



AFRL-AFOSR-UK-TR-2020-0005

Anticipating Future by Analogy-Making

**Georgi Petkov
NEW BULGARIAN UNIVERSITY
21 MONTEVIDEO STR.
SOFIA, 1618
BG**

**04/30/2020
Final Report**

DISTRIBUTION A: Distribution approved for public release.
--

Air Force Research Laboratory
Air Force Office of Scientific Research
European Office of Aerospace Research and Development
Unit 4515 Box 14, APO AE 09421

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Executive Services, Directorate (0704-0188). Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.</p>					
1. REPORT DATE (DD-MM-YYYY) 30-04-2020		2. REPORT TYPE Final		3. DATES COVERED (From - To) 30 Sep 2015 to 30 Sep 2019	
4. TITLE AND SUBTITLE Anticipating Future by Analogy-Making				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER FA9550-15-1-0510	
				5c. PROGRAM ELEMENT NUMBER 61102F	
6. AUTHOR(S) Georgi Petkov				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) NEW BULGARIAN UNIVERSITY 21 MONTEVIDEO STR. SOFIA, 1618 BG				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) EOARD Unit 4515 APO AE 09421-4515				10. SPONSOR/MONITOR'S ACRONYM(S) AFRL/AFOSR IOE	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S) AFRL-AFOSR-UK-TR-2020-0005	
12. DISTRIBUTION/AVAILABILITY STATEMENT A DISTRIBUTION UNLIMITED: PB Public Release					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT This research extended the DUAL cognitive architecture with a new model of categorization for both relation-based and feature-based concepts. New tools for human-machine interaction were integrated into DUAL. In addition, a purely connectionist model for relational thinking was constructed. Finally, psychological experiments for testing the cognitive validity of the models were conducted. These developments were captured in 13 peer reviewed journal articles published about this research.					
15. SUBJECT TERMS Analogical Reasoning, Distributed Artificial Intelligence, C2 Planning, EOARD					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT SAR	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON FRIEND, MARK
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER (Include area code) 011-44-1895-616292

FA9550-15-1-0510

ANTICIPATING FUTURE BY ANALOGY MAKING (AFAM)

09/30/2015 - 09/29/2018

**PI: Georgi Petkov, Ivan Vankov,
New Bulgarian University**

FINAL REPORT

Table of Contents

ABSTRACT	4
SUMMARY	5
INTRODUCTION.....	7
Methods, Assumptions and Procedures	8
1. The RoleMap Model of Categorization	8
2. A (Deep) Neural Network Model of Relational Reasoning	8
Results and Discussion	10
1. The RoleMap Model of Categorization	10
Simulating formation of various categories and recognition	10
Categorization on different levels of abstraction	11
Summary.....	12
2. A (Deep) Neural Network Model of Relational Reasoning	12
Learning connectionist relational representations	16
Reasoning with connectionist relational representations.....	17
Using VARS to make out-of-the box neural networks achieve combinatorial generalization	17
3. Online publishing the Python code for the DUAL architecture, and the RoleMap model	19
4. Experimental work	19
5. Publications	20
Conclusions.....	22
List of Symbols, Abbreviations and Acronyms	23

LIST OF FIGURES:

Figure 1. A layout of a convolutional deep learning network for object recognition. The network is only able to represent a single token of a given type.

Figure 2. Representing multiple tokens of the same type. There are number of representational slots ($r_1..r_5$) and each of them can encode a token of any type. The choice of a slot to represent a particular token is arbitrary. Therefore the two representations depicted are functionally equivalent.

Figure 3. Connectionist representation of the statement “The brown dog chases the white cat”. There are six representational slots (r_1, \dots, r_6) and two argument mapping spaces (A_1 and A_2). The variable binding (i.e. binding the first argument of “chase-1” token to “dog-1” and second two to “cat-1”, as well as the argument of “white-1” to “cat-1” and the argument of “brown-1” to “dog-1”) is implemented by activating the corresponding unit in the argument mapping matrices.

Figure 4. Examples if the stimuli used in the simulation with a convolutional network trained to encode the relations between two objects. The task was to encode the two spatial relations above(X, Y) and left-of(X, Y) as well as higher-order relation indicating which of the two spatial relations had greater magnituded. One of the possible relational representations of the third stimulus is shown at the bottom.

Figure 5. Description of a sample trial in Simulation 1. During encoding, the model was presented with a series of three role-filler pairs (subject - dog, verb - ate, patient - steak). At the last (fourth) time step, the model was only presented with a test role (in the example above: subject) and it had to output the corresponding filler which was associated to it (dog). When the model was trained to produce VARS representations, the role-filler pairs were accompanied by a random permutation of tokens during encoding which determined the allocation of symbols to representational slots (for example, the fact that dog was paired to token 3 meant that dog has to be represented in the third representational slot). The verb was treated as a binary relation and its arguments (i.e. the relational roles subject and patient) were bound to the corresponding fillers (in this example, subject to dog and patient to steak). Note, the VARS representations were only trained in time step 4.

ABSTRACT

The DUAL cognitive architecture is based on the assumption that analogy-making underlies most of the human cognition. Analogy-making is a fundamental cognitive ability that let us perceive the world and behave both predictively and effectively by taking into account our past experience even in an environment with dynamically changing context.

During the AFAM project realization, the DUAL cognitive architecture was further developed. A new model of categorization of both relation-based and feature-based concepts was designed. A purely connectionist model of relational thinking was constructed. New tools for human-machine interaction were integrated. Psychological experiments for testing the cognitive validity of the models were conducted.

KEYWORDS: analogy-making, cognitive modeling, categorization, relational thinking

SUMMARY

Two main models are realized – a symbolic model of categorization and a neural network implementation of relational reasoning. In addition, some other simulations that cover the objectives of the project are conducted. Some psychological experiments are performed.

