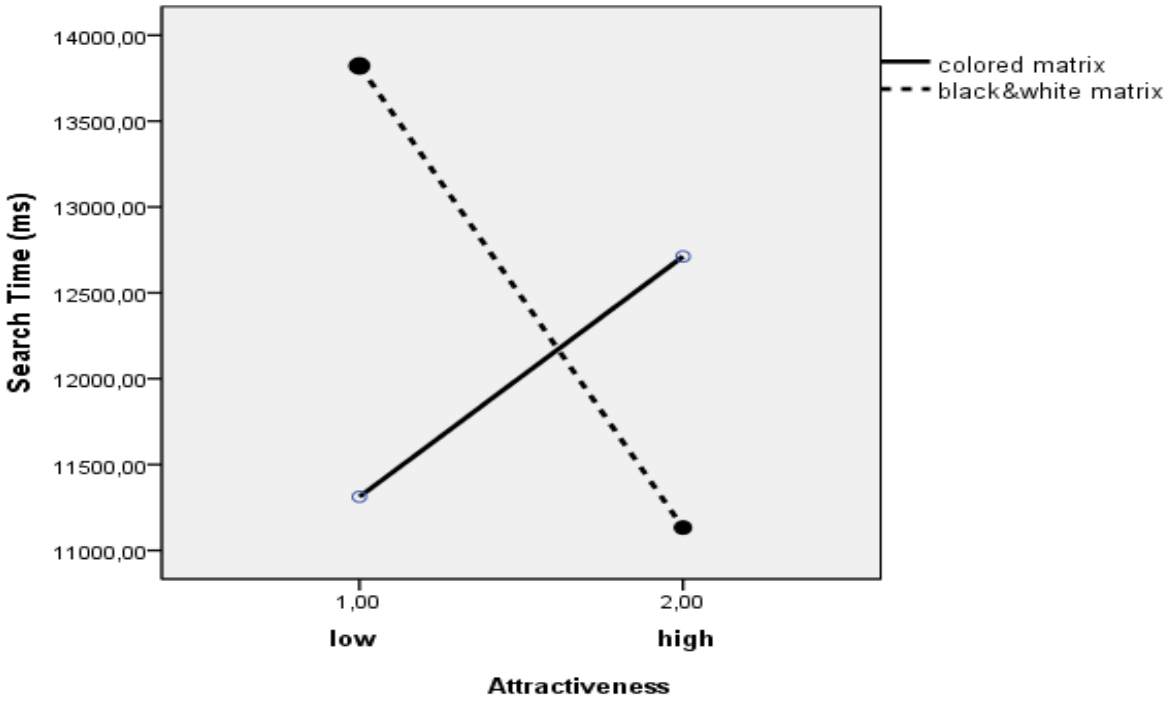


Fig. 3. Search time as a function of attractiveness and the color grade of the stimulus matrix.

4. CONCLUSION

Searching for target objects in real or virtual surroundings is one of the most important life tasks. In order to succeed, one has to set the right direction or select the right segment of the environment while disregarding everything that is not related to the target before its detection and then, finally, confirming the accuracy of his or her choice. It is necessary to possess a more or less clear representation (or description) of the searched object to be able to compare it with the perceived information. The results of our empirical analysis showed that the target template has a big impact on the effectiveness of the search process as well as the manner of scanning the environment. It was discovered that a textually set target stimulus requires a longer time period to be found. Moreover, the nature of cognitive processing changes – longer fixations occur, including slower saccades with shorter amplitudes. A pattern of eye movement of this kind indicates deeper information processing that reaches the level of semantics.

It appeared that the color grade of a stimulus, as a low-level factor, does not play a significant role on its own, although it interacts with other factors. It was illustrated that the color grade helps the subject to detect the target stimulus within a matrix when it is set in the form of a picture and, on the contrary, can hinder the search process when it is presented as a word. This knowledge suggests that the format the target stimuli are presented in affects the search process as well as determines the role of lower-level

factors. If the target stimulus is presented as a picture, the search is conducted through surface-level cognitive processing. At this level, the various physical features of the matrix can play an important role – in particular, the facilitating effects of the color grade become more apparent. If the target stimulus is set as a word, its search and detection requires a deeper semantic analysis. In this case, the role of the physical characteristics is the reverse. What made it easier to detect the target object at the surface-level of processing now has a negative effect.

The role of subjectivity was also described. We were able to see that the familiarity of the stimuli speeds the search process and makes it wider. Notably, more familiar objects could be found more quickly when set as words. However, when the target was set as an image, its familiarity did not make a difference. The attractiveness of the stimulus had the opposite effect – it slowed down the process. Furthermore, the factor of attractiveness correlated with the color factor as a low-level factor but did not have any connection with the high-level factors.

The resulting data could be applied in the area of interface development. In particular, it is important to take into account the type of search query and task that a user will potentially come across. When dealing with visual search, it is essential to keep in mind the habitual format of a search query. Since a web page user's task is formulated verbally in most cases, the developers should cautiously apply tools that can increase the esthetic and emotional value of the interface and its elements. It is also wise to limit the variety of graphic elements where diversity could in fact decrease subjective familiarity and, thus, decrease the effectiveness of the search.

Acknowledgement

This research is supported by the Russian Foundation of Basic Research; Grant №14-06-00371.

References

[Atkins et al.,2006] Atkins M. S., Moise A., Rohling R. (2006) An application of eye gaze tracking for designing radiologists' workstations: Insights for comparative visual search tasks. *ACM Transactions on Applied Perception* (TAP), 3(2).

[Burmistrov et al., 2015] Burmistrov I., Zlokazova T., Izmalkova A., Leonova A. (2015). Flat design vs traditional design: Comparative experimental study.*Human-Computer Interaction*, (INTERACT-2015):C_ID 700.

[Eckstein, 2011] Eckstein, M. P. (2011) Visual search: A retrospective // Journal of Vision. — V.11(5). – P.1–36.

[Faraday, 2000] Faraday P. (2000) Visually Critiquing Web Pages. *Proceedings of the 6th Conference on Human Factors the Web*, 1-12.

[Garcia et al., 1994] García M., Badre A. N., Stasko J. T. (1994) Development and validation of icons varying in their abstractness. *Interacting with Computers*, 6 (2), 191–211.

[Goonetilleke et al., 2001] Goonetilleke R. S., Martins Shih H., Kai on H., Fritsch J. (2001) Effects of training and representational characteristics in icon design. *International Journal of Human-Computer Studies*, 55 (5), 741–760.

[Hall-Phillips et al., 2013] Hall-Phillips A., Yang R., Djamasbi S. (2013) Do ads matter? An exploration of web search behavior, visual hierarchy, and search engine results pages. *HICSS '13 Proceedings of the 2013 46th Hawaii International Conference on System Sciences*, 1563-1568.

[Halverson, Hornof, 2011] Halverson T., Hornof F.J. (2011) A Computational Model of ‘‘Active Vision’’ for Visual Search in Human–Computer Interaction. *Human–Computer Interaction*, 26, 285–314.

[Hou, Lu ,Ho, 2011] Hou K. C. H., Lu Y. N., Ho C. H. (2011) Balance between Beauty and Usability in Kansei Interface for Smartphone. *IASDR 2011 4th World Conference on Design Research*, pp. 1–6.

[Hout, Goldinge, 2015] Hout, M.C., Goldinge, S.D. Target templates: the precision of mental representations affects attentional guidance and decision-making in visual search// Attention, Perception, Psychophysics. 2015 January ; 77(1): 128–149.

[Malcolm,Henderson, 2009] Malcolm, G. L.,Henderson, J. M. (2009). The effects of target template specificity on visual search in real-world scenes: Evidence from eye movements. Journal of Vision, 9(11):8, 1–13.

[Malinowski, Hübner, 2001] Malinowski, P.,Hübner, R. (2001). The effect of familiarity on visualsearch performance: Evidence for learned basic features. Perception & Psychophysics, 63, 458–463.

[Pan et al., 2004] Pan, B., Hembrooke, H.A., Gay, G.K., Granka, L.A., Feusner, M.K., Newman, J.K. (2004) The Determinants of Web Page Viewing Behavior: An Eye-Tracking Study. *Proceedings of the 2004 symposium on Eye tracking research & applications,* pp. 147-154.

[Pogue, 2013] Pogue D. (2013) Out with the Real. *Scientific American*, 308 (2), 29–29.

[Qin, Koutstaal, Engel, 2014] Qin, X. A., Koutstaal, W., & Engel, S. A. (2014). The hard-won benefits of familiarity in visual search: Naturally familar brand logos are found faster. Attention, Perception, and Psychophysics, 76, 914-930.

[Rhodes at al, 2003] Rhodes, G., Jeffery, L., Watson, T. L., Clifford, C. W. G., & Nakayama, K. (2003). Fitting the mind to the world: Face adaptation and attractiveness aftereffects. Psychological Science, 14, 558–566.

[Salimun et al, 2012] Salimun, C., Purchase H.C., SimmonsD. (2012) Visual aesthetics in computer interface design: does it matter? In: 34th European Conference on Visual Perception, Toulouse, France, 28 Aug-1 Sept 2011,

[Shen, Reingold, 2001] Shen J., Reingold EM. (2001) Visual search asymmetry: the influence of stimulus familiarity and low-level features //Perception & Psychophysics. 63(3):464-75.

[Sutcliffe, Ennis, 1998] Sutcliffe A.G., Ennis M. (1998) Towards a cognitive theory of information Retrieval. *J Interacting with Computers*, 10, pp. 321–351.

[Treisman, Gelade, 1980] Treisman, A., Gelade, G. (1980). A feature-integration theory of attention. Cognitie Psychology, 12 , 97 –136.

