Role-Governed Categorization and Category Learning as a Result from Structural Alignment: The *RoleMap* Model

Yolina A. Petrova, Georgi I. Petkov

Abstract—The paper presents a symbolic model for category learning and categorization (called *RoleMap*). Unlike the other models which implement learning in a separate working mode, role-governed category learning and categorization emerge in *RoleMap* while it does its usual reasoning. The model is based on several basic mechanisms known as reflecting the sub-processes of analogy-making. It steps on the assumption that in their everyday life people constantly compare what they experience and what they know. Various commonalities between the incoming information (current experience) and the stored one (long-term memory) emerge from those comparisons. Some of those commonalities are considered to be highly important, and they are transformed into concepts for further use. This process denotes the category learning. When there is missing knowledge in the incoming information (i.e. the perceived object is still not recognized), the model makes anticipations about what is missing, based on the similar episodes from its long-term memory. Various such anticipations may emerge for different reasons. However, with time only one of them wins and is transformed into a category member. This process denotes the act of categorization.

Keywords—Categorization, category learning, role-governed category, analogy-making, cognitive modeling.

I. INTRODUCTION

THE ability to classify entities is constantly manifested in our everyday life [1]. When we have the urge to pet a *dog* on the street, for example, usually first we identify that moving object as a *dog*. We can further infer that it eats *dog food* and knowing that *dogs* usually *bark*, we can also make a prediction that the same dog could *bark* any moment now [2].

Usually, to study how we categorize and learn new categories, the researchers use tasks focused on the intrinsic features of different category members. Importantly, the category members are rarely included in any relational structure [2], [3]. Those tasks have advanced our understanding about how we acquire and manipulate the so-called *feature-based categories* – categories which are relatively self-sufficient when defined. For example, we can explain what a *cat* is just by its intrinsic properties – it would be something that *meows*, has *whiskers*, *fur*, etc. Importantly, most *cat* members share those intrinsic features. Yet, we cannot generalize this understanding to all types of concepts, because in our everyday life we interact with category members which appear in relations with other entities. Such concepts are all *relational*, *role-governed* and *theme-based categories* [4], [5]. The commonality between them is that their category membership depends on their external relationships with other categories.

If we take a simple situation such as "*A cat hunting a mouse.*", *hunts* would be an example for a *relational category* [5]. *Hunts* cannot be explained only through its intrinsic features without any references to other categories, namely *predator* and *prey*. The *role-governed categories*, on the other hand, denote specific roles in a relational structure [1]. Thus, *cat* would be a member of the role-governed category *predator*, because it is an animal hunting another animal. Analogously, *mouse* would be member of the role-governed category *prey*. Finally, *theme-based categories* combine members performing different roles but relationally connected in the same event [6]. In this situation, as interacting members in the same event, both *cat* and *mouse* will comprise a single theme-based category.

The role-governed categories have recently attracted the attention of several authors [1], [5], [7]. Asmuth and Gentner [8], for example, reported that, according to the British National Corpus, around half of the most frequently used English nouns have role-governed meaning. There are also empirical evidences showing that, in some circumstances, people prefer to categorize through role-governed categories. Such are the findings of Goldwater et al. [7] who presented their participants with videos of novel objects moving in similar ways (like two objects chasing each other, while a third one stays impartial to the chasing event). In the test phase, they presented the same objects and asked the participants to classify which object (of two alternatives) goes best with the target one to form a category. According to the results, role-governed classification largely dominated the people's preferences.

II. MODELS OF CATEGORIZATION AND CATEGORY LEARNING

Even though the role-governed categories are frequently used, they have been overshadowed in the cognitive modeling domain. Various influential models such as the *Generalized Context Model* [9] and *SUSTAIN* [10] treat the categorization and category learning as depending only on intrinsic similarity between the features of the new instance and the stored category representations. The feature-based models in general rarely consider any relational information between the members of different categories, disabling them to account for

Yolina Petrova and Georgi Petkov are with the East and Central European Center of Cognitive Science; Department of Cognitive Science and Psychology, New Bulgarian University, 21 Montevideo Street, Sofia 1618, Bulgaria (e-mail: yolina.petrovaa@gmail.com, gpetkov@nbu.cogs.bg).

any relation-based categories.