1. Constructive model for categorization based on the DUAL cognitive architecture and on the assumptions that the sub-processes of analogy-making are at the core of human cognition. The simulations with the model highlight its abilities to automatically create various types of categories and use them for recognition and thinking:
 - a. Role-governed categories (ex. 'domestic animals');
 - b. Goal-driven categories (ex. 'dairy animals'), as well various ad-hoc concepts;
 - c. Theme-based categories ('the animals at the village');
 - d. It is demonstrated that the typical feature-based categories (ex. 'cat') could be created and recognized on the same basic principles;
 - e. It is demonstrated that time separation into situations, events, etc. could be modeled on the same basic principles;
 - f. Categorization on different levels of abstraction (ex. 'cat' – 'mammal' – 'animal')
 - g. Context dependent categorization (ex. 'predator')
 - h. Probabilistic categories
2. A neural network account of representing and learning relations. The network is designed as a part of the large goal to model the main sub-processes of analogy-making by more psychologically plausible mechanisms and to deal with a large-scale data. The model successfully:
 - a. Represents structure of varying complexity
 - b. Learns to represent spatial relations (ex. 'right-of' ('circle', 'triangle'));
 - c. Works non-symbolically, with real pictures of various scenes;
 - d. Captures implicitly magnitudes;
 - e. Learns relations of second order, i.e. relations between relations (ex. 'the magnitude of the relation "right" is more than the magnitude of the relation "above."
3. Tools that allow the system to ask the user for information are designed, e.g. for giving names of the automatically created categories. Tools for testing the model with a large database, as well as parser for transforming simple sentences onto DUAL-agents, are implemented. Both models are posted online.
4. Experimental work on:

- a. Testing the hypothesis that analogy-making is important part of understanding the intentions of the others in social situations.
 - b. Testing the hypothesis that the famous Garner interference effects are product of relational mapping;
 - c. Testing a prediction of the RoleMap model that the known inverse-base effect may emerge from basic mechanisms of retrieval and inhibition.
5. Several publications have been produced to disseminate some of the results.

Thus, the objectives of the project are fulfilled.

INTRODUCTION

It is a long-standing dispute between the so-called direct and inferential approaches to many cognitive processes. This debate may be found in the field of vision (ecological vs. constructivist approach), social science (direct vs. inferential understanding of other's intentions and goals), memory and many others cognitive processes.

Whereas the constructivist approach is supported by many psychological data, there are few computational (or other type of formal) models based on it. According to the constructivist approach, people continuously generate inferences and hypotheses about the past, the present, and the future, and continuously try to verify or reject them. However, there are various possible ways to argue where these hypotheses come from.

The answer that we try to explore is: by the sub-mechanisms for analogy-making. Analogy-making is a fundamental cognitive ability that let us perceive the world and behave effectively by taking into account our past experience. Recent theoretical and empirical studies have suggested that analogy-making underlies cognitive processes such as context specific memory retrieval, categorization, higher-level perception, learning and generalization; planning; understanding of complex social situations.

The DUAL cognitive architecture (Kokinov, 1994, 1998) implements a context-sensitive, emergent and intrinsically parallel view of analogy-making. The major components of DUAL are the processes of associative memory retrieval, analogical mapping, and the more recently developed anticipation formation and analogical transfer. All these processes operate in parallel and permanently interact with each other.

However, new challenges and potential advantages and implications of the model emerged. Thus, during the last few years, the transfer mechanism was developed further on and a lot of simulations tested successfully the abilities of the architecture to account to cognitive processes like dynamic re-representations; abstraction, concept acquisition, generalization of problem solutions.

A model of categorization of both relational and future-based concepts was developed. A purely connectionist model of relational thinking was designed. Tools for user-machine interaction were integrated. Some psychological experiments were obtained for providing empirical data and testing the cognitive validity of the models.

In general, nevertheless that this project was initially designed to deal with fundamental research, it could be potentially used for providing relevant mechanisms for developing intelligent application-oriented intelligence systems.

Methods, Assumptions and Procedures

1. The RoleMap Model of Categorization

The model is built under the assumptions that few basic mechanisms (usually thought as sub-mechanisms of the analogy-making process) underlie most of the human cognition. These are (1) Associative spreading of activation, which allows the pattern of activation to represent dynamically the relevance of each item to the current context. (2) Creation of local mappings, which represent that the system has found something in common between two items. (3) Self-organize of these local mappings into larger supporting or competing each other coalitions on the basis of the relational structures. (4) Transfer, modeled as a formation of anticipations that fill the gaps in the target representation on the basis of the mappings performed.

The mapping themselves represent the knowledge that there is something in common between two items. Thus, they are the embryos of new concepts. Unlike the most of the other models of learning of new categories (in general, of learning of anything), the current model doesn't assume different stages of learning and using the learned. Instead, the mapping creation is thought as a basic process that occurs automatically. However, some of the mappings could exceed a certain threshold for staying active enough for enough time. They are transformed into concepts for additional usage. Thus, the category formation is modeled as a side effect of the natural cognitive processes, not as a separate process. However, these new categories will affect in turn the subsequent thinking.

In an analogical way, some of the transferred anticipations represent the knowledge that some of the items should be grouped together. In the same way, some of these anticipations are transformed into instances. This is the process of categorization into already existing categories: the system groups some items together because they are analogical to memorized items that are already grouped into categorized instance. Again, the process of categorization is not modeled as a separate process. Instead, it emerges as a side effect of the assumed natural sub-processes of mapping and transfer.

2. A (Deep) Neural Network Model of Relational Reasoning

Recent advances in neural network modeling have lead to spectacular results in fields as diverse as object, face and scene recognition, speech perception and machine translation. One of the most appealing aspects of this success story is that the same set of tools such as the back propagation learning algorithm or the convolutional network architecture have been applied to very different tasks with minor or no modifications. Therefore, neural network modelling are productive not only in advancing artificial intelligence but also in providing a uniform modelling framework for understanding cognition. However there is still no agreement on whether there are cognitive processes which can't be accounted for in connectionist terms.

One of the major arguments of the opponents of connectionism is that it lacks the ability to represent symbolic structures and is thus unsuitable for modelling tasks requiring such representations (Fodor & Pylyshyn, 1988). Other authors (Hummel & Holyoak, 2003; Kokinov & Petrov, 2001; van der Velde & de Kamps, 2006) have argued that connectionist systems need to be augmented with additional representational mechanisms in order to be able to account for symbolic reasoning. For example, the AMBR model of relation reasoning (Kokinov & Petrov, 2001), built on the foundations of the DUAL hybrid architecture, is based on the assumption that symbolic reasoning requires integrating connectionist and symbolic mechanisms.

There are several advantages of developing a purely connectionist account of relational reasoning. First, a purely connectionist model will be based on a widely used and popular set of modelling tools and this will be able to benefit from the existing expertise in the field. Second, neural network models benefit from recent progress in developing hardware which allows to train and run connectionist models in parallel, which drastically improve their performance. Most importantly, developing a neural network model of relational reasoning will help to address one of the major deficiencies of the symbolic and hybrid connectionist-symbolic accounts - the problem of how symbolic structures are learned and extracted from experience.