[Treisman, Vieira, Hayes, 1992] Treisman, A., Vieira, A., Hayes, A. (1992) Automaticity and preattentive processing. American Journal of. Psychology, 105, 341-362.

[Q. Wang, P. Cavanagh, M. Green, 1994] Wang, Q., Cavanagh, P., & Green, M. (1994). Familiarity and popout in visual search. Perception & Psychophysics, 56 , 495–500.

[Wolfe, 1998] Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), Attention (pp. 13-73). Hove, U.K.: Psychology Press.

Authors' Information

Irina V.Blinnikova *– senior researcher*

Moscow State Lomonosov University, Faculty of psychology, Mokhovaya st., 11/9 Moscow 125009 Russia, e-mail: blinnikovamslu@hotmail.com

Major Fields of Scientific Research*: cognitive psychology, spatial cognition, visual experience, emotional effects, eye movement, mental chronometry*

Anna I.Izmalkova*- researcher*

Moscow State Lomonosov University, Faculty of psychology, Mokhovaya st., 11/9 Moscow 125009 Russia, e-mail: mayoran@mail.ru

Major Fields of Scientific Research*: cognitive psychology, verbal cognition, levels of cognitive processing, bilingual lexicon, eye movement*

Maria A.Semenova*- postgraduate student,*

Moscow State Lomonosov University Faculty of psychology, Mokhovaya st., 11/9 Moscow 125009 Russia, e-mail: 154mmi223@gmail.com

Major Fields of Scientific Research*: cognitive ergonomics, human-computer interaction, visual search, eye movement*

INTELLIGENT TECHNIQUES FOR PREDICTION OF CYTOSINE-PHOSPHATE-GUANINE (CPG) SITES ASSOCIATED WITH LUNG CARCINOMA

Abdel-Badeeh Salem, Mohamed Gawish, Mohammed Maher, Yasmin Amr, Amany Hussein, Keryakous Zarif, Basant Mohamed, Hebatullah Mohamed

Abstract: *DNA Methylation is a process by which cell assure the regulation of gene expression. Improved methods of detecting methylated sites are needed, instead of experimental methods which are very expensive and time consuming. In this paper, metaheuristic techniques namely; Genetic Algorithm, Artificial Immune System, and Hybrid Immune Genetic Algorithm are implemented to solve the problem of feature selection and select the susceptible CpG sites from dataset. Reducing the dimensionality of the dataset by applying previous algorithms resulting the following sets. After running the three algorithms many numbers of iterations, the average number of CpG sites determined by each algorithm is found to be less than 10% of the original dataset size. A new signature set was created by gathering all the common CpG sites from the three generated sets. Its size is equal to 0.1% of the original dataset size. Then it is used to generate the proteins regulatory network.*

Keywords: *CpG sites, Genetic Algorithm, Artificial Immune System, Clonal selection Algorithm, Hybrid Algorithms, Kruskal Wallis Test, Bioinformatics.*

ACM Classification Keywords: *I.5.2 Computing Methodologies - Pattern Recognition - Design Methodology-Feature evaluation and selection; I.2.8 Computing Methodologies – Artificial Intelligent - Problem Solving, Control Methods, and Search - Heuristic methods*

Introduction

DNA is a code of life; it is a polymer of four simple nucleic acids units called nucleotides. It consists of four nucleotide bases Adenine (A), Guanine (G), Cytosine (C) and Thymine (T). It is responsible of passing of genetic data between generations. CpG sites are where a Cytosine nucleotide exists next to a Guanine nucleotide separated by a phosphate group. DNA Methylation is one of the main factors causing gene silencing leading to an epigenetic change. Epigenetics involve inactivation of tumor suppressors and activation of oncogenes (Aine et al., 2015; Waterland & Michels, 2007). DNA Methylation involves addition of a methyl (CH3) group covalently at carbon 5 in the pyrimidine ring of a cytosine base, DNA methylation generally occurs in the context (5'-CG-3')dinucleotides, the

methylation pattern differs from one cell to another (depending on the functions which the cell needs to be active) and from a disease to another(Ahn & Wang, 2013; Meng, Murrelle, & Li, 2008).

CpG islands are present on the promoter of genes, found un-methylated in a normal state but methylated in Cancer with presence that's five times larger than the normal state. Distinguishing the methylated Cytosine (5mC) is based on principles like Bisulfite conversion which differentiates between methylated and un-methylated Cytosine by treating the DNA with sodium Bisulfite while un-methylated Cytosine turn into Uracil , methylated Cytosine are not affected, the changes resulting in the DNA sequence can be detected through PCR amplification proceeded by DNA sequencing as (Guo et al., 2015).

A number of CpG sites in a high throughput methylation arrays are irrelevant and don't provide information to distinguish the normal cells of cells with Cancer. In this paper efficient computational intelligence techniques such as Genetic Algorithm, Artificial Immune System, and Hybrid Immune Genetic Algorithm have been utilized to reduce the number of CpG sites resulted from a high throughput methylation array resulting in CpG sites most likely causing Cancer. Although wrapper methods generally outperform filter methods, they are computationally intensive and may become inefficient in practice for large datasets (Guyon, Weston, Barnhill, & Vapnik, 2002; Li, Zhang, & Ogihara, 2004; Zhang et al., 2006).

This paper is organized in the following manner. First we show existing work in predicting CpG sites. Then we present the three techniques used on our work for predicting CpG sites. Experimental design, results, and statistical analysis are presented afterwards. Followed by a section that provides discussion. Finally, conclusions is provided.

Related Work

There are two types of feature selection methods. The first type is filter-based methods that assess the relevance of features by looking only at the intrinsic properties of the date. Filter-based methods are quite popular because they are more efficient, more scalable, and independent of the classification algorithm. On the other hand, they have limitations and the classification accuracy of the selected genes is less accurate. The other type of feature selection methods is the wrapper methods, which employ classifiers to determine feature selection based on the predictive accuracy of the classifier(Guyon et al., 2002; Li et al., 2004; Zhang et al., 2006).

One work that is worth mentioning here is the work done by Model, Adorjan, Olek, and Piepenbrock (2001). In that work, the simple Fisher criterion was used as a feature selection strategy combined with SVM in order to discriminate between acute lymphoblastic leukemia and acute myeloid leukemia using methylation pattern data.

Another work is done by Meng, Murrelle, and Li (2008). In that work a two-stage feature selection method was developed to select a small optimal subset of DNA methylation feature to distinguish lung cancer tissue samples from normal lung tissue samples using DNA methylation data.

Current work

In this work, the problem of predicting CpG sites has been tackled by three metaheuristic techniques (Genetic Algorithm, Artificial Immune System, and Hybrid Immune Genetic Algorithm).As shown in Figure 1, the three techniques are applied on the same dataset to get three different sets of CpG sites sets(Kim, Park, & Kon, 2013; Xu & Zhang, 2005). After that, a new set called signature set was created by gathering the common CpG sites of the three generated sets in order to find the most related CpG sites to lung carcinoma. Finally, the signature set was used to produce the proteins regulatory network.

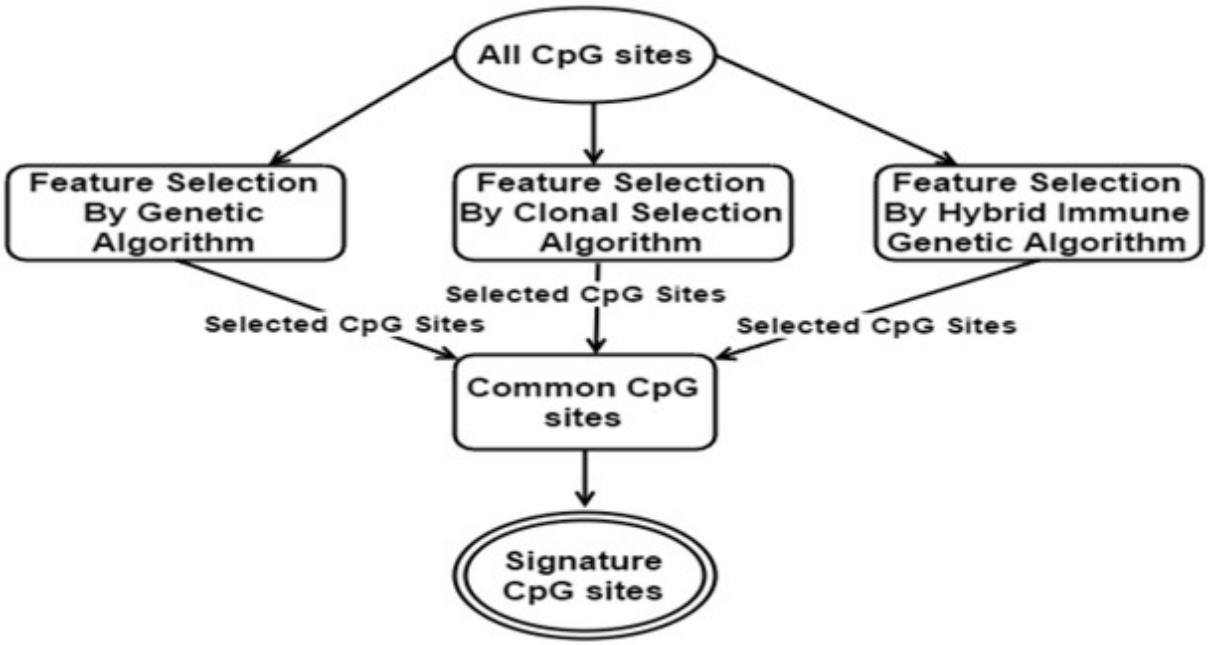


Figure 1.Generating the signature CpG sites

Solving CpG sites selection by Genetic Algorithms

Overview

In the early 1970s, John Holland introduced the concept of genetic algorithm (GA) as a class of evolutionary algorithms, which generate solutions to optimization problems using techniques inspired by Darwin's classical theory of natural evolution, such as inheritance, mutation, selection, and crossover.