The neural network models (especially the deep networks [11]) advanced the domain, demonstrating great success in classification tasks for various types of stimuli and categories. The separate hidden units in them may sometimes capture within-category structure [12]. However, often their knowledge is represented in a distributed way, making it hard to explicitly interpret this structure sensitivity, and more specifically, how exactly the input features are linked to categories on the output. Thus, even though the deep neural networks have a lot of practical uses, it is hard to evaluate their contribution to understanding the exact structure and use of the human concepts. Most importantly, at least for now, the neural networks cannot explicitly represent relations; that is why, the impact of the structural information cannot be applied.

Because the role-governed categories need structural information, it is natural to turn to the *structure-based models* of categorization and category learning. For example, *SEQL* [13] is a category learning model using the structural alignment process modeled in *SME* [14]. Unlike the feature-based models, *SEQL* works with structured descriptions of episodes. As new instances are given to the model, they are structurally aligned (comparison process known from the analogy-making domain) to all available generalizations (and eventually to the stored exemplars, if a suitable generalization was not found). The structural comparison (handled by *SME*) produces structural similarity scores, and the new instances are assigned to the category to which generalization the target was found to be most similar. Unfortunately, even though *SEQL* can handle relational representations with higher-order relations as well as feature representations, it is still not developed to learn or categorize any relation-based categories (including role-governed categories).

DORA [15], another analogy-based model, is supposed to learn relational concepts. Importantly, it does not do it directly. First, the model extracts common features of objects fulfilling the same role, treating them as role representations. Then, the roles are combined into a relational structure, which serves as a relational concept. Unfortunately, the model cannot use what it has learned (it cannot categorize). More importantly, *DORA* assumes that a relation could be learned only if its roles (arguments) have been learned in advance. Contrary to that assumption, empirical evidences show that role-governed categories are themselves instantiated when a new relational structure is acquired [16].

Kemp et al. [17] developed a model that also highlights the importance of the extrinsic relations for developing many of the relation-based concepts. However, their model relies on enormous computational power, making it not psychologically plausible enough.

To summarize, even though the relation-based categories (role-governed specifically) are of huge importance, they are highly underestimated in the cognitive modeling field.

III. *ROLEMAP'S STEPPING STONE*

The *RoleMap* model proposed here is a category learning and categorization model aiming to address the role-governed categorization gap in the modelling domain. The model is implemented in the cognitive architecture *DUAL* [18], initially developed to account for analogy-making process. Yet, *DUAL* assumes that its employed mechanisms are of such an importance that various higher cognitive processes can be implemented through them.

The *RoleMap* model rests on the idea that, in our everyday life, we constantly make structural alignments between the things that we know and see. As already noted, the structural alignment is a form of structural comparison known from the analogy-making domain. When different episodes are structurally aligned, a common description between them is produced in the form of coalition of mappings – where each mapping represents a single commonality. The model assumes that because some of the mappings are highly important (indicated from their activation level), they could stay for further use. This coalition forms the distributed representation of a newly created concept. That is how both the learning of new role-governed categories and the learning of new situational schema concepts emerge as a result from the structural alignment process. Categorization, on the other hand, results from the transfer of base knowledge (another analogy-making sub-process) missing in the new episode. More precisely, if there is a mapping between un-categorized target element and a base element which is part of a categorized object, then an anticipation pointing to the expectation that the un-categorized target element is also part of such object is formed (more details below).

The model's base knowledge is comprised of hundreds of hand-coded agents. There are semantic agents, representing the model's semantic knowledge, and episodic agents, representing the model's prior experience. Each individual agent carries some knowledge about a single entity. Thus, a single agent can hardly perform any significant work alone. On the contrary, the model's behaviour emerges from a huge number of local processing interactions between the agents, possible because of the specific links interconnecting them. All *RoleMap*-agents interact with each other through several basic mechanisms reflecting the analogy-making sub-processes. Importantly, the mechanisms, implemented in *RoleMap*, overlap in time and constantly influence each other, meaning that neither the category learning nor the categorization are separate from other processes – retrieval, mapping, etc.