Results and Discussion

1. The RoleMap Model of Categorization

Simulating formation of various categories and recognition

In one of the simulations, the knowledge that “grandmother feeds chickens” is encoded in the model with a coalition of three agents. The knowledge about some concepts – “chicken”, “bird”, “animal”, “feeds”, etc. is also encoded as a semantic knowledge of the model. When a new situation appears on the target, namely that “grandmother feeds a pig”, then the processes of spreading activation and mapping formation start to work. Relatively easily the model found out that both the memorized agent for ‘chicken’ and the target one for ‘pig’ are analogous for two reasons – both they are animals and both they are the second argument of the relation “feeds”. The mapping between them emerged. Without any inhibitory pressures, it survived active enough for enough time and a procedure for creating new concepts started its work. It is a tool modeled in DUAL that allows the system to ask the user for new names. Namely, the model highlighted the mapping between the chicken and the pig and asked: “what is this?”. The eventual answer of the user “domestic animal” served as a name for the newly created concept (this functionality, however, could be disabled if the system is run on a mode for statistical collection of data from many runs). The simulation, however, doesn’t stop here. Following the same mechanisms, a new concept of “grandmother feeds animals”, a sub-concept of ‘grandmother’ was created; as well a concept of ‘feeds animals’, a sub-concept of ‘feeds’. In addition, a whole situation “feeding domestic animals” emerged as result of the mapping between the situation agents that bind the memorized and the target structures. Note, the model doesn’t distinguish the representations of the spatial and time related items, i.e. the same principles are used for encoding objects, their parts, and situations and their sub-situations.

The newly created concepts stayed in the knowledge base and a new simulation could be run without any reloads. If a new target appears: ‘grandmother feeds a cow’, the mappings between the cow and the retrieved memorized instances of ‘pig’ and ‘chicken’ emerge. However, they are already encoded as parts of situations, that are both instances of the situational concept “feeding domestic animals”. Thus, the anticipatory mechanism starts its work. If the memorized elements of the mappings are organized in a group (i.e. the base element of a mapping is part of something, that has not analog in the target), then the system creates an anticipation, which tries to fill the gap. In the example, the system creates an anticipation situation agent that groups the respective target elements. Without any inhibitory pressures, this anticipation survives and is transformed as a new instance. With other words, the system *recognizes* the new situation as “feeding animals” and in turn, recognizes its respective elements.

Of course, by the same mechanisms, it will be a pressures to create a new concept that combines the target ‘cow’ and the base ‘chicken’, as well the ‘cow’ and the ‘pig’. It is now, however, an important difference between the previously described simulation (when only a ‘chicken’ was in the long-term memory and a ‘pig’ on the target) and the current one. In the first simulation, new

concepts were created, whereas in the second one finished with a recognition of already existing ones. The reason is that in the second simulation the mappings 'cow'—'chicken' and 'cow'—'pig' inhibit each other. Thus, without any other pressure they will not survive as new concepts. In a certain context, however, if an additional contextual pressure is presented, this may happen. This illustrated the work of the last main mechanism of the model – the constraint satisfaction mechanism. It combines the different pressures and depending on the context determines the final behavior of the model.

In more complicated simulations, one after the other, more pressures were added; different types of situations were encoded and presented to the model, thus the ability of the model both to create new concepts and to recognize in old ones was demonstrated. The model is able both to learn new categories, and to recognize new instances into already existing ones. This happened with various types of relational categories – goal-driven ones, theme-based ones; context dependent categories.

In addition, following the same basic mechanisms, the model successfully created and recognized categories typically assumed as featured-based ones. for example, a 'cat' was represented as a coalition of about ten agents, representing parts like 'head', 'tail', etc., as well the relations between them. The relational interconnection between the parts is important: part of the anticipatory mechanism combines different locally and asynchronously created anticipations, if they belong to relationally interconnected items. with other words, if for example, there is not a path through relations and their arguments between the 'head' and the 'tail', the system will anticipate that there are *two* cats presented! The 'head' belongs to the first one; the tail – to the second one. From one side, we thing this as an advantage of the model – many other models for categorization are unable to recognize two items from one and the same category presented simultaneously. From the other side, this proposal challenges the theoretical models that assume that purely feature-based categories, without any relational information between these features, exist. In turn, this peculiarity of the model makes it more falsifiable and produces predictions that could be empirically tested.

Finally, the model doesn't distinguish principally the representations of objects and situations. Thus, various time-prolonged events were also successfully categorized and additionally recognized.

Categorization on different levels of abstraction

The basic functionality of the DUAL based model of categorization can be described as following: DUAL can do abstractions by analogy. If a certain input information is given, DUAL can find some structurally similar episodes from the memory, as well some abstract concepts, and on the basis of a certain parameters values, can decide whether to categorize the input in a certain category (if the best structural fitting is with a memorized concept), or to create new category (if the best fitting is with a concrete memorized exemplar), or just to leave the input situation as a non-categorized unit. More precisely, the parameters determine the work of the constraint satisfaction mechanism. However, the first simulations with this model couldn't capture the possibilities to categorize on different levels on abstraction. More concretely, DUAL can categorize the concrete structural description of a certain dog as "dog". It can categorize it also

as “mammal”, or as an “animal”. However, DUAL couldn’t do this simultaneously, with one and the same values of the parameters. This is a common problem for other models of categorization, based on structural similarities, for example, for SEQL (Kuehne et al., 2000). The solution we proposed is to use simultaneously several (two in the performed simulations) sets of the parameters values. If a certain mappings satisfy the more conservative set, then these mappings survive and are transformed into concepts as described above. However, if they don’t satisfy them, but do for the second set of parameters, according to the newly designed mechanisms, they survive too and are transformed into concepts too but on a higher level of abstraction.

Summary

The existing DUAL architecture (Kokinov, 1994, Petrov, Kokinov, 1998) is an excellent instrument to model various cognitive processes on the basis of the paradigm that human cognition is constructive in its nature and that manipulation of relations is at the basis of the top-down creation of hypotheses. Thus, within this project, we developed a number of new reasoning mechanisms and integrated them in the AMBR model of analogy-making. As a result, one of the big problems of the DUAL architecture – that most of its knowledge representations are manually encoded – was partially solved. The new mechanisms allow the system to automatically create new relational categories (as well, many feature-based ones) and to categorize the consequently coming items into them. The model has properties, which are not presented in many of the other model of categorization: it allows for “fast learning”, just by a single exemplar; it allows automatically categorization on different levels of abstraction; it deals with relational categories, but as well with feature-based ones; it is able to categorize also situations and events.

