

Review on

“Analogies and understanding intentions”, a thesis submitted by Luiza Shahbazyan in partial fulfillment of the requirements for the degree of Ph.D. in Cognitive Science and Psychology

Shahbazyan presents a 167 pages MS for partial fulfillments of the requirements for a Ph. D. degree. The page 1 is a title one and the page 2 is dedicated to Ass. Professor Boicho Kokinov, the first Ph. D. advisor of Luiza Shahbazyan. The page 3 is an abstract, the pages 4 - 8 are a table of contents and the page 9 – acknowledgements. The page 10 is a list of figures and the pages 11 – 12 are a list of tables.

The chapter 1 entitled “Introduction: Analogies and understanding intentions” (pages 13 – 19) is a short review of the basic questions studied and discussed in this Ph. D. thesis such as: “How we are able to infer the mental states of others” or “How people ascribe content to the intentions (the plans of actions in pursuit of a goal) behinds others’ actions.” In this chapter Shahbazyan also presents different approaches addressed the question of how people understand the intentions of others: 1) Direct perception theories, 2) Simulation theories, 3) Theory, theories. Direct theories suggest that we employ sensory information derived from our perception of the action. Simulation theories suggest that we use episodic memories – representations of how the cognizer acts in such a situation. Theory theories, on the other hand, suggest that people resort to semantic memory in order to retrieve some generalized knowledge that might be relevant to the situations.

There is also a fourth type of knowledge –relational which could be defined as “a binding between a relation symbol and a set of ordered tuples of elements” The author has noted that there is evidence that action perception involves coding motor information bound by relational structures and people represent perceptual information in terms of higher order relations. This allows suggesting that relational knowledge is intrinsic to the information that is supposedly used to generate inferences about others’ intentions. Shahbazyan has noted that this raises the question to what extend similarity in the relational structures of the current (target situation) and some prior source knowledge (perceptual, semantic or episodic) determines the inferences that are going to be generated. The process of finding relational similarity between two or more representations is known as “analogy making” and there is robust empirical support for its role in humans’ ability to generate inferences about unknown aspects of the world. Thus, the aim of this thesis is to explore the role which analogy making plays in understanding others’ intentions in ambiguous situations.

The chapter 2 entitled “Major approaches to understanding mental states” (pages 20 – 40) is a more detailed review devoted to the different theories explaining how people understand the intentions of others, mentioned in Chapter 1. The first type of theories discussed in Chapter 2 are so called “Direct perception theories”. According to these theories we are able to perceive directly states such as intentions, desires and beliefs without any additional processing. Such information can be derived from the body movements or the face expression. Shahbazyan has noted that under some conditions “sensory data seem to unambiguously reveal the intentions behind the action, but direct perception is not able to account for the complex and ill-constraint inferences that we generate every day”.

The other theory described in Chapter 3 is so called “Simulation theory”. It is assumed “that perceivers use their own cognitive system to pretend that they are in the situation of the other person, so that they can simulate what the other person would do. The output of the simulation is assigned to the other person. We still need to have knowledge about the possible intentions of the other person, but this knowledge does not need to be generalized”. According to this theory we “would not expect any influence of information such as the mental state of the perceiver or his contextualized knowledge about the target situation. Nevertheless, such influence has been extensively documented”. Shahbazyan has pointed out that “this approach limits mindreading only to situations, which are identical or at least very similar to our previous experiences, which is a significant and not a very realistic constraint”.

The next type of theories discussed in Chapter 3 is so called “Theory Theory”. According to this theory mentalizing is dependent on prior knowledge and “understanding of mental states depends on a rich set of mental representations containing substantial amounts of information (or, sometimes, misinformation) about mental states and their interactions with environmental stimuli, with behaviour, and with each other”. The author has cited experimental evidence in support of this theory, as well as findings which might be interpreted as being not in accordance with this theory (slow process of learning from “repetitive experience of concrete episodes”).

The author discussed also so called “Hybrids between Theory Theory and Simulation theory” models which might be interpreted as an attempt to explain how people understand the intentions of others. According to these Hybrid models “some aspects of mentalizing are served by simulations, while others are served by theorizing”. It has been pointed out that “these theories face several challenges. First, by proposing that people use both theorizing and simulation, the hybrid models inherit some of the limitations of each of the individual mechanisms that have been already outlined above, including ignoring the role of relational

knowledge to derive inferences. Second, hybrid models should specify the conditions under which each of the mechanism is employed, which is problematic and leads to conflicting claims or claims that do not lead to testable predictions. Third, the hybrid models imply that people selectively use either episodic or semantic memory despite evidence that the two systems interact. The hybrid models do not address the possibility for interaction”.

The fourth type of knowledge – relational, suggested as intrinsic to the information probably used to generate inferences about others’ intentions is discussed in Chapter 3 entitled “Relational Knowledge and Analogical Inferences” (pages 40 – 55). The author has noted that the ability to acquire and manipulate the relational knowledge is considered as a power tool of the human brain. This ability “allows re-presenting the sensory stream in terms of non-obvious relations such as cause, preventing, executing, promising and also wanting, feeling and intending. Relations allow us to express ideas such as causations, chains of implications, counterfactuals and others, which will not be possible otherwise. Furthermore, they enable us to think analytically beyond what is given in the situation. But probably the most important characteristic of relational knowledge is the role it plays in analogical inference – inferring unknown properties of the current situation based on relational similarity to prior episode or schema”. Shahbazyan has also cited some experimental evidence in support of the idea that “people may use analogy spontaneously to infer unknown aspects of a situation or a solution to a problem based on previous example”.