We already said that *RoleMap* steps on structural comparison as a base for the categorization and category learning. Yet, the structural comparison is not happening between what is on the model's input as target and the whole base knowledge. Only relevant episodes are considered. Which knowledge is relevant enough to be included in the correspondences' search is determined by the *spreading of activation mechanism*. The *RoleMap* agents may be thought also as nodes in a classical neural network, and an activation (representing the relevance of the respective agents to the current context) spreads by a linear activation function with threshold for entering in the active memory:

new_activation = decay_rate_parameter * (w * old_activation + (1 - w) * net_input)

The parameters are fixed as following: decay rate parameter: 0.90; weight for the old activation level w: 0.75; respectively weight for the net input (1 – w): 0.25; threshold for entering into the active memory: 0.20. During some of the

simulations, however, we added a random noise to the weight of the links, in order to simulate random variations of the knowledge base and to obtain statistical results (bellow in the article). The threshold for entering in the active memory is fixed at 0.2.
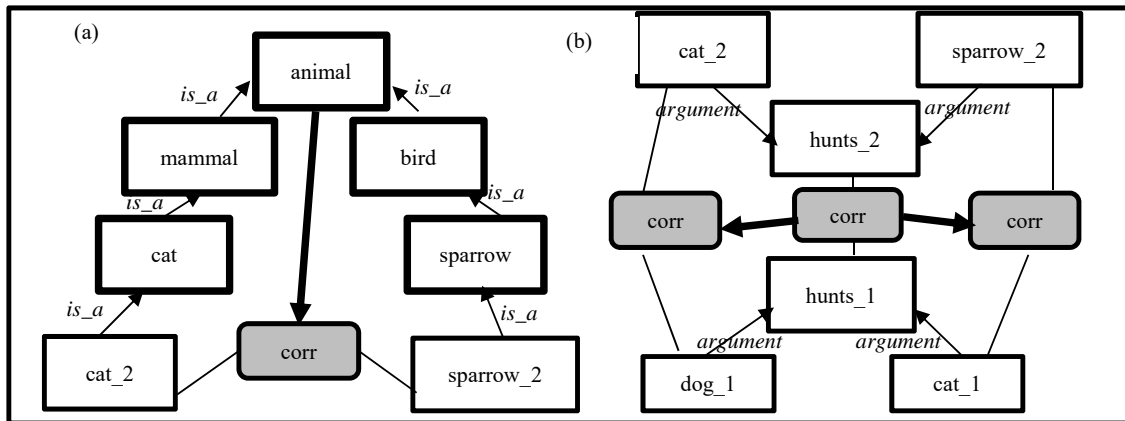

Fig. 1 Creation of new mapping-agents (correspondences). (a) points to the marker-passing mechanism – cat_2 and sparrow_2 correspond to each other because they are both animals. (b) points to the structural correspondence mechanism – dog_1 corresponds to cat_2 because these are respective arguments in the already mapped relations hunts_1 and hunts_2; the same applies for cat_1 and sparrow_2
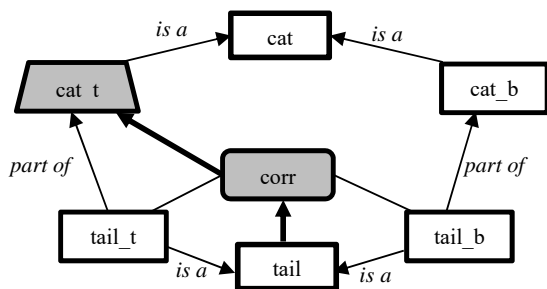

Fig. 2 Creation of anticipations. When the target agent tail_t finds a correspondence with the base one tail_b, the mapping agent (corr) registers that the base element is part of an already categorized object (cat_b) and creates an anticipation (the gray trapeze cat_t) that the target tail_t is a part of a cat as well

As soon as the model receives target agents on its input, by using the agents' outgoing and incoming links, this mechanism starts activating the target's closest neighbors. Thus, they are the first to get enough activation to enter the model's working memory. The activation continues to spread, and as the time passes, relevant base episodes are being retrieved. That connectionist side of the model allows trading-off between the flexibility and the efficiency of the model. It also allows for the simulation of various context effects.

As soon as an agent enters the model's working memory, it can participate in the work of the other mechanisms. Each instance agent (representing a concrete manifestation of a concept) checks for another relevant (i.e. active enough) instance with whom it shares a concept upward in the hierarchy. If there is such an instance, that would mean that the two are semantically similar, because of which the *marker*

passing mechanism forms a mapping (correspondence) between them (Fig. 1 (a)). If two relations are found to correspond to each other, their arguments should also correspond. Thus, based on mappings between relations, the *structural correspondence mechanism* forms new mappings between their respective roles. The mappings between the respective relational roles denote the structural commonalities between them (Fig. 1 (b)).