GA steps

Figure 2 presents the steps of GAs (Jourdan, Dhaenens, & Talbi, 2001; Vafaie & De Jong, 1992),

- Step 1: initialize population of chromosomes randomly: The population has a set of binary chromosomes that are of fixed size. The size equals the number of features in the dataset, Fill chromosomes randomly 0 or 1; 0 means that feature is deactivated and 1 means that feature is activated.
- Step 2: evaluate fitness for each chromosome by getting all active feature (1's index) in this chromosome, then create subset dataset by getting expressed data for these features from the original dataset then pass it to support vector machine to calculate validation error. The fitness of this chromosome is 1/validation error (Guyon et al., 2002; Kim et al., 2013).
- Step 3: select two chromosomes for reproduction by using roulette wheel (Banzhaf & others, 1999).
- Step 4: Crossover between the two selected chromosomes, single point crossover is applied to get offspring as illustrated in Figure 3.
- Step 5: Bit-Flip mutation for offspring as illustrated in Figure 4.
- Step 6: Replacement by using elitist strategy by evaluating fitness for each offspring; if its fitness greater than its parents, replace.
- Step 7: Repeat Steps 2 through 6 while the validation error is minimum.

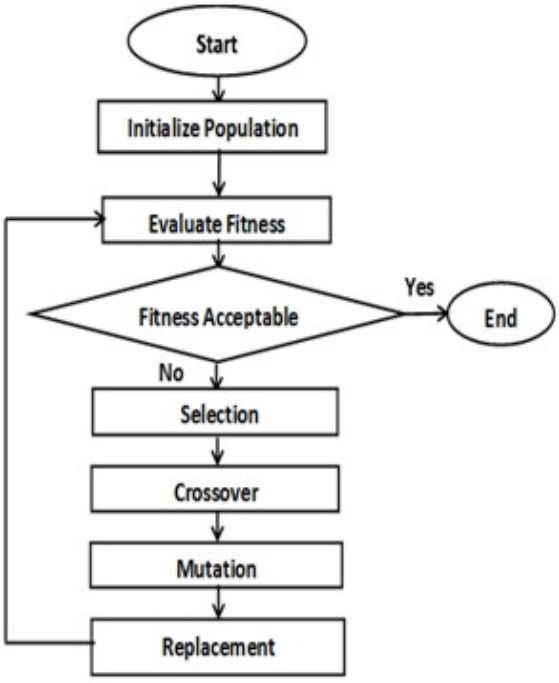


Figure 2. GA flowchart

Crossover

Applying one point crossover on selected chromosomes to get offsprings by generating numbers R1 and R2 randomly then compare if R2 is smaller than or equal to probability of crossover (0.5) do crossover between parents to get offsprings, otherwise, offsprings equal parents.

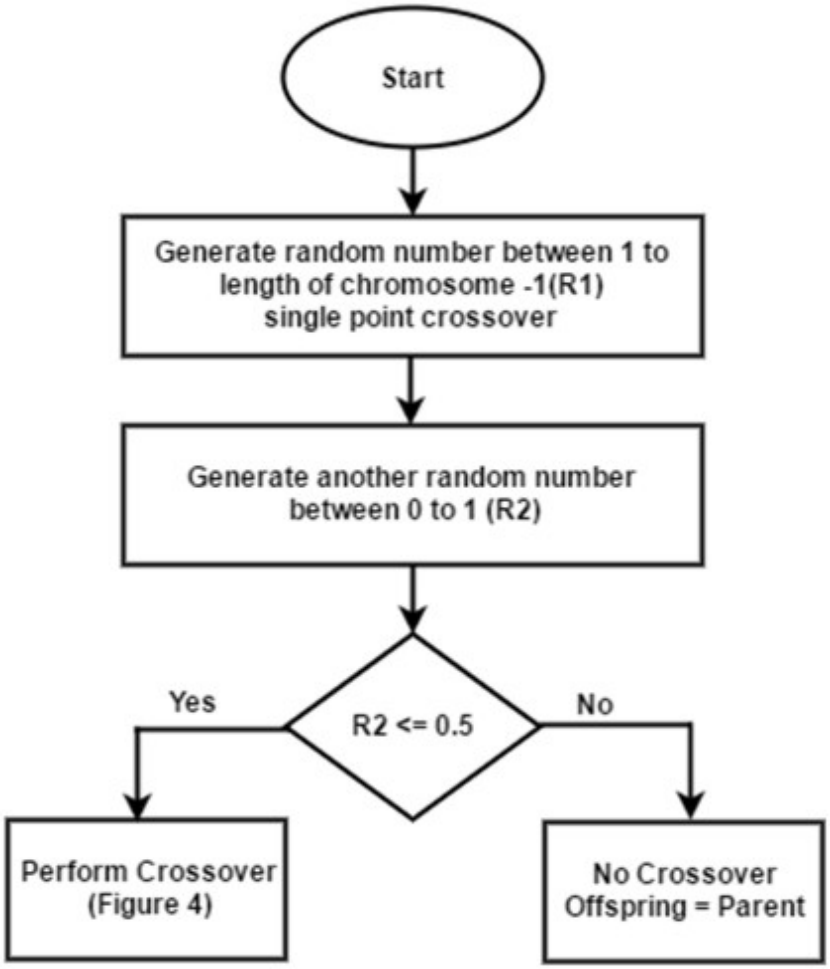


Figure 3. Crossover flowchart

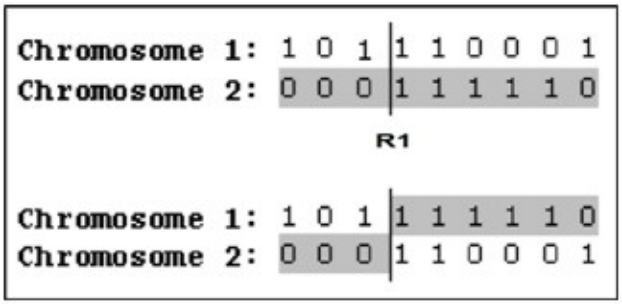


Figure 4. One-Point Crossover

Mutation

Two different probabilities are applied one for mutating to one (small value 0.05) and the other for mutating to zero (large value 0.5) ;to reduce the amount of 1s to get the final result faster as illustrated in Figure 5 .

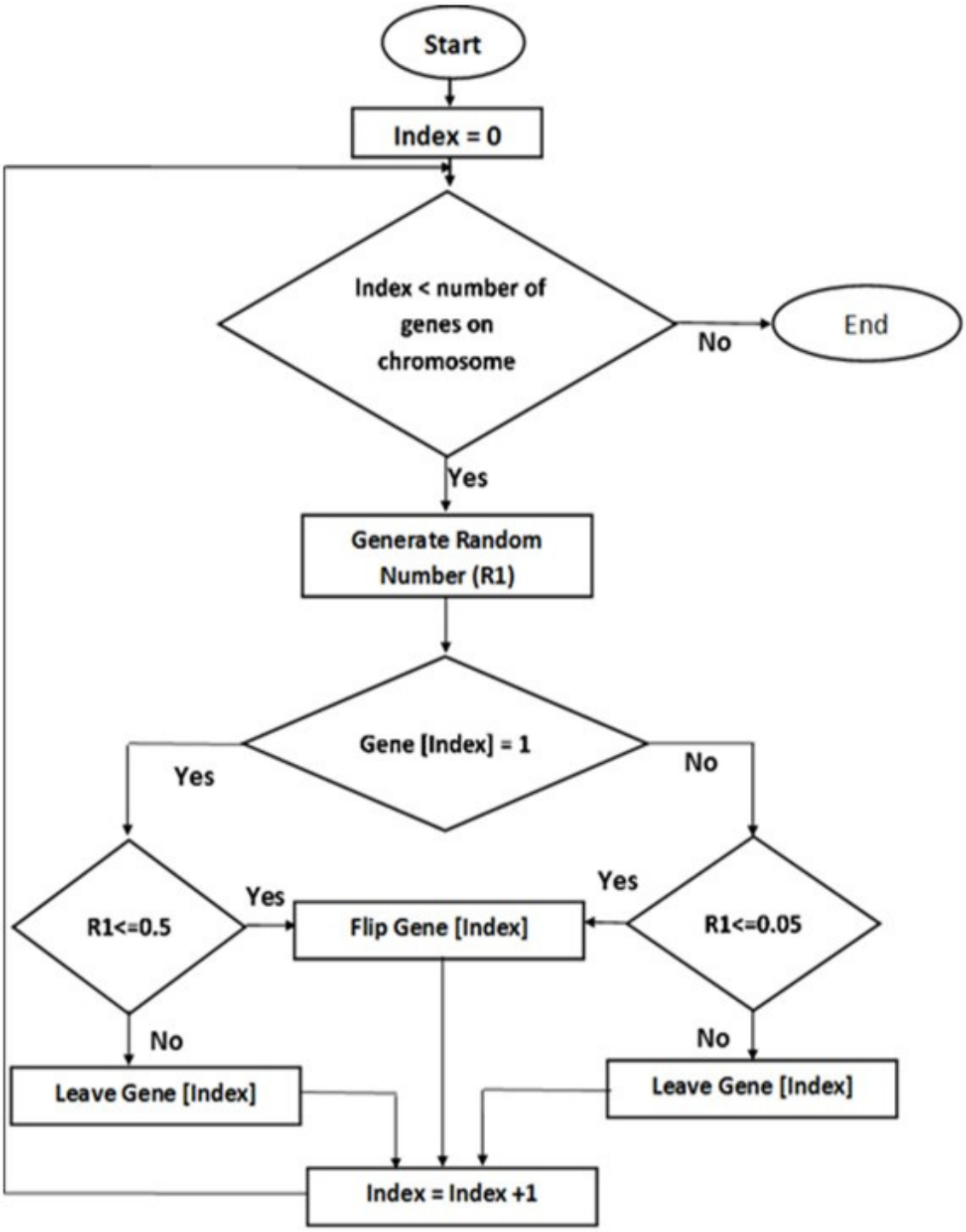


Figure 5. Mutation flowchart

Replacement

The fitness of the offspring is calculated, and then if it's higher than the fitness of the parents, the parents will be replaced by the offspring to get the new population which has a higher fitness as in genetic algorithm the fittest only can survive.