2. A (Deep) Neural Network Model of Relational Reasoning

VARs - the vector approach to representing symbols

There are two major problems which need to be solved in order to represent relational information in connectionist fashion (i.e. by using vectors of real numbers only). The first problem is the type/token distinction, i.e. the ability to have more than one representations of the same thing, for example two instances (tokens) of a cat (type). This problem is also known as the problem of two (van der Velde & de Kamps, 2006). Typical neural networks, such as the widely popular convolutional models of object recognition, lack this ability - they can only represent one occurrence of a given category in the perceptual input (Figure 2.1). This won’t work if the task is to encode the relation between two object which belong to the same category (for example, same-shape(circle-1, circle-2)). One solution is to use recurrent neural networks which unfold relational representations in time (for example, in the way LISA, Hummel & Holyoak, 2003, does). However this approach poses a number of other problems which are hard to be addressed (for example, how such temporal representations are orchestrated). Therefore we propose another approach to solving the type/token distinction problem (Figure 2.2). In our view, knowledge (including relational) is represented redundantly in a limited number of “representational slots”. Each representational slot is capable of representing all the knowledge of the system and the choice of slot to represent a given fact is arbitrary.

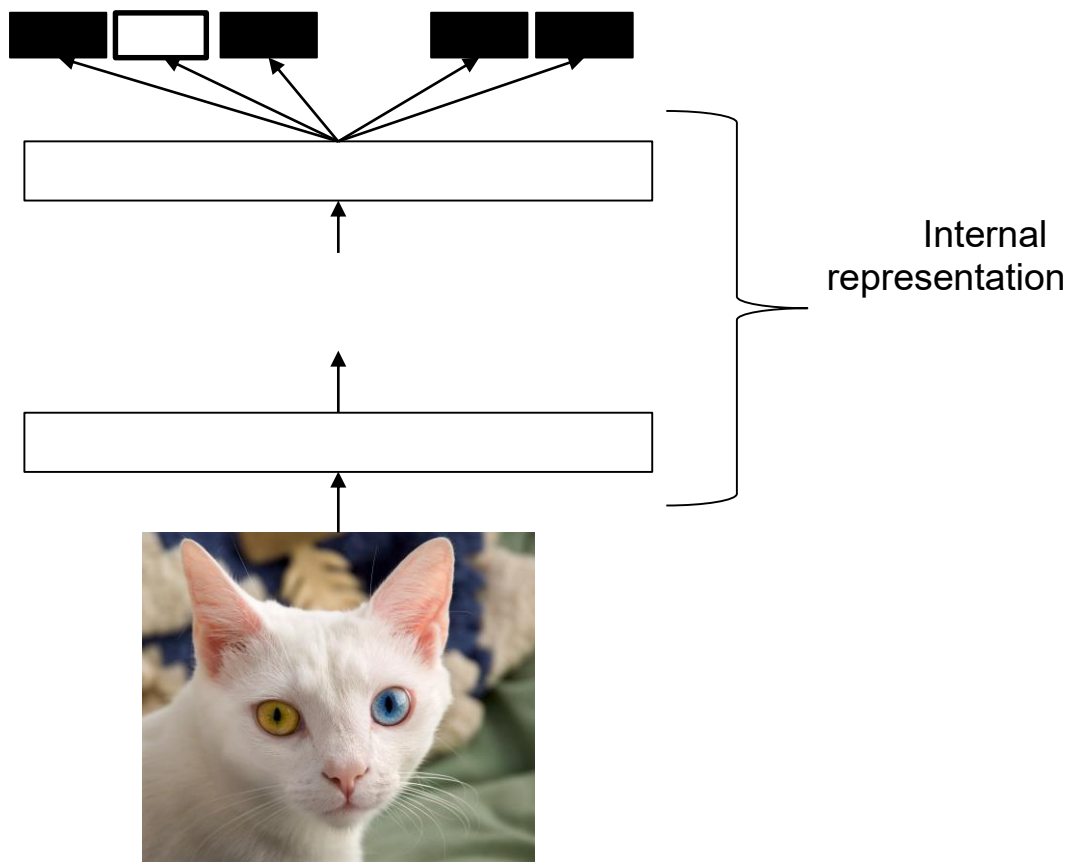


Figure 1. A layout of a convolutional deep learning network for object recognition. The network is only able to represent a single token of a given type.