Dynamically, with the appearance of every mapping, the *constraint satisfaction mechanism* sets justification and inhibitory links between the supporting each other and the competing with each other mappings. At the end, the mappings part of the most coherent to the target global structure take lead in their activation and prevail. Because the mappings are treated as some initial state of new concepts, if the activation of a mapping exceeds a pre-defined threshold, the model considers this mapping as capturing highly important knowledge, which can be used in other situations. Thus, the *abstraction transformation mechanism* transforms this mapping into a concept. When presented on the input, the target agents are not bound by a single agent. However, when a set of interconnected with relations mappings is transformed into concepts, then one additional concept is created – capturing the whole schema. It may be thought as kind of a binding node for the lower level concepts of the part-of hierarchy. That is how *RoleMap* learns new concepts. The newly learned concepts influence the model's performance in the same way that the hand-coded ones do. To categorize a novel target instance, the model uses the *anticipatory mechanism*. Basically, when two instance agents are mapped, the model tries to "fill the gaps" in the target episode. That

happens through transfer of base knowledge in the form of anticipations (Fig. 2).

Like the creation of mappings, several competing category anticipations can be created, the constraint satisfaction mechanism again makes sure that a single anticipation is promoted for a winner. In turn, the winning anticipation is transformed into an instance, which could be interpreted as an act of categorization. When the model categorizes the whole target episode as a situation (for which it has a schema concept), its respective parts are categorized accordingly.

A. RoleMap's Competing Pressures

There is something important that should be noted. When the number of instances of a certain concept increases, the number of mappings between target entity similar to those instances also increases. In turn, the overall inhibition between the mappings increases, and consequentially the chance for a single mapping to be promoted as a concept decreases. At the same time, such mappings support one and the same anticipation. Importantly, the anticipation's overall support does not increase. This is achieved by a balance between the model's parameters: the strength of each mapping decreases proportionally to the number of its competitors.

This peculiarity has an important implication. As the number of concepts' instances increases, the probability to categorize a new entity into the same category also increases (the mappings are responsible for the process of category learning, whereas the anticipations – for the categorization into old categories). Based on this, we could model various empirically known category learning and category use data – for example, frequency effects, base-level effects, etc.

IV. ROLEMAP IN ACTION

A. Simulation1: Learning a Role-Governed Category

To be easily traceable how everything comes into play, so the model can learn and categorize, we simulated a simple set up. To start its work, we provided the model with a single base episode of "A neighbor milking a cow." (Fig. 3). Concepts for neighbor, grandmother, cow, pig, chicken, animal, etc. and relations such as feeds and milks were given to the model as prior semantic knowledge. Thus, the base elements point to the respective concepts with is_a links.

In the moment where we introduced the model's input with a target situation of "A grandmother feeding a chicken", it started searching for a base episode which corresponds to that one. The spreading of activation mechanism retrieved the base episode of "A neighbor milking a cow.", because many of the agents comprising the two episodes are close in the semantic network. For example, the relations feeds and milks are subclasses of cares; both the grandmother and the neighbor are type of person, etc.

With the retrieval of the base agent representing the grandmother, the marker passing mechanism created a mapping between the grandmother and the neighbor – because they share the concept person upper in the hierarchy. Similar semantic commonalities were established for the

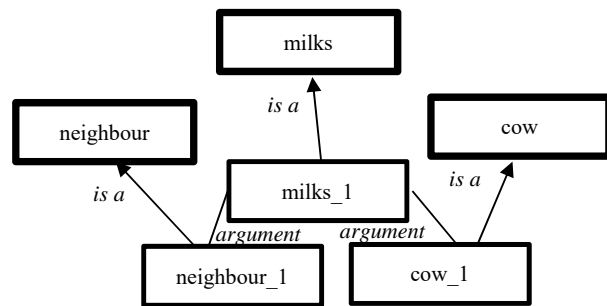relations feeds and milks – because they are both type of care.

Fig. 3 The hand-coded situation, representing "A neighbor milking a cow." The agents neighbor_1, cow_1 and milks_1 are instance-agents; the others – concept-agents. The agent milks_1, as well as its respective concept milks are relations.

With the creation of the mapping between the two relations (feeds and milks), the structure correspondence mechanism made sure that their corresponding roles are mapped as well. Thus, the mapping between grandmother and neighbor became more important (i.e. was justified by two reasons). Except for their semantic commonality, the two agents shared structural commonality as well – they performed similar roles in the structures that they were part of (both took care of something else). Accordingly, the second relational role – filled by chicken and cow – was also mapped. Eventually, the model retrieved information that they are both animals, because of which the mapping's strength increased.