Solving CpG sites selection by Artificial Immune System

Overview

Artificial Immune System (AIS) is a metaheuristic search inspired by the theoretical immunology and observed immune functions principles and models, which are applied to complex problem domains. Its flowchart is presented in Figure 6.

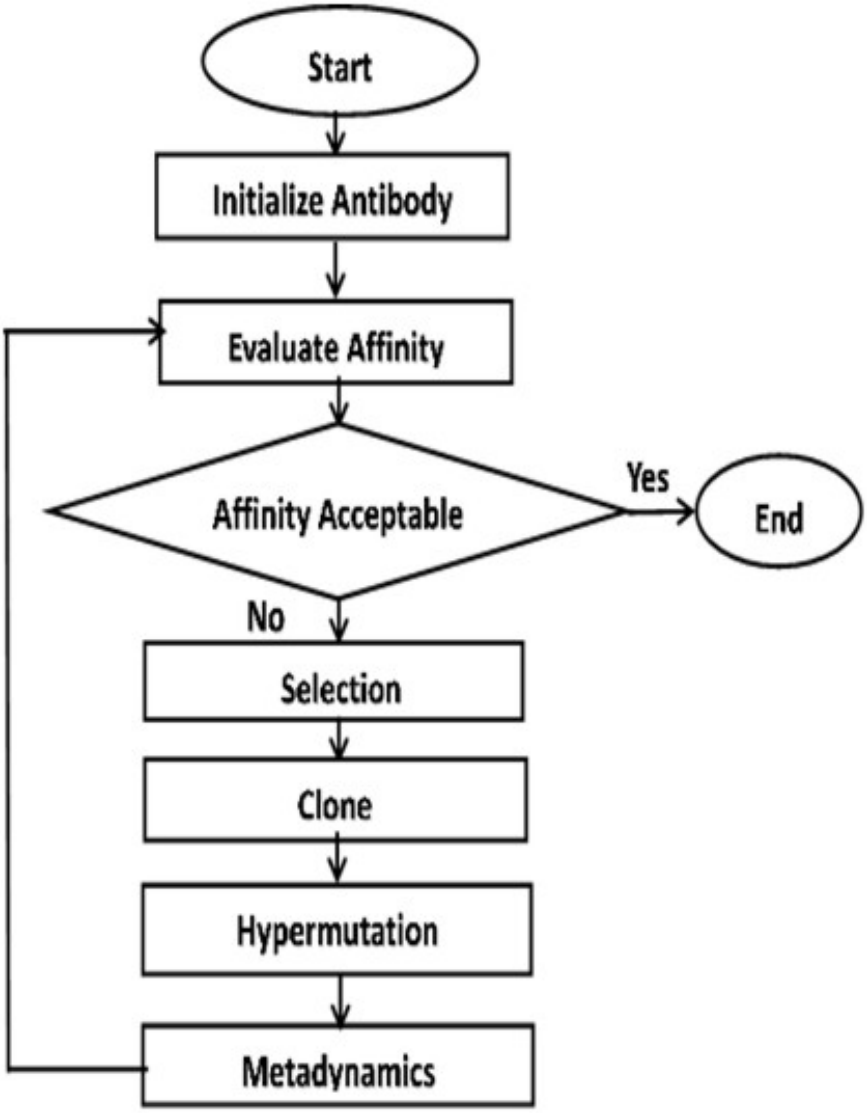


Figure 6. AIS flowchart

AIS Steps

- Step 1: initialize population of antibodies randomly: The population has a set of binary antibodies that are of fixed size. The size equal to the number of features in the dataset. Fill antibodies randomly 0 or 1; 0 means the feature is not active and 1 means the feature is active.
- Step 2: evaluate affinity for each antibody by getting all active feature (1's index) in this antibody then create subset dataset by getting expressed data for these features from the original dataset then pass it to Support vector machine to calculate validation error, the affinity of the antibody is 1/validation error(Guyon et al., 2002; Kim et al., 2013).
- Step 3: select populations for Lowest 20 percentage of Affinity that is better matching between antibody and antigen
- Step 4: Copy lowest 20 percentage to replace it in the next lowest 20 percentage of population
- Step 5: Hybermutation for selected and cloned (result on step 3 and 4) 40 percentage; by generating a random number; if the random number is less than probability of mutation flip bit.
- Step 6: Randomize lowest 60 percentage of population that's called Metadynamics
- Step 7: Repeat Steps 2 through 6 while the validation error is minimum.

Solving CpG sites selection by Hybrid Immune Genetic Algorithm

For increasing the exploration of a search space, Hybrid Immune Genetic Algorithm (HIGA) is proposed by Nabil, Badr, and Farag (2009), which is a result of hybridizing AIS with the GA's crossover operator. Its flowchart is presented in Figure 7.

HIGA steps

- Step 1: initialize population of antibodies randomly: The population has a set of binary antibodies that are of fixed size. The size equal to the number of features in the dataset. Fill antibodies randomly 0 or 1; 0 means the feature is not active and 1 means the feature is active.
- Step 2: evaluate affinity for each antibody by getting all active feature (1's index) in this antibody then create subset dataset by getting expressed data for these features from the original dataset then pass it to Support vector machine to calculate validation error, the affinity of the antibody is 1/validation error(Guyon et al., 2002; Kim et al., 2013).
- Step 3 select populations for Lowest 20 percentage of Affinity that is better matching between antibody and antigen
- Step 4: Copy lowest 20 percentage to replace it in the next lowest 20 percentage of population

- Step 5: Hybermutation for selected and cloned (result on step 3 and 4) 40 percentage; by generating a random number; if the random number is less than probability of mutation, flip bit.
- Step 6: Do Crossover on selected population after Hybermutation not cloned population and replace the selected population by the new offspring.
- Step 7: Randomize lowest 60 percentage of population that's called Metadynamics
- Step 8: Repeat Steps 2 through 7 while the validation error is minimum.
-

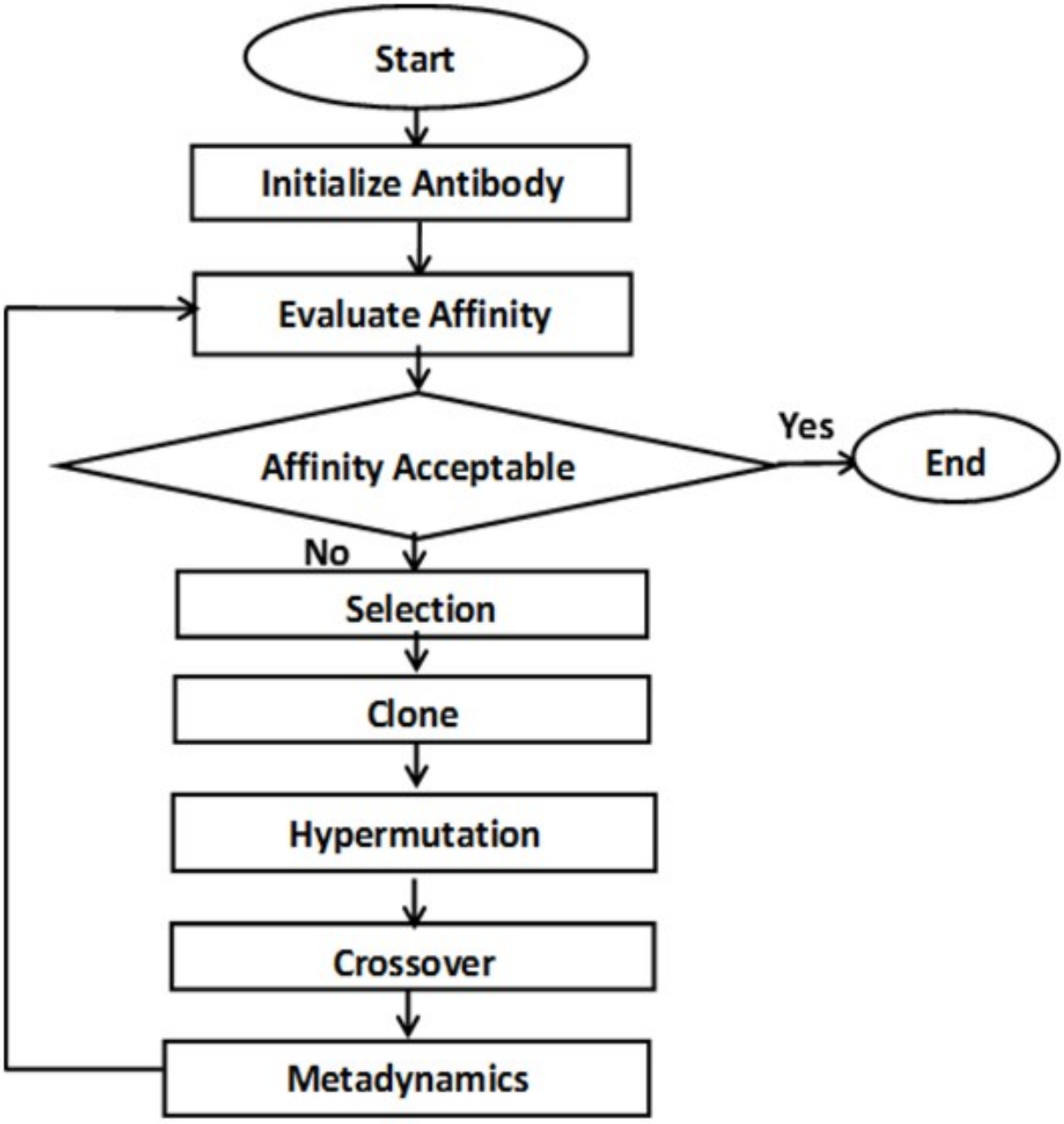


Figure 7. HIGA flowchart

The evaluation function

The same evaluation function was used in the three metaheuristic techniques in order to calculate the fitness or affinity value. It takes the dataset and the binary (chromosome/antibody) as inputs. The binary (chromosome/antibody) are used to select a subset from the original dataset by choosing only features that has a corresponding active value in the (chromosome /antibody).