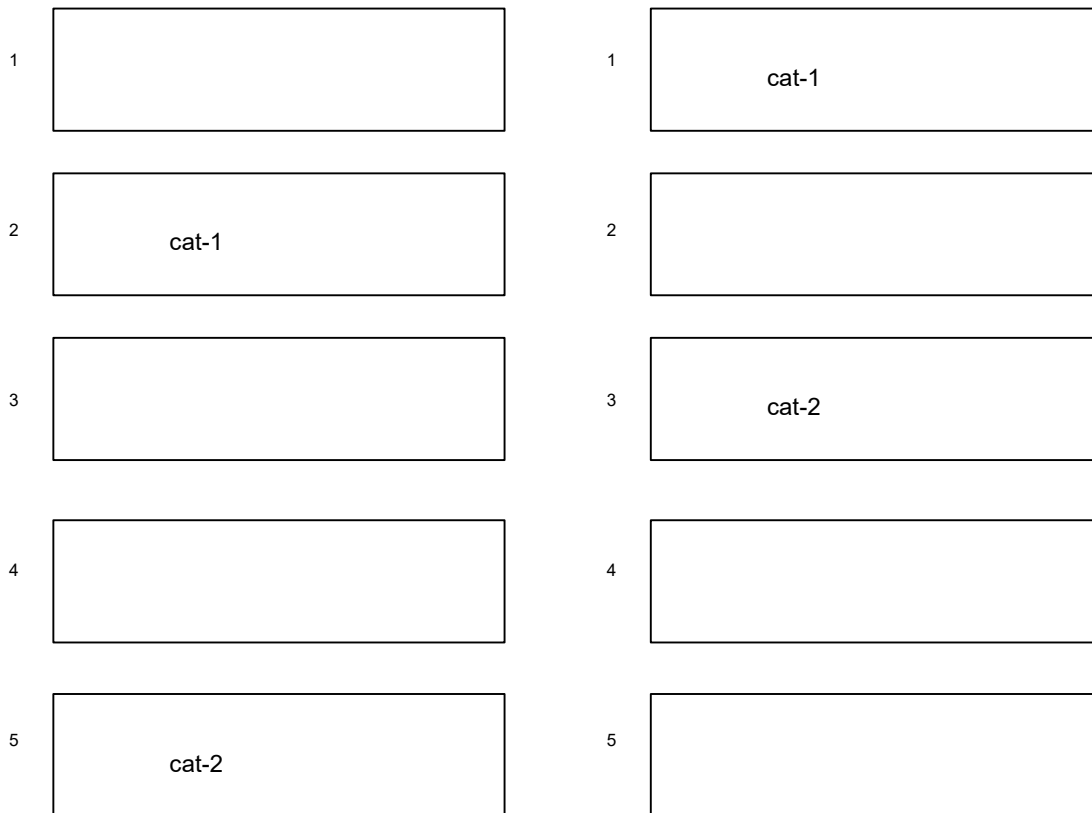


Figure 2. Representing multiple tokens of the same type. There are number of representational slots ($r_1..r_5$) and each of them can encode a token of any type. The choice of a slot to represent a particular token is arbitrary. Therefore the two representations depicted are functionally equivalent.

The second major challenge to representing relations in pure connectionist systems is the variable binding problem (also known as the role-filler binding problem). The capacity to bind the argument variables (i.e. roles) of a relational concept to particular instances of objects and other relations is fundamental to relational reasoning. Consider the spatial relation “left-of(X , Y)”. The meaning of the relation poses little constraint on the nature of its possible arguments. In fact, the relation can hold between any two material objects, not to mention its metaphoric use. There are infinitely many possible bindings involving X and Y and it is unfeasible to represent them using conjunctive coding which relies on pre-existing structures. What is needed is a way to bind a variable to any argument, even if they have never been bound so far. Variable binding is essential for implementing systematicity - a fundamental characteristics of symbolic reasoning (Fodor & Pylyshyn, 1988), which posits that higher order representations should be independent of the their lower level constituents. Systematicity entails that if we are able to understand the proposition “The hospital is to the left of the church.”, we should be equally able to understand “The church is to the left of the hospital” or even “The dax is to the left of the fisp”, although we have no idea

what a dax or a fisp is. In more general terms, variable binding is needed for representing any kind of combinatorial knowledge.

Our solution to the variable binding problem is in having a number of argument mapping spaces (Figure 2.3). The number of such spaces is the maximum arity of the predicates which can be represented (for example, if the system can only represent unary and binary relations, then there will be two argument mapping spaces) and the n -th argument mapping space represents the binding of the n -th arguments of the representation relations to the corresponding lower-order predicates.

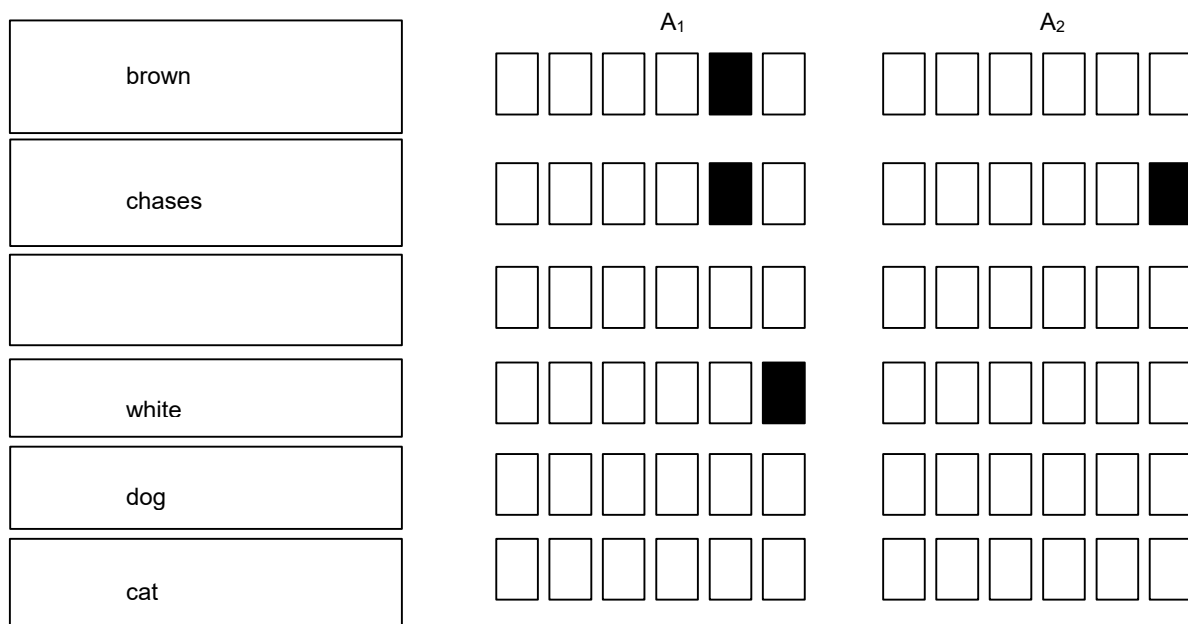
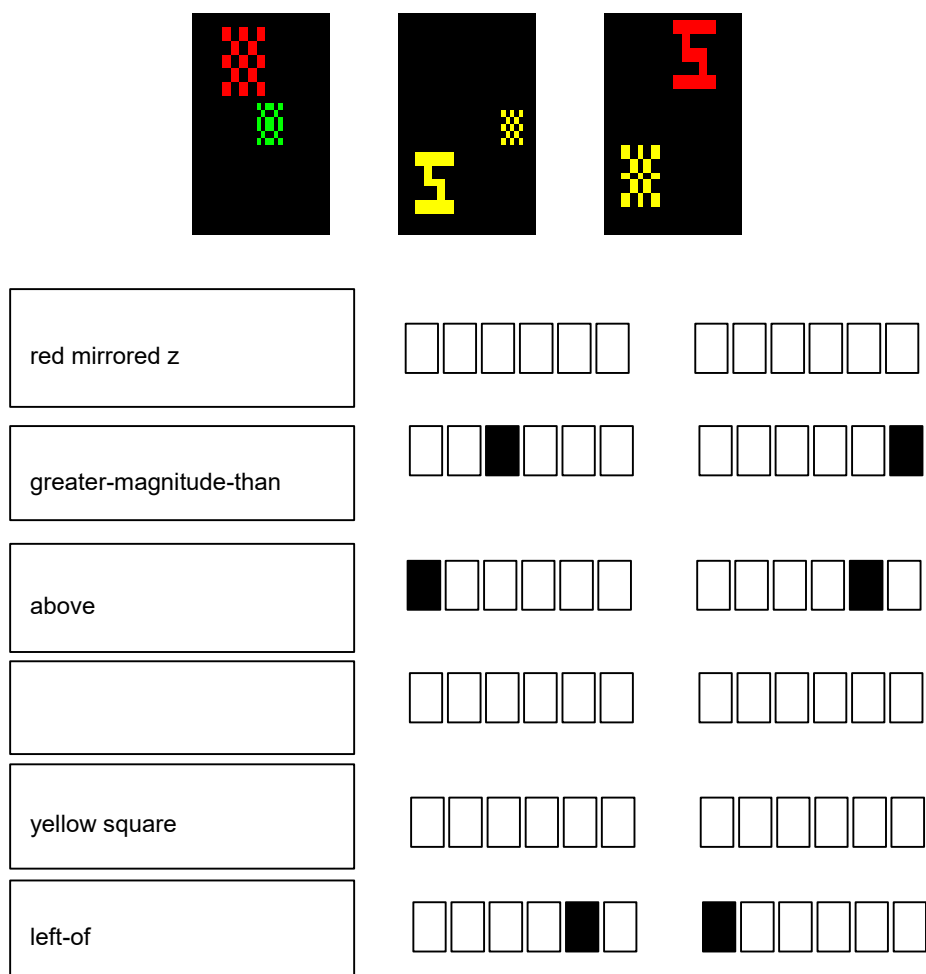