The simulation, being described, consisted of a base knowledge with single episode in it. Therefore, the formed mappings all supported each other (being in the same structure) and had no competitors. Because of that, one by one, the three mappings reached the transformation threshold. For each mapping exceeding that threshold, the model underlined the commonality between the mapped elements and asked for a name of the concept which was to be created[1] – so it received the names domestic animal (for the mapping between the chicken and the cow), animal care (for feeds and milks) and animal caregiver (for the mapping between the grandmother and the neighbor).

The new concepts domestic animal and animal caregiver are examples of role-governed categories. In both cases, the category was created because both agents who created the initial mapping (based on which the concept was established) played one and the same role across the two situations.

In addition, the model created a new "schema" concept representing the common parts of the two situations (Fig. 4). We called it – domestic animal care. The two mapped episodes became instances of that schema. Accordingly, the three newly created concepts – animal care, animal caregiver and domestic animal – became parts of the "schema" (with bi-directional links to and from it). Even though the concepts

---

[1] The model will work the same way with names like agent001, agent002, etc. However, there is a procedure, allowing the user to give mnemonic names for better description. The model can also set formal names, working without any human intervention.

above were initially created as mappings, once they were transformed into concepts, they could have and did affect the

subsequent behavior of the model, just as the rest of its (hand-coded) semantic knowledge.
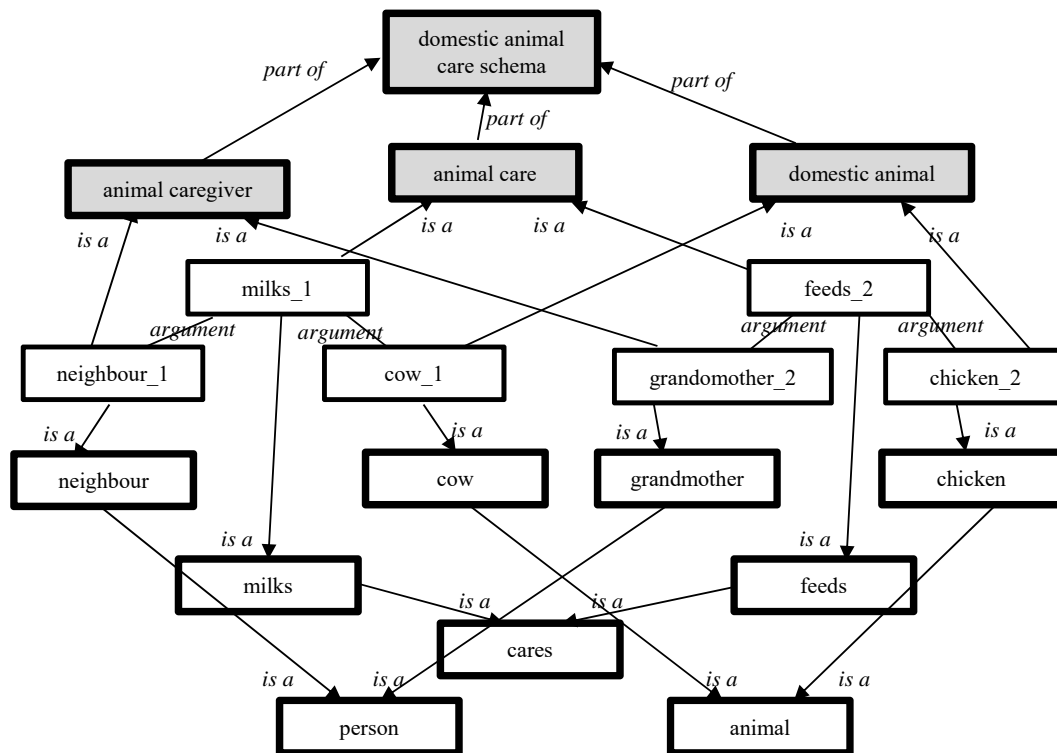

Fig. 4 Illustration of Simulation 1. The hand-coded situations "Neighbour milks cow" and "Grandmother feeds chicken" map each other because of the shared concepts upward in the class hierarchy (on the bottom part of the picture). As a result, new agents are created (with gray on the upper part of the picture), as well a new schema concept

Note that the mechanism for concepts creation may seem to be too unconstrained in the current implementation of the model, meaning that the model probably will create more concepts that people would. Thus, we plan additional mechanisms for learning. Just by a simple decay mechanism, the concepts that the cognitive system does not use often will disappear. For the current simulation, the important thing is the structure based creation of potential concepts.