The SVM that is one of state-of-the-art classification method was chosen to calculate the test error. It has been widely used in microarray data analysis (Guyon et al., 2002).

Leave-one-out cross-validation was employed to evaluate the classification performance of each subset. Each sample was excluded from the training set, one at a time, and then classified based on the SVM trained from the remaining samples. This procedure was repeated, in turn, for all samples, and the cross-validation error was defined as the sum of misclassifications. Finally the function returns the cross-validation error as its output.

Finding the Signature CpG sites

After applying the previous three metaheuristic techniques on the same dataset, three different sets of CpG sites are generated. In order to find the smallest and most important CpG sites set, a new set (signature) was created that has only the common CpG sites of the three generated sets as depicted in Figure1.

Generating the regulatory network

After generating the signature CpG sites, the dataset has been used in order to return the related genes (Network & others, 2014). A set of proteins that relate to the genes are entered into string-db to produce the proteins regulatory network as shown in Figure 8(Szklarczyk et al., 2010).

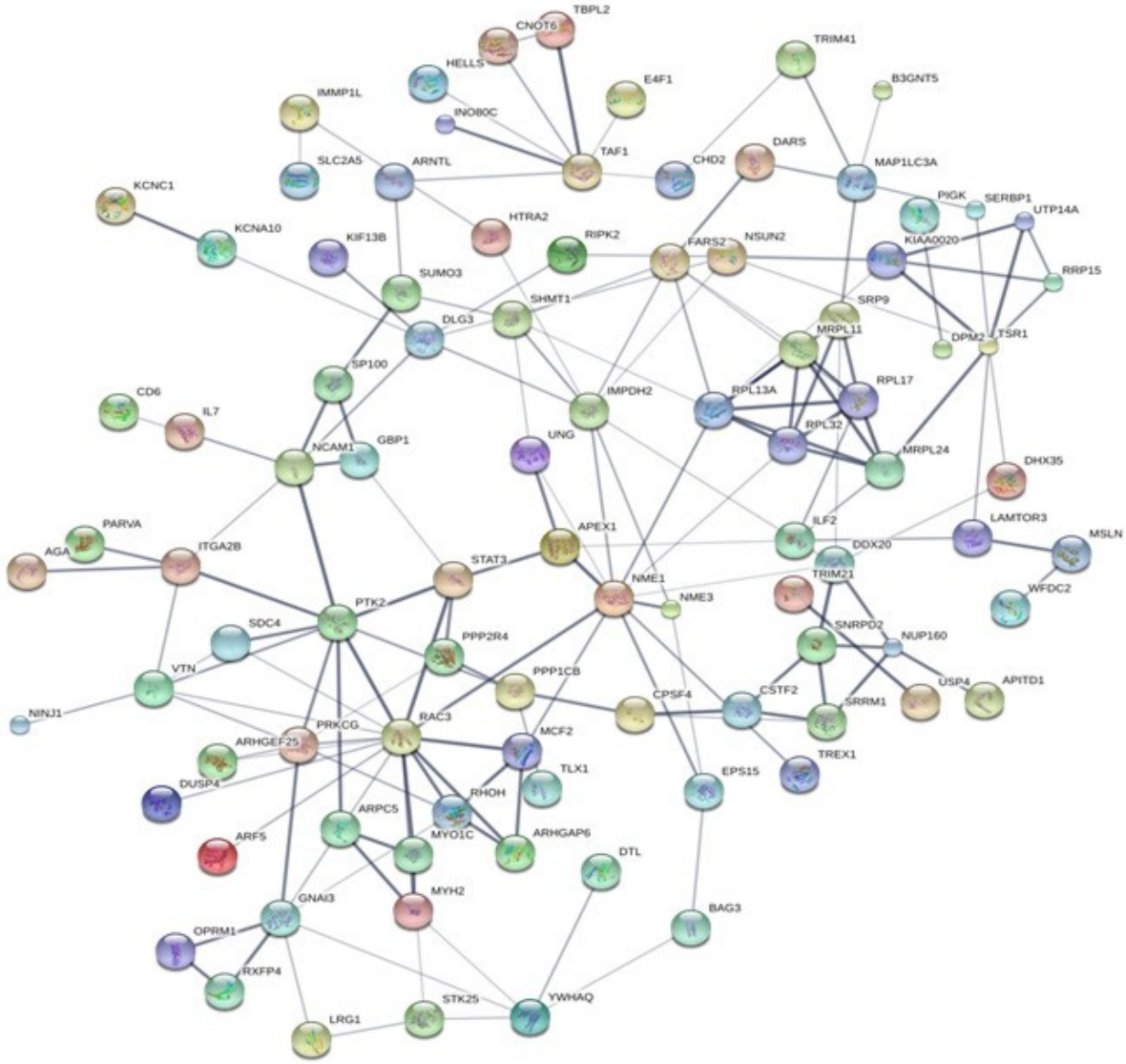


Figure 8. The proteins regulatory network

Results and Discussion

Dataset

The dataset used in this work is collected through the analysis of tumor and matched normal material from previously untreated lung adenocarcinoma patients (Network & others, 2014). And it is extracted using an Illumina array platform which is based on Bisulfite conversion, with an accuracy of 99.9% which can detect a 17% difference in DNA methylation and 2.5% DNA methylation as (Müller, Assenov, & Lutsik, 2015).It consists of twenty persons; each person has twenty seven thousand (27,000) CpG sites, and labeled by the predicted output, as the CpG sites of the highest fitness only can survive.

Array based genome wide Methylation analysis methods based on Bisulfite Conversion generates high throughput of about 27000 assays per sample, also known for their ability to derive information at single nucleotide resolution which make these methods very specific .

Parameter settings

In the used metaheuristic techniques several parameters have to be assigned. The maximum number of generations is set to 500 for GA, AIS, and HIGA. For all techniques, the population's size is set to 10.Both GA's crossover probability and HIGA's crossover probability are set to 0.5. The [1 to 0] mutation probability is set to 0.5. The [0 to 1] mutation probability is set to 0.05 as illustrated in table 1.Each technique has been run 100 times.

Table 1. Parameter Settings

Parameter name	GA	AIS	HIGA
Population size	10	10	10
Max No. of generations	500	500	500
Cross over probability	0.5	-	0.5
[1->0] mutation probability	0.5	0.5	0.5
[0->1] mutation probability	0.05	0.05	0.05
No. of run	100	100	100

Experimental Results

In the experiment, the performance of the used techniques has been evaluated over the dataset using SVM classifier. The comparison between the different techniques as shown in Figure8is in term of the best accuracy classification returned by SVM. It can be seen from Figure 9 that the HIGA technique obtains the best accuracy classification compared to the other techniques. The GA technique obtains the second best accuracy classification. Finally the AIS technique comes at the end.

After running the three algorithms for 100 iterations, the average number of CpG sites determined by the Genetic algorithm found to be from 2050 ~2100 , And Clonal selection algorithm from 2100~ 2200 and Immune genetic hybrid algorithm from 2000~2065 , The Combination between 3 algorithms resulted 25CpG site.