Figure 3. Connectionist representation of the statement “The brown dog chases the white cat”. There are six representational slots (r_1, \dots, r_6) and two argument mapping spaces (A_1 and A_2). The variable binding (i.e. binding the first argument of “chase-1” token to “dog-1” and second two to “cat-1”, as well as the argument of “white-1” to “cat-1” and the argument of “brown-1” to “dog-1”) is implemented by activating the corresponding unit in the argument mapping matrices.

The advantage of our approach, which we refer to as VARS (vector approach to representing symbols) is that it allows to represent symbolic structures of varying complexity in conventional connectionist systems, without the need to introduce additional representational mechanisms. The proposed representational scheme meets the requirements put forward by Fodor & Pylyshyn (1988) as long as it is both productive and systematic. In this view, the ability to represent symbols is only constrained by the number of representational slots and the maximal arity of the represented predicates. Both these constraints are neurally and cognitively plausible (one may argue they reflect the limits of working memory capacity) and computationally feasible.

Learning connectionist relational representations

One of the motivations to develop a purely connectionist account of relational reasoning is that it will allow to learn relational representations from experience, using the powerful tools of (deep) learning in neural networks. To demonstrate this, we developed a convolutional deep neural network which was trained to discover two spatial relations (above and left-of) and a higher-order relation (greater-magnitude-than) between two objects with varying shape, color and size (Figure 2.4). The network not only managed to fit the training set, but also to generalize across pairs of objects which did not appear in the training set. The results of the simulations show that the proposed representations of relational information are learnable within state-of-the-art



connectionist architectures. We implemented the neural network using the Tensorflow modelling framework, which allowed to run the simulations on parallel hardware (we acquired a NVidia GTX1080 GPU within the project) and thus drastically increase performance.

Figure 4. Examples of the stimuli used in the simulation with a convolutional network trained to encode the relations between two objects. The task was to encode the two spatial relations

above(X, Y) and left-of(X, Y) as well as higher-order relation indicating which of the two spatial relations had greater magnitude. One of the possible relational representations of the third stimulus is shown at the bottom.

Reasoning with connectionist relational representations

One of the advantages of the proposed connectionist representations of structure is that they support analogical comparison in a straightforward and elegant way. In order to determine how analogous two representations are and which the corresponding elements are you need just to rearrange the elements of one of the representations across the representational slots until the vector similarity (measure by cosine similarity, for example) of the two representations is maximized:

$$\text{Analogy} = \text{Max}[(1-\sigma) * \cos(\text{semantics}_1, \text{semantics}_2) + \sigma * \cos(\text{structure}_1, \text{structure}_2)]$$

, where *semantics* is the concatenation of the representation slots of the corresponding representation, *structure* is the concatenation of the flattened argument mapping matrices and

σ is the relative weight of superficial (semantic) and structural similarity.

It is possible to apply the same algorithm to compare a target representation to multiple other representations (bases) in order to implement simultaneously analogical retrieval and mapping. Both analogical mapping and retrieval have been implemented in Tensorflow and can be run on parallel hardware. This is particularly beneficial for analogical mapping when the number of bases is large - simulations showed speed ups up to a factor of 10 when using parallel computations.

We have also run a number of simulations which showed the proposed connectionist model of analogical mapping meets the systematicity requirements, as defined by Gentner (1983) - if two higher order predicates are mapped, then their arguments should be mapped as well. Another set of simulations investigated how well the model copes with situations involving cross-mapping (when the semantic and the structural similarity are at odds). In all of the simulations the model performed similarly to classical (and far more complicated) models of analogy-making such as SME, LISA and AMBR.

Using VARS to make out-of-the box neural networks achieve combinatorial generalization

In this project we demonstrated how VARS can be used to make conventional neural network architectures solve a problem - combinatorial generalization - which they are supposed to fail on. The idea is to train an ordinary neural network to explicitly represent the symbolic structure of its input while solving its main task. We run two such simulations showing how two conventional neural architectures can achieve impressive success in tasks which are usually

considered very difficult for out-of-the-box architectures without specialized mechanisms for symbolic processing.

In the first simulation, we replicated the simulation by Kriete, Noelle, Cohenm & O'Reilly (2013) which aimed to show that combinatorial generalization is impossible in simple recurrent networks. Contrary to their finding, we show that a standard recurrent architecture based on long short term memory cells (LSTM) can achieve perfect combinatorial generalization if VARS is used to press it to explicitly represent the symbolic structure of its input (Figure 2.5). In another simulation, we obtain a similar result in a visual reasoning task and convolutional neural network architecture. Our findings pose a serious challenge to the widespread assumption that conventional neural network architectures are incapable of symbolic processing unless and they are equipped with mechanisms which are engineered specifically for the purpose of representing symbols. Interestingly, our results are in tact with recent results by the DeepMind team (Hill, Santoro, Barrett, Morcos, & Lillicrap, 2019; Barrett, Hill, Santoro, Morcos, & Lillicrap, 2018) and by Schlag & Schmidhuber (2018), showing that implementing pressure to encode relational information may improve performance of deep neural networks in relational reasoning tasks. We plan to follow this line of research and investigate the extent to which one can push the ability of neural networks to solve problems requiring symbolic reasoning.