B. Simulation 2: Categorization through Newly Learned Categories

RoleMap continued its work and demonstrated its ability to use what it has learned by categorizing a new target situation. The previous target episode of "*A grandmother feeding a chicken.*" was already incorporated in the model's long-term knowledge. In addition, the newly learned categories were also incorporated into the semantic knowledge. In that moment we gave to the input a new target episode – "*A girl brushing a horse.*". Analogously to what was described above, the agents, representing the new target started activating other agents to which the target episode can be compared. The structural alignment started producing mappings denoting what goes with what. Contrary to the previous simulation, the mappings inhibited each other decreasing their activation level. During that time, the anticipatory mechanism started its

work. Each mapping created an anticipation for the situation that the model is witnessing. The mappings between the target *girl* and the base *grandmother* and between the target *girl* and the base *neighbor* both created an anticipation that the whole target episode is instance of the *domestic animal care* schema (which was created in the previous simulation). That happened because when part of the target situation is missing (in that case the situation category), it could be potentially transferred and expected. The other mappings tried to create the same schema anticipation. But, because such an anticipation already existed, only additional justifications were added to the first one[2]. When the anticipation became active enough, it was transformed into an instance, meaning that the model made a categorization (in a certain context, it may happen that some of the mappings win before the anticipation. In such case, a new category will be made). Subsequently, the model incorporated the new target episode by correctly adjusting it as an instance of the *domestic animal care* schema concept. The corresponding bi-directional links were set up accordingly for all target parts as well. In that case, the new *girl* was categorized as an *animal caregiver*; the new *brushes* became

---

[2] During this simulation, there were no any competing anticipations. If there were, they would have inhibited each other. In the same way, as with the mappings' competition, the interplay between the supporting and inhibiting each other anticipations would have resulted in a single winning anticipation.

instance of the *animal care*, and the *horse* was categorized as a *domestic animal*.

TABLE I
DESCRIPTION OF THE SITUATIONS ENCODED FOR THE SIMULATION 4 AND THE RESULTS

Set	Base situation 1	Base situation 2	Target situation	Categorization of the Target situation in %s
First set	*Grandmother feeding a chicken.*	*Grandmother feeding a pig.*	*Grandmother feeding a cow.*	*75%*
Second set	*Grandmother feeding a chicken.*	*Neighbour milking a cow.*	*Grandmother feeding a horse.*	*11%*
Third set	*Grandmother feeding a chicken.*	*Neighbour milking a cow.*	*Neighbour feeding a horse.*	*30%*
Fourth set	*Grandmother feeding a chicken.*	*Neighbour milking a cow.*	*Girl brushing a horse.*	*52%*

C. Simulation 3: Learning a Role-Governed Category through Episodes with Higher-Order Relations

RoleMap can also deal with more complex structures containing higher-order relations. To demonstrate that ability, we gave the episode of "*A grandmother feeding a chicken which causes the chicken to give eggs.*" as target. The episode "*A neighbor feeding a cow which causes the cow to give milk.*" was encoded as a base. Again, long-term concepts such as *grandmother*, *chicken*, *cow*, *milk*, *eggs*, etc. were encoded. Some relational concepts *feeds*, *causes*, as well as *gives* were also predefined. The base episode was represented through 6 instance agents (*feeds_1*, *neighbor_1*, *cow_1*, *causes_1*, *gives_1*, and *milk_1*). Importantly, the relational instance *feeds_1* and the relational instance *gives_1* were both encoded as arguments of the relational instance *causes_1*. The target episode was represented in an analogues way (through the agents: *feeds_2, causes_2*, etc.). As in all previous simulations, when the target agents appeared on the input, the spreading activation mechanism caused the relevant base agents to enter the working memory one by one. Eventually, the whole base episode and the corresponding concepts to the target and base episodes also entered the working memory. Meanwhile, various mappings between the instances started to be created – total of six mappings emerged. Eventually, all mappings were additionally activated through justification links from up to three supporting coherent mappings. For example, the mapping between the relational concepts *feeds_1* and *feeds_2* was supported by the mappings between *neighbor _1* and *grandmother_2* and *causes_1* and *causes_2*, because they all formed a coherent global structure.