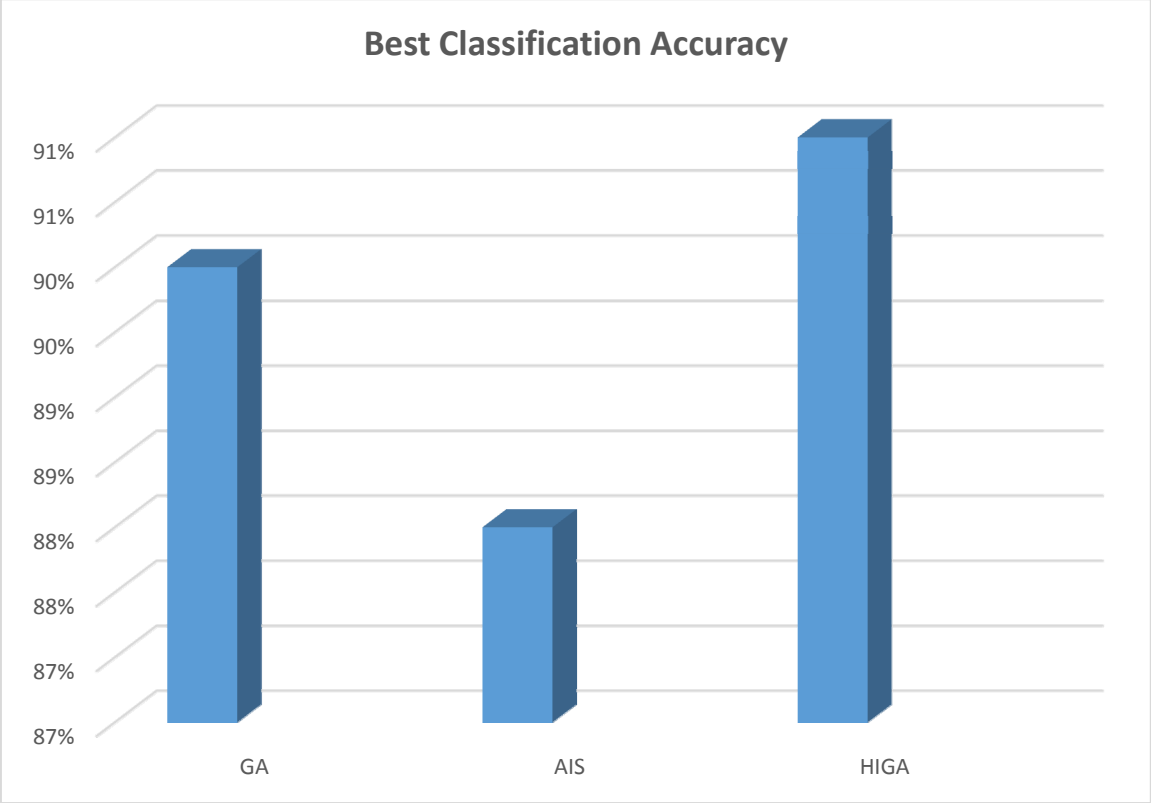


Figure 9. The best classification accuracy

Statistical analysis

In order to illustrate that the experimental results are statistically significant, the Kruskal-Wallis test has been performed on the results. The Kruskal-Wallis test is a rank-based nonparametric test. It can be used to determine if there are statistically significant differences between two or more groups of an independent variable on a continuous or ordinal dependent variable (McKight & Najab, 2010).

Conclusion and Future Work

Features selection is necessary in order to get a reliable predication results by using a small size training datasets. So, this work has tackled the problem of features selection by using three different techniques (GA, AIS, and HIGA).The three techniques have produced three different subsets each one contains different features. In term of best classification accuracy, the HIGA algorithm has outperformed both GA and AIS.

In order to get the most relevant features a new subset (signature) was created by gathering the common features from the three subsets. The proteins regulatory network that relevant to the signature subset was produced using the string-db.

Protein ligand interaction can also be investigated in conjunction with protein-protein interaction. Likewise, other metaheuristics can be investigated such as cuckoo search bat algorithm, firefly algorithm, differential evolution and others.

This study can be extended to the Four hundred Fifty thousand CpGs once a superior hardware is available. This incurs to the use of deep learning neural networks such as Convolutional neural network, deep belief networks and restricted Boltzmann machines, this is because of the huge amount of features (450,000) that will necessitate the use of deep learning.

Acknowledgement

"The paper is published with partial support by the ITHEA ISS (www.ithea.org) and the ADUIS (www.aduis.com.ua)".

Bibliography

[Ahn & Wang ,2013]. Ahn S. & Wang T., „A powerful statistical method for identifying differentially methylated markers in complex diseases'' , *Pacific Symposium on Biocomputing , p.69.*

[Aine, Sjödahl, Eriksson, Veerla, Lindgren, Ringnér, Höglund, 2015], Aine M., Sjödahl G., Eriksson P., Veerla S., Lindgren D., Ringnér M. & Höglund M., „Integrative epigenomic analysis of differential DNA methylation in urothelial carcinoma'', Genome mid, 7(1), 23.

[Banzhaf & others,1999] *Foundations of genetic algorithms';* Morgan Kaufmann Publishers Inc.

[Guo, Yan, Xu, Bao, Zhu, Wang, … others, 2015] . Guo, S.,Yan, F., Xu, J., Bao, Y., Zhu, J., Wang, X., … others, „ Identification and validation of the methylation biomarkers of non-small cell lung cancer (NSCLC)'', *Clinical Epigenetics*, *7*(1), 1–10.

[Guyon, Weston, Barnhill & Vapnik, 2002]. Guyon, I., Weston, J., Barnhill, S. & Vapnik, V., „ Gene selection for cancer classification using support vector machines'' *Machine Learning*, *46*(1-3), 389–422.

[Jourdan, Dhaenens & Talbi, 2001]. Jourdan, L., Dhaenens, C. & Talbi, E.-G, „ A genetic algorithm for feature selection in data-mining for genetics'', *Proceedings of the 4th Metaheuristics International ConferencePorto,(MIC'2001)*, 29–34.

[Kim, Park & Kon, 2013], Kim, S., Park, T. & Kon, M. A, „Computational methods for cancer survival classification using intermediate information'' , *IWBBIO* (pp. 517–525).

[Li, Zhang & Ogihara, , 2004]. Li, T., Zhang, C. & Ogihara, M. , „A comparative study of feature selection and multiclass classification methods for tissue classification based on gene expression.'', *Bioinformatics*, *20*(15), 2429–2437.

[McKight & Najab, 2010]. McKight, P. E. & Najab, J. „ Kruskal-Wallis Test''. *Corsini Encyclopedia of psychology.*

[Meng, Murrelle & Li, 2008]. Meng, H., Murrelle, E. L. & Li, G. , „Identification of a small optimal subset of CpG sites as bio-markers from high-throughput DNA methylation profiles.'', *BMC Bioinformatics*, *9*(1), 457.

[Model, Adorjan, Olek & Piepenbrock, 2001]. Model, F., Adorjan, P., Olek, A. & Piepenbrock, C. , „Feature selection for DNA methylation based cancer classification.'' ,*Bioinformatics*, *17*(suppl 1), S157–S164.

[Müller, Assenov & Lutsik, 2015]. Müller, F., Assenov, Y. & Lutsik, P.„RnBeads-Comprehesive Analysis of DNA methylation Data.''

[Nabil, Badr & Farag, 2009]. Nabil, E., Badr, A., & Farag, I.„ An immuno-genetic hybrid algorithm'',Int J Comput Commun Control, 4(4), 374-385.

[Network & others, 2014]. Network, C. G. A. R. & others „Comprehensive molecular profiling of lung adenocarcinoma''. *Nature*, *511*(7511), 543–550.

[Szklarczyk, Franceschini, Kuhn, Simonovic, Roth, Minguez, 2010]. Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguez, P., … others. , „ The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored'', *Nucleic Acids Research*, gkq973.

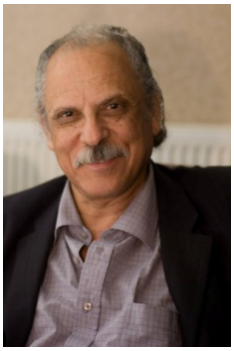
[Vafaie & De Jong, 1992]. Vafaie, H. & De Jong, K. „ Genetic algorithms as a tool for feature selection in machine learning. In *Tools with Artificial Intelligence.'', 1992, TAI'92, Proceedings., Fourth International Conference on* (pp. 200–203).

[Waterland & Michels, 2007]. Waterland, R. A. & Michels, K. B. „ Epigenetic epidemiology of the developmental origins hypothesis.'', *Annu . Rev . Nutr.*, *27*, 363–388.

[Xu & Zhang, 2005]. Xu, X. & Zhang, A. , „Selecting informative genes from microarray dataset by incorporating gene ontology.'' , In Bioinformatics and Bioengineering, 2005, BIBE 2005, Fifth IEEE Symposium on , (pp. 241–245).

[Zhang, Lu, Shi, Xu, Hon-chiu, Harris, Wong, 2006]. Zhang, X., Lu, X., Shi, Q., Xu, X., Hon-chiu, E. L., Harris, L. N., … Wong, W. H. , „Recursive SVM feature selection and sample classification for mass-spectrometry and microarray data.'', BMC Bioinformatics, 7(1), 1.

Author's information



Abdel-Badeeh M. Salem - *Professor of Computer Science, Head of Artificial Intelligence and Knowledge Engineering Research Labs,*

Faculty of Computer and Information sciences,

Ain Shams University, Cairo, Egypt .

e-mail: abmsalem@yahoo.com

Mohamed Gawesh - *Teacher assistant of faculty of computer and information science,*

Ain shams university, Cairo, Egypt.

e-mail: mygawish@cis.asu.edu.eg

Mohammed Maher - *Bsc student at Department of Bioinformatics*

Faculty of Computer and Information Sciences,

Ain Shams University, Cairo, Egypt.

e-mail: mohammed.maher2013@hotmail.com

Yasmin Amr - *Bsc student at Department of Bioinformatics*

Faculty of Computer and Information Sciences,

Ain Shams University, Cairo, Egypt.

e-mail: yasminamrfcis@gmail.com

Amany Hussein *- Bsc student at Department of Bioinformatics*

Faculty of Computer and Information Sciences,

Ain Shams University, Cairo, Egypt.

e-mail: amanyhusseinhasan@hotmail.com

Keryakous Zarif *- Bsc student at Department of Bioinformatics*

Faculty of Computer and Information Sciences,

Ain Shams University, Cairo, Egypt.

e-mail: keryakous.zarif@hotmail.com

Basant Mohamed *- Bsc student at Department of Bioinformatics*

Faculty of Computer and Information Sciences,

Ain Shams University, Cairo, Egypt.

e-mail: basant.mohamed.ramdan@fcis.asu.edu.eg

Hebatullah Mohamed *- Bsc student at Department of Bioinformatics*

Faculty of Computer and Information Sciences,

Ain Shams University, Cairo, Egypt.

e-mail: hebafouad.cs2016@gmail.com

POINT CLOUDS REGISTRATION AND GENERATION FROM STEREO IMAGES

Aram Gevorgyan Vladimir

Abstract*: Registration of 3D point sets is a basic problem in computer vision and 3D modeling. The goal of registration is to transform two or more overlapping point clouds into one common coordinate system. In this paper we discuss point clouds generation from stereo images and registration problem. We also present our realization, we capture object images from different angles with two cameras, then we generate point clouds from these stereo images and merge these point clouds to get 3D model of the object.*

Keywords*: point clouds, registration, stereovision, ICP, stereo image*

ACM Classification Keywords: *I.4.4 Image Processing and Computer Vision - Restoration*

1. Introduction

Point cloud is a set of points in some coordinate system. Point clouds are typically used to measure physical world surfaces. Point clouds are used for many purposes, including to create 3D CAD models for manufactured parts, and at a multiple of visualization, rendering, animation and mass customization applications. They have applications in robot navigation and perception, depth estimation, stereo vision, visual registration, and in advanced driver assistance systems (ADAS).