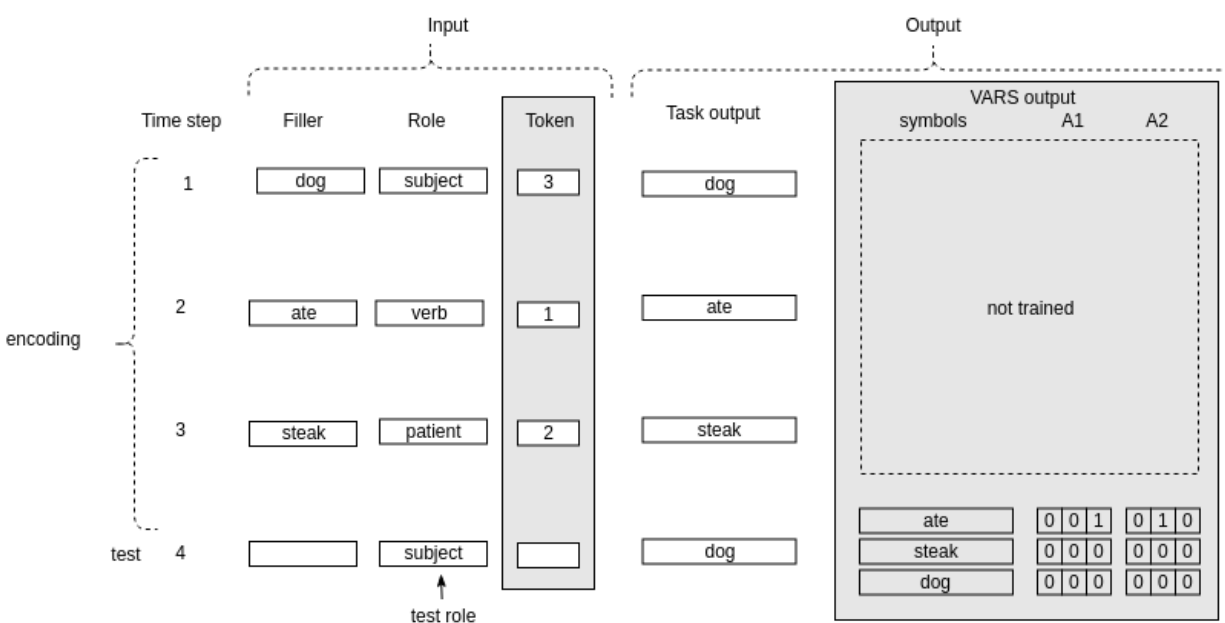


Figure 5. Description of a sample trial in Simulation 1. During encoding, the model was presented with a series of three role-filler pairs (subject - dog, verb - ate, patient - steak). At the last (fourth) time step, the model was only presented with a test role (in the example above: subject) and it had to output the corresponding filler which was associated to it (dog). When the model was trained to produce VARS representations, the role-filler pairs were accompanied by a random permutation of tokens during encoding which determined the allocation of symbols to

representational slots (for example, the fact that dog was paired to token 3 meant that dog has to be represented in the third representational slot). The verb was treated as a binary relation and its arguments (i.e. the relational roles subject and patient) were bound to the corresponding fillers (in this example, subject to dog and patient to steak). Note, the VARS representations were only trained in time step 4.

3. Online publishing the Python code for the DUAL architecture, and the RoleMap model

The whole work on the code was improved, documented, and put on the Web. It could be found on <https://github.com/nbuDUAL/DUAL>. In addition, a work on integrating the WordNet ontology (<https://wordnet.princeton.edu/>) into the DUAL architecture is performed. This will allow the DUAL based models to work in a real-world environment, with a large semantic memory. More specifically, the concepts from the WordNet are transformed onto interconnected DUAL agents. In addition, parsers using the Python's NLTK package are responsible for constructing the episodic memory. Thus, the situations, which the user wants to encode in the DUAL models should be written on an almost natural language (with sentences with simple structure) and the parsers are responsible for transforming them into episodic DUAL agents.

This overcomes the limitation of the DUAL architecture for manual encoding of the structures.

4. Experimental work

Apart from modeling, we also used psychological experimentation for both understanding more about human cognition and to verify our model.

One series of experiments were performed to understand more about how analogical thinking is related with the process of understanding the intentions of the others. We provided five experiments based on similar designs. We provided people with a single story and asked them to judge what the intention behind the behavior of one of the characters of the story was. During the experimental series we manipulated the overall mood of the story, its reality, etc. In all of the experiments we provided people also with one another story that they should read before the target one. The contextual story varies according to its similarity to the target one – it was either structurally similar (shared the same relations), or superficially similar (sharing the same features), or emotionally similar (sharing the same overall mood it suggests), or not similar at all. In this way, by linking theoretical and empirical findings about analogy making and intentions' understanding, we proposed a novel hypothesis about the mechanisms underlying the process of understanding intentions in ambiguous situations. This hypothesis states that people are able to spontaneously put to use concrete episodes to infer the intentions of a target agent in a novel, but structurally similar situation. In support of this hypothesis, the results of three experiments (Experiments 1, 2 and 5) showed that the participants were more likely to attribute a negative intention to the actor in an ambiguous situation if the latter is preceded by a negative, structurally similar episode

(“analog”). However, structural similarity tends to interact with activated stereotypes as the change of the characters in the base story from negative to positive attenuated the effect of the negative analog (Experiment 3 and 4) and enhanced the effect of the positive analog (Experiment 3). Furthermore, across three experiments (Experiment 2, 4 and 5) it is demonstrated that, under certain conditions, participants are more likely to attribute a negative intention to the actor in an ambiguous situation if the latter is preceded by a positive analog. This phenomenon, which we termed inverted effect, is suggested to occur as a result of a failure to evaluate the analogy between the analog and the target story as useful due to their alienable differences. Taken together, the findings of the five studies provide moderate support for the proposal that analogy processing plays an important role in understanding intentions and suggest that analogical inference might be the looked for unitary mechanism which underlies the use of both general knowledge and concrete episodes in understanding intentions in ambiguous situations.

To sum up, we provided an empirical investigation that is a progress of the second objective of the project – modeling understanding of others in social situations.