As in the first simulation, there were again no inhibitory pressures coming from competing mappings. Thus, all created six mappings survived long and active enough to exceed the upper activation threshold and were transformed into concepts. The newly created concepts were called *animal feeding, animal feeder, domestic animal, animal product*. In addition, a new relational concept was created with name *gives food* (with domestic animal and *animal product* as arguments). This concept was itself an argument of the relational concept *cause to give food*. The model adjusted each newly created concept with the mapped target and base elements as instances.

D. Simulation 4: Statistical Results from Many Runs

To explore the model's behavior in more detail, we designed several sets (each containing three situations about people caring for various animals), similar to the described above. Each set was run sequentially, presenting the situations one by one. The model always stored the first situation and created new concepts during the second, and finally, the third situation was either categorized onto the already created schema-concept, or created a new schema-concept, combining the third situation with one of the two situations presented before that. We explored exactly this decision of the model – whether it will categorize the third situation, or it will form a new category. This depended on the competition between the anticipations (pressuring to categorize) and the mappings (pressuring to create new categories) – which will pass its threshold first. The sets of situations differed according to the similarities between the three situations. In the first set, there were three highly similar situations. In the second set, the first two situations were relatively different, while the third situation was similar to the first one. This simulated initial creation of a sparse concept followed by a situation more similar to one of the two categorized situations. In this case, we expected to form the model higher tendency to create new sub-category, combining the two similar exemplars, diverging them from the different base. In the third set, we changed the third situation only (the target), making it a combination of the other two. Thus, we expected the tendency for categorization to increase in exchange to creating new concepts. Finally, in the fourth set, we again kept the same two bases and introduced a target situation, very different from both bases. All the sets are presented in Table I[3].

Finally, in order to obtain statistical results, we run each of the sets 100 times, by putting random noise in the strength of all links of each situation (noise from N(0, 0.25)), as well as the initial activation of all agents (noise from N(0, 0.05)). We counted the percentage of times that the model categorized the third situation instead of creating new concepts between it and some of the base situations (i.e. the anticipation won before the mappings). For the first set (the control one), the parameters were fitted to produce 75% of categorization, so the tendencies from the other sets could be explored.

The results were as follows (Table I): the model categorized the third situation in 11% for the second set; in 30% for the third set; and in 52% of the cases for the last one (note that the exact values do not have much sense, only the tendency does).

---

[3] The simplicity of the bases was deliberate. We wanted to explore the influence of a single similarity or dissimilarity between them, namely the length of the path between them through the class hierarchy. However, any other similarities or dissimilarities between the entities influence the work of the model in the same way.

Thus, the results confirmed our expectations – the number of categorizations increased when all situations were different. The categorization decreased (in exchange for the formation of new categories), when the target situation was similar to only one of the base situations.

E. Simulation 5: Varying the Thresholds by Keeping One and the Same Episode Descriptions

Basically, one of the most important parameters for the trade-off between categorization and category learning is the difference between the respective thresholds for creating mappings into concepts and for transforming anticipations into instances. Thus, we took the control set of the three similar episodes (described above) and run it 100 times varying the links and the initial activations of the agents. However, this time we also varied the thresholds values for categorization and/or category creation. Not surprisingly, when the two parameters were equal, the model tended to categorize and to form new categories in equal rate. With the increase of only one of the thresholds, the tendency shifted to the respective direction.

What was more interesting, however, was to explore only the cases of category learning – which of the mappings were transformed into concepts first. We excluded all runs finished with categorization and explored the rest. When we made the model biased to categorize (thresholds 15 and 10 for transformation of anticipations and mappings respectively), but it still created new categories, in 70% of the cases the arguments of the relation (*grandmother* or specific *animal*) were transformed. Vice versa, when the model was biased to learn new categories (the opposite pattern of thresholds), in all 100% of the cases when new categories were learned, the first one was about the relation *feeding*.

Extremely speculating (and planning to work on this in the future), we link this result with the reviewed results of Gentner Rattermann [19] who found out that when children learn new concepts, they prefer superficial similarity, whereas adults, who rarely create new concepts – prefer structural ones. In *RoleMap,* this result emerged naturally from the inherited from the DUAL architecture property the structural pressures to evolve slower in the time.