There are different ways of getting point clouds, such as, laser scanners, time-of-flight cameras, special 3D cameras, like PrimeSensor or Microsoft Kinect, or stereo cameras. In Section 2 we describe the process of getting point cloud from pair of images taken from two cameras simultaneously that is called stereo image or stereo pair.

For different purposes, such as 3D modeling or 3D scene reconstruction, it is usually needed to merge point clouds from different scenes or partially overlapped point clouds into one common point cloud - this process is called registration. In other words, registration transforms multiple 3D point clouds into the same coordinate system so as to align overlapping components of these sets.

Point clouds registration is very important problem that is used in 3D modeling. In Section 3 we discuss registration problem and different registrations algorithms and in section 4 we present our realization of merging object point clouds from different views to get object 3D model.

2. Point cloud generation from stereo pair

Stereovision is a technique that estimates 3D information of a scene from two or more cameras. In the case of two cameras it is usually called binocular stereovision. One of the fundamental problems in stereovision is calculating disparity map which is also known as stereo correspondence problem. Disparity is the distance between correspondence points in the left and right images. The goal of stereo correspondence algorithms is to find correspondent points in the left and right images. There are plenty of stereo correspondence algorithms, good taxonomy and review of stereo correspondence algorithms is at [Sharstein, 2002].

Let us consider simple case (Figure 1): suppose we have two cameras with parallel optical axes, we know geometrical arrangement and internal parameters of the cameras and have disparity map calculated. So we can compute point's coordinates by triangulation.

The line connecting the cameras lens centers is called the baseline. Let the baseline be perpendicular to the optical axes of the cameras and parallel to x-axis. The distance between cameras is d and cameras have equal focal length f.

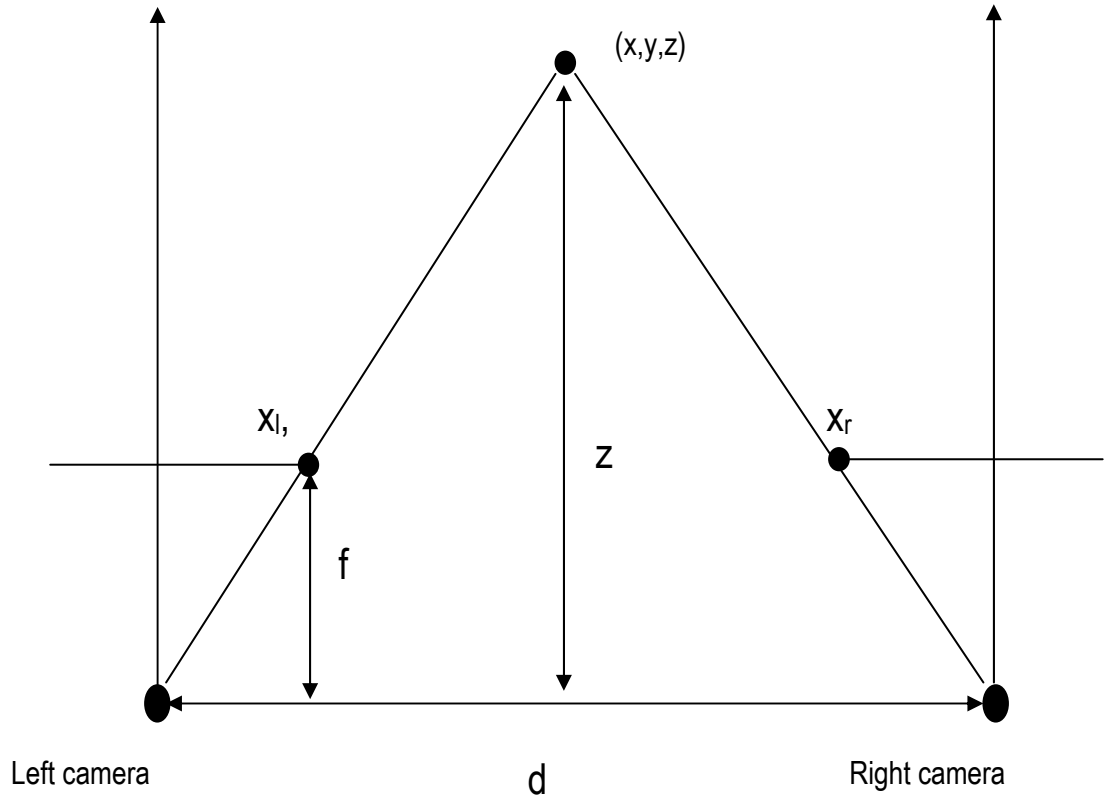


Figure 1. Two cameras with parallel optical axes

Consider (x, y, z) are coordinates of a point in three-dimensional world, where z is a depth. Let this point have (xl, yl) and (xr, yr) coordinates in the left and right image planes of the respective cameras.

Then from the similar triangles we have:

$$\frac{z}{d} = \frac{f}{x_l - x_r} \Rightarrow z = \frac{df}{x_l - x_r},$$

where $x_l - x_r$ is disparity. Thus, the depth at various scene points may be recovered by knowing the disparities of the corresponding image points.

Also we can compute:

$$x = \frac{d\left(x_l + x_r\right)}{2\left(x_l - x_r\right)}, \quad y = \frac{d\left(y_l + y_r\right)}{2\left(x_l - x_r\right)}.$$

So having these coordinates we can construct point cloud. In more details how to compute depth from stereo images is described at [Bagga, 2013], [Mahammed, 2013].

3. Registration

As it was mentioned, 3D registration is problem of merging overlapping point clouds and bringing them into one common coordinate system. There are two types of registration algorithms that uses rigid and non-rigid transformations. Rigid algorithms assume that point clouds may differ only by rotation and translation but preserve the distance between every pair of points, which means that transformation can be modeled using only 6 Degrees of Freedom (DOF). On the other hand, non-rigid approaches are able to cope with articulated objects or soft bodies that change shape over time.

Both rigid and non-rigid algorithms can be also classified as pairwise registration and multi-view registration methods. Pairwise registration algorithms calculate a rigid transformation between two subsequent point clouds while the multi-view registration process takes multiple point clouds into account to obtain better results.

Rigid transformations algorithms are usually divided into 2 classes: feature-based registration algorithms and iterative registration algorithms.

Feature-based registration algorithms use geometric features of the objects for registration. Usually consists of the following steps:

1) estimate the associated descriptors of the point clouds
2) correspondence calculation
3) reject bad correspondences
4) transformation estimation

Iterative algorithms that are Iterative Closest Point(ICP) algorithm and it's variants iteratively register points clouds, trying at each iteration minimize distance between point clouds. In contrast to feature-based registration, iterative registration algorithms do not match salient feature descriptors in order to find correspondences between source and target point clouds, but instead search for closest points and align the found point pairs. ICP algorithm consists of the following steps:

1) Selection subsets of two point clouds
2) Matching corresponding pairs
3) Weighting the corresponding pairs appropriately
4) Rejecting certain pairs based on looking at each pair individually or considering the entire set of pairs
5) Compute an error metric based on the point pairs
6) Minimizing the error metric

ICP algorithm needs aligned point clouds to have small rotation degree and to be near each other to achieve good results, it is also called fine registration. So it is frequently combined with feature-based registration algorithms, feature-based registration is used for initial alignment and then alignment is improved with ICP. More about ICP and its variants at [Besl, 1992], [Rusinkiewicz, 2001], [Johnson, 1997].

More about registration algorithms and their survey is given in [Bellekens, 2015].

4. Realization

We reconstruct object 3D model from the set of stereo pair images taken from the left and right cameras (see examples on Figure 2 and Figure 3). We realize our project in C++ using OpenCV [OpenCV], [Bradski, 2008] – the biggest image processing library for calculating disparity map and generating point clouds and Point Cloud Library (PCL) [PCL] – point clouds processing and viewing library for 3D registration.

We try our realization with image sets captured from two Logitech HD Pro Webcam C920 web cameras and Minoru 3D stereo camera [Minoru]. We place an object at about 40-50 cm in front of the cameras and rotate it at small degrees capturing images for each rotation.

Then for each stereo pair we calculate disparity map using Stereo-Global Matching (SBM) algorithm [Hirschmuller, 2005] that is one of the best stereo correspondence algorithms combining speed and accuracy. From disparity maps we generate point clouds and segment our object from the scene to have point cloud only of the object.

Then we register these point clouds. We perform pairwise registration with ICP. At first step we take the first point cloud as a global model. Then at each next step we take the next consecutive point cloud and merge it to the global model with ICP registration algorithm, transforming it to the global model's coordinates and adding to it. . We also try to combine ICP with feature-based registration algorithms, but it didn't give good results in our case.