Second series of experiments check the hypothesis that relational correspondences may play role in the known Garner interference effects. Crossmodal correspondences are innate, language-based and statistically derived. They occur across all sensory systems and in different cultures. Despite their multiformity, they are exhibited analogously, mainly through robust congruency effects. One plausible explanation is that they rely on a common underlying mechanism, reflecting our fundamental ability to transfer relational patterns across different domains. We investigated the pitch-height correspondence in a bimodal sound-discrimination task, where relative sound context was changed online. The intermediate sound frequency was presented successively with two lower or higher equidistant sound frequencies, and squares at two fixed vertical positions. Congruency effects were transferred across sound contexts with ease. The results supported the assumption about the relational basis of the crossmodal associations. In addition, vertical congruency depended critically on the horizontal spatial representations of sound. Thus, probably crossmodal correspondences are indeed represented as relations.

5. Publications

The results from the AFAM project were published in the following papers:

Shahbazyan, L., Petkov, G. & Gurova, L. (2014). Relational priming enhances hostile attribution bias (HAB) in adults. *Folia Medica*, 56(1), 19.

Popov, V., Petkov, G. (2014) The level of Processing Affects the Magnitude of Induced retrograde Amnesia. In: *Proceedings of the 36th Annual Cognitive Science Conference*, Quebec, pp. 2787-2792

Shahbazyan, L., Petkov, G., Gurova, L. (2014) Analogical Transfer of Intentions. In: *Proceedings of the 36th Annual Cognitive Science Conference*, Quebec, pp. 2907-2912

Gurova, L., Schahbazyan, L., Petkov, G. (2015). Understanding Hostile Attribution bias: Dodge's Model Revised. In: *Drozdostoy Stoyanov (Ed.), Towards a New Philosophy of Mental Health: Perspectives from Neuroscience and the Humanities*. Chapter 11. Cambridge Scholar Publishing, ISBN (10): 1-4438-7661-5, ISBN (13): 978-1-4438-7661-2.

Petkov, G., Pavlova, M. (2016) The role of similarity in constructive memory: evidence from tasks with children and adults. In: *Proceedings of the 38th Annual Cognitive Science Conference*, Philadelphia, pp. 325-330

Vankov, I. I., & Bowers, J. S. (2017). Do arbitrary input–output mappings in parallel distributed processing networks require localist coding? *Language, Cognition and Neuroscience*, 32 (3), 392-399.

Bowers, J. S., Vankov, I. I., & Ludwig, C. J. (2016). The visual system supports online translation invariance for object identification. *Psychonomic bulletin & review*, 23 (2), 432-438.

Bowers, J. S., Vankov, I. I., Damian, M. F., & Davis, C. J. (2016). Why do some neurons in cortex respond to information in a selective manner? insights from artificial neural networks. *Cognition*, 148 , 47-63.

Petrova, Y., Petkov, G. (2018) Role-Governed Categorization and Category Learning as a Result from Structural Alignment: The Rolemap Model. In: *Proceedings of the 20th International research conference, Barcelona*, pp. 1762-1769

Petkov, G., Vankov, I., Petrova, Y.(2018) Discovering the Dimension of Abstractness – Structure-Based Model That Learns New Categories and Categorizes on Different Levels of Abstraction. In: *Proceedings of the 20th International research conference, Barcelona*, pp. 1820-1825

Vankov, I. I., Bowers, J. S. (under review). Out-of-the box neural networks can support combinatorial generalization. Submitted to *Philosophical Transactions of the Royal Society B: Biological Sciences*.

Zafirova, Y., Petrova, Y., & Petkov, G. (2019) Crossmodal Spatial Mappings as a Function of Online Relational Analyses?. In: *Proceedings of the 41st Annual Conference of the Cognitive Science Society. Montreal, Quebec, Canada*

Petkov, G., Petrova, Y. (2019). Relation-based categorization and category learning as a result from structural alignment. The RoleMap model. *Frontiers in Psychology*, 10, 563. doi: 10.3389/fpsyg.2019.00563., ISSN: 1664-042

Conclusions

We have managed to address the main goals of the project. Some of our work has been dedicated to problems which have not been initially planned, as our understanding of the role relational reasoning plays in solving problems evolved. We have paved the way for many interesting and promising research lines which, in our view, have valuable theoretical and practical implications.

References:

- Barrett, D. G., Hill, F., Santoro, A., Morcos, A. S., & Lillicrap, T. (2018). Measuring abstract reasoning in neural networks. *arXiv preprint arXiv:1807.04225*.
- Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28(1-2), 3-71.
- Hill, F., Santoro, A., Barrett, D. G., Morcos, A. S., & Lillicrap, T. (2019). Learning to make analogies by contrasting abstract relational structure. *arXiv preprint arXiv:1902.00120*.
- Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological review*, 110(2), 220.
- Kokinov, B. (1994). A hybrid model of reasoning by analogy. *Advances in connectionist and neural computation theory*, 2, 247-318.
- Kokinov, B. (1998). Analogy is like cognition: Dynamic, emergent, and context-sensitive. *Advances in analogy research: Integration of theory and data from the cognitive, computational, and neural sciences*, 96-105.
- Kokinov, B., & Petrov, A. (2001). Integrating memory and reasoning in analogy-making: The AMBR model. *The analogical mind: Perspectives from cognitive science*, 59-124.
- Kriete, T., Noelle, D. C., Cohen, J. D., & O'Reilly, R. C. (2013). Indirection and symbol-like processing in the prefrontal cortex and basal ganglia. *Proceedings of the National Academy of Sciences*, 110(41), 16390-16395.
- Kuehne, S., Forbus, K., Gentner, D., & Quinn, B. (2000, August). SEQL: Category learning as progressive abstraction using structure mapping. In *Proceedings of the 22nd annual meeting of the cognitive science society* (pp. 770-775).
- Petrov, A., & Kokinov, B. (1998). Mapping and access in analogy-making: Independent or interactive? A simulation experiment with AMBR. *Advances in analogy research: Integration of theory and data from the cognitive, computational, and neural sciences*, 124-134.
- Schlag, I., & Schmidhuber, J. (2018). Learning to reason with third order tensor products. In *Advances in Neural Information Processing Systems* (pp. 9981-9993).
- Van der Velde, F., & De Kamps, M. (2006). Neural blackboard architectures of combinatorial structures in cognition. *Behavioral and Brain Sciences*, 29(1), 37-70.

List of Symbols, Abbreviations and Acronyms

RoleMap – The name of the model of categorization

VARs – The name of the connectionist application of relational thinking