V. SUMMARY AND FUTURE WORK

This paper explored how the *RoleMap* category learning and categorization model can learn a new role-governed category and further use it to categorize new members in the same category. To summarize, the model successfully demonstrated a single-shot learning, which is extremely difficult for the neural networks for example. More importantly, the model accounts both for category learning and for categorization in already existing categories. More specifically, it proposes an idea about how to resolve the trade-off between these two controversial tendencies.

Even though the model is focused on relational and role-governed categories, it is not limited to them. The same principles for evaluating similarity between a target episode and similar base ones could be applied (at least theoretically)

for the feature-base categories as well. In addition, the model does not assume any principle difference between how people separate space into objects and group them into categories and how they separate time into situations and events and group them into schemas.

Learning mechanisms should put additional constraints on category-learning. Finally, the scalability of the model should be tested by larger knowledge bases that are not hand-coded. We are aware of the limitations that the hand-encoded knowledge base imposes, but we consider this as an important first step for the model's validation.

Even though the model is still in its embryo phase and still has lots of limitations, the arguments above makes us think that the described approach can be fruitful for further exploration of the way that people form and use role-governed categories. We further plan to address problem such as categorization on different levels of abstraction and context-sensitivity effects.

REFERENCES

[1] Markman, Arthur B., and Brian H. Ross. "Category use and category learning." *Psychological bulletin* 129.4 (2003): 592. W.-K. Chen, Linear Networks and Systems (Book style). Belmont, CA: Wadsworth, 1993, pp. 123–135.
[2] Goldstone, Robert L., Alan Kersten, and Paulo F. Carvalho. "Concepts and Categorization" In I. B. Weiner (Ed.), *Handbook of Psychology*, Second Edition (2012).
[3] Yamauchi, Takashi, and Arthur B. Markman. "Learning categories composed of varying instances: The effect of classification, inference, and structural alignment." *Memory & Cognition* 28.1 (2000): 64-78.
[4] Gentner, Dedre, and Kenneth J. Kurtz. "Relational categories." Categorization inside and outside the lab (2005): 151-175.
[5] Markman, Arthur B., and C. Hunt Stilwell. "Role-governed categories." *Journal of Experimental & Theoretical Artificial Intelligence* 13.4 (2001): 329-358.
[6] Lin, Emilie L., and Gregory L. Murphy. "Thematic relations in adults' concepts." *Journal of experimental psychology: General* 130.1 (2001): 3.
[7] Goldwater, Micah B., Rebecca Bainbridge, and Gregory L. Murphy. "Learning of role-governed and thematic categories." *Acta psychologica* 164 (2016): 112-126.
[8] Asmuth, Jennifer A., and Dedre Gentner. "Context sensitivity of relational nouns." *Proceedings of the Annual Meeting of the Cognitive Science Society*. Vol. 27. No. 27. 2005.
[9] Nosofsky, Robert M. "Attention, similarity, and the identification–categorization relationship." *Journal of experimental psychology: General* 115.1 (1986): 39.
[10] Love, Bradley C., Douglas L. Medin, and Todd M. Gureckis. "SUSTAIN: a network model of category learning." *Psychological review* 111.2 (2004): 309.
[11] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." *Nature* 521.7553 (2015): 436.
[12] Rips, Lance J., Edward E. Smith, and Douglas L. Medin. "Concepts and categories: Memory, meaning, and metaphysics." *The Oxford handbook of thinking and reasoning* (2012): 177-209.
[13] Kuehne, Sven, et al. "SEQL: Category learning as progressive abstraction using structure mapping." *Proceedings of the 22nd annual meeting of the cognitive science society*. Vol. 4. 2000.
[14] Falkenhainer, Brian, Kenneth D. Forbus, and Dedre Gentner. "The structure-mapping engine: Algorithm and examples." *Artificial intelligence* 41.1 (1989): 1-63.

[15] Doumas, Leonidas AA, John E. Hummel, and Catherine M. Sandhofer. "A theory of the discovery and predication of relational concepts." *Psychological review* 115.1 (2008): 1.

[16] Goldwater, Micah B., Arthur B. Markman, and C. Hunt Stilwell. "The empirical case for role-governed categories." *Cognition* 118.3 (2011): 359-376.

[17] Kemp, Charles, et al. "A probabilistic model of theory formation." *Cognition* 114.2 (2010): 165-196.

[18] Kokinov, B. "The context-sensitive cognitive architecture DUAL." *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*. 1994.

[19] Gentner, Dedre, and Mary Jo Rattermann. "Language and the career of similarity." *Perspectives on language and thought: Interrelations in development* 225 (1991).