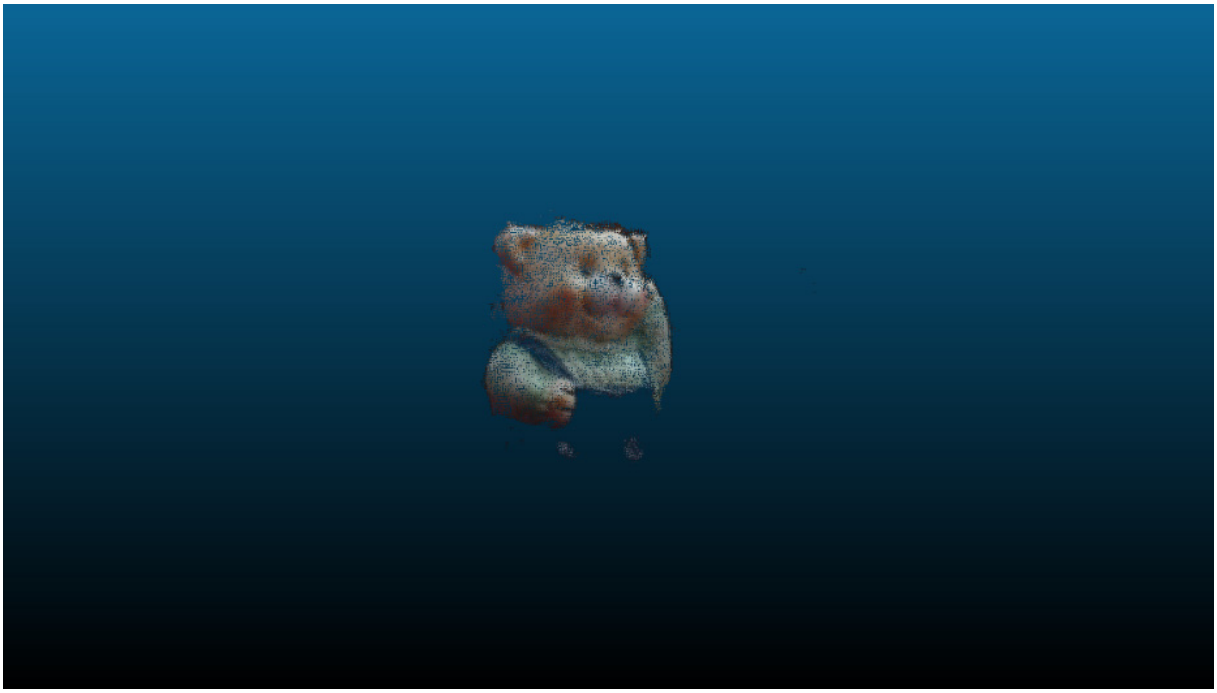


Figure 2. Reconstructed 3D model from different angles - example 1

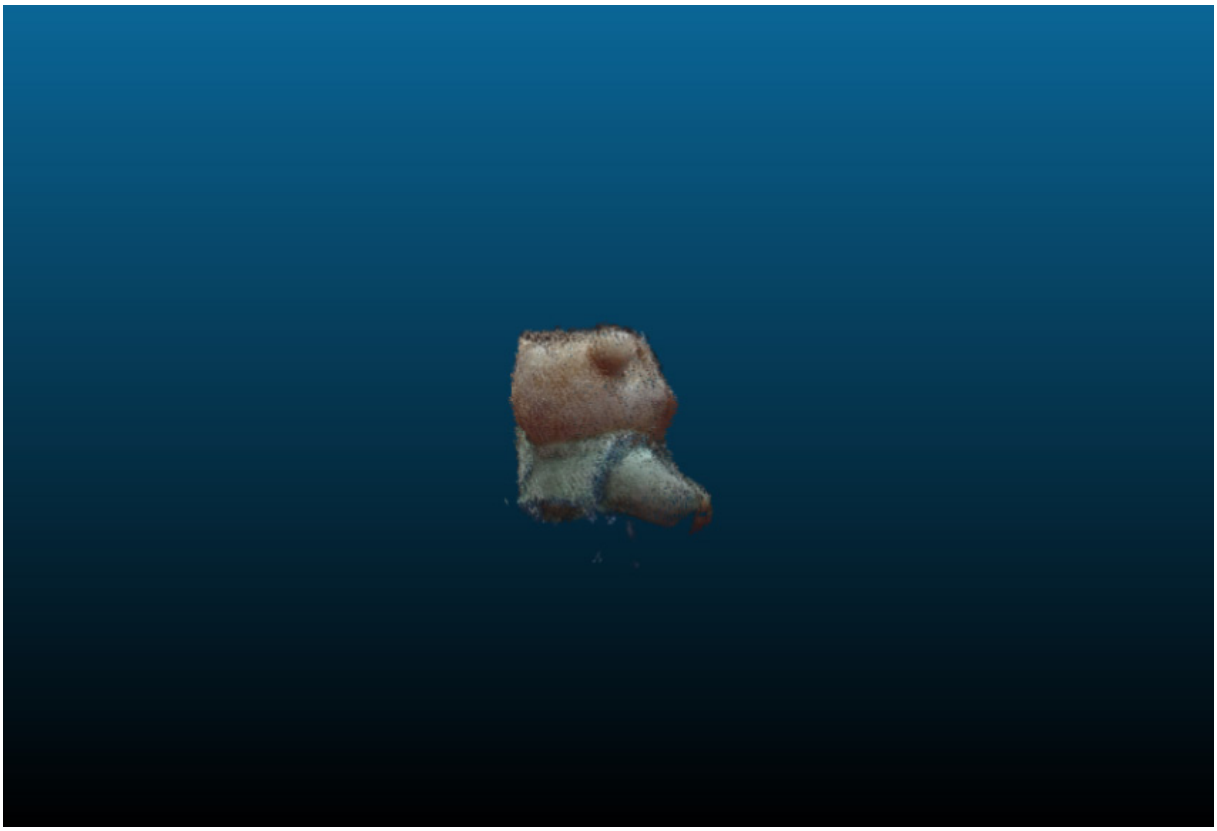


Figure 3. Reconstructed 3D model from different angles - example 2

5. Conclusion

At this paper we discuss point clouds registration problem and registration algorithms. We show how point clouds can be generated from stereo images using stereovision. And we present our realization, we construct object's 3D model from the set of stereo images. For each stereo image pair we generate point cloud and then merge these point clouds with ICP algorithm.

Bibliography

[Bagga, 2013] Prashant Jagdishchandra Bagga, “Real time depth computation using stereo imaging”, Journal Electrical and Electronic Engineering, vol. 1, pp. 51-54, 2013.

[Bellekens, 2015] Ben Bellekens, Vincent Spruyt, Rafael Berkvens, and Maarten Weyn, "A Survey of Rigid 3D Pointcloud Registration Algorithms", International Journal on Advances in Intelligent Systems, vol. 8, 2015.

[Besl, 1992] P. J. Besl and H. D. Mckay, "A method for registration of 3-Dshapes", IEEE Trans. Pattern Anal. Mach. Intell., pp. 239–256, 1992.

[Bradski, 2008] Gary Bradski, Adrian Kaehler, Learning OpenCV: Computer Vision with the OpenCV Library, 1st ed., O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, 2008.

[Hirschmuller, 2005] H. Hirschmuller, "Accurate and Efficient Stereo Processing by Semi Global Matching and Mutual Information." IEEE Computer Vision and Pattern Recognition, vol. 2, pp. 807-814 2005.

[Johnson, 1997] A. Johnson and S. Kang, "Registration and integration of textured 3D data", In Proc. 3DIM '97, Ottawa, pp. 234–241, 1997.

[Mahammed, 2013] M. A. Mahammed, A. I. Melhum and F. A. Kochery, "Object distance measurement by stereo vision", International Journal of Science and Applied Information Technology, vol. 2, pp. 05-08, March 2013.

[Minoru] Minoru 3D stereo camera: http://www.minoru3d.com/

[OpenCV] OpenCV official site: http://opencv.org/

[PCL] Point cloud library: http://pointclouds.org/

[Rusinkiewicz, 2001] S. Rusinkiewicz and M. Levoy, "Efficient variants of the ICP algorithm", 3-D Digital Imaging and Modeling, pp. 145–152, 2001.

[Sharstein, 2002] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms", International Journal of Computer Vision, vol. 47, pp. 7-42, 2002.

Authors' Information

Aram Gevorgyan *– Institute for informatics and automation problems of NAS RA, 1, P. Sevak street, Yerevan 0014, Armenia; Russian-Armenian University, Hovsep Emin steet, 123, Yerevan, Armenia; e-mail: aramgv@gmail.com*

Major Fields of Scientific Research*: Computer Vision, Stereo Vision, 3D reconstruction*

TABLE OF CONTENTS

Gender Differences in the Use of Noun Concept Categories – A Statistical Study Based on Data from Child Language Acquisition

Velina Slavova, Dimitar Atanasov, Filip Andonov 103

Cognitive Agent Based Simulation Platform for Modeling Large-Scale Multi-Level Social Interactions with Experimental Games

Maurice Grinberg, Emiliyan Todorov 117

Processing Differences between Near and Far Analogies

Alexandra Alexieva, Penka Hristova 133

Empathy and Moral Judgment in Trolley-Like Dilemmas

Veselina Kadreva, Evgeniya Hristova 148

The Factors of Effectiveness and the Organization of the Search of Elements within a Graphic Interface

Irina Blinnikova, Anna Izmalkova, Maria Semenova 162

Intelligent Techniques for Prediction of Cytosine-phosphate-Guanine (CpG) Sites associated with lung carcinoma

Abdel-Badeeh Salem, Mohamed Gawish, Mohammed Maher, Yasmin Amr, Amany Hussein, Keryakous Zarif, Basant Mohamed, Hebatullah Mohamed 176

Point clouds registration and generation from stereo images

Aram Gevorgyan Vladimir 193

Table of contents 200