In the Chapter 4 entitled “Linking Analogy Making and Understanding Mental States” the author demonstrates parallelism between analogy making and understanding intentions. She notes “that understanding intentions is very much like role-based relational reasoning - differentiating relational roles from the entities that fill those roles, which is observed in analogy making. The second parallel between analogy making and understanding intentions is that both processes involve generating inferences based on the perceived similarity between a target situation and preexisting knowledge. Similarly to analogy making, understanding intentions seems to depend on activation of relevant past knowledge, either concrete or generalized”.

Shahbazyan notes that the question of the role of relational knowledge in understanding intentions is not systematically studied. Thus, the aim of the 5 studies that follow is to explore the role which analogy making plays in understanding others’ intentions in ambiguous situations. In these studies the base information available to the participants was varied. Based on the suggested role that analogy making plays in transfers from both concrete and generalized knowledge, it was hypothesized that participants will be able to transfer

intentions from structurally similar, superficially dissimilar episodes. It was also hypothesized that participants will be more likely to transfer intentional information from episode that shares relational information, rather than from episode with which it shares common objects or attributes such as mood valence.

The next Chapter 5 entitled “Experimental Studies” (pages 64 - 124) describes 5 different experiments aimed to study the role of analogy making in understanding of other’s intentions.

Experiment 1 is aimed to test the hypothesis that single episodes which are structurally similar to an ambiguous target situation will influence the attribution of intention to the ambiguous action in the target. It was predicted that if the target situation is ambiguous, relevant relational prior concrete knowledge will facilitate the representational process by ‘filling-in’ missing information. This follows from the suggestion that relational knowledge and analogical processing play a role in intentional understanding. In agreement with the hypothesis of the study, there was a significant interaction between *similarity* and *base content* over the ratings of the *negative target intention*. There was no significant interaction between *base similarity* and *base content* over the ratings of the *positive target intention*. Participants rated the *negative target intention* significantly higher when the target was preceded by structurally similar negative base than when it was preceded by superficially similar (marginal difference), dissimilar story or a control condition. They rated the base inconsistent (positive) *target intention* significantly lower when the target was preceded by structurally similar negative base than when it was preceded by superficially similar, dissimilar story (marginal difference) or a control condition, suggesting interaction between *base similarity* and *base content* over the ratings of the positive *target intention*. Nevertheless, an effect not initially expected, was observed: Rather than observing the positive structurally similar episode increasing the ratings of the positive *target intention*, it was observed that the negative structurally similar episode decreased the ratings of the positive *target intention*. At the same time, the positive structurally similar episode did not exert effect on the ratings of the positive *target intention*, nor on the rating of the negative *target intention*.

Experiment 2. In this experiment the role of structural similarity was also explored by two corrections of the procedure. A measure of aggression was added to the experimental tasks in order to control for any potential role of aggression in creating bias towards negative intentions that is not induced by the experimental manipulation. A forced choice response as a more sensitive measure of the effects of the structural similarity, instead of scale ratings was provided, because it was found to assess higher level processing of the information rather than

operation of learnt schemas. The results of Experiment 2 replicated the effect of negative structurally similar base on increasing the preference towards negative intention. No effect of the positive structurally similar story on increase of preference towards the positive base-consistent *target intention* was observed. Quite the opposite, there was a tendency participants to prefer the base-inconsistent negative intention after having seen the positive, structurally similar story.

Experiment 3. Experiment 3 aims to answer the question to what extent the effect of the negative structurally similar base in Experiments 1 and 2 is aided by activation of a negative stereotype. The hypothesis that the positive analog will increase the ratings of the positive *target intention*, yielding an effect of *base content* on these ratings was tested. To summarize, in line with the hypothesis of the study, there was a significant effect of *base content* on the rating of the positive *target intention*, reflecting higher ratings of this intention in the positive analog condition than in the negative analog condition. This confirms the prediction that activation of a positive or negative stereotype accounts for why the negative analog story exerts effect when the main characters are wolves, but the positive analog exerts effect when the main characters are ghosts.

Experiment 4 aims to answer the question to what extent the preference towards negative intentions in the structurally positive condition in Experiment 2 is a result of activation of a negative stereotype. The results of the experiment demonstrated that the inverted effect – preference towards negative intention after seeing a positive analog, observed in Experiment 2 remained even after the characters of the base story were altered in order to activate a positive stereotype. Furthermore, seeing a negative analog does not lead to preference towards negative intentions similarly to the lack of effect of the negative episode on the ratings of the negative intention in Experiment 4. Similarly to Experiment 3, participants memorized the control/neutral base episode significantly worse than the analogs.

Experiment 5 was aimed to further scrutinize the proposed influence of deeper processing in producing the inverted effect in Experiments 2 and 4. More specifically, it was suggested that the inverted effect of the positive analog (increased preference towards the negative intention) in Experiment 2 and 4 was due to participants' more thorough processing of the material due to the forced choice format. The main hypothesis of the study was that there would be an effect of *base task* on the ratings of negative *target intention*, reflecting an increase of the ratings of the negative intention after inducing both positive and negative schema. In line with the hypothesis of the study, comparison between two negative analogical bases do increase the rating of the base-consistent intention, demonstrating that people are

more likely to generate analogical inferences that affect the rating of the *target intention* following manipulation that enhances the encoding of the common relational structure. people tend to rate the negative base-inconsistent intention significantly higher after they have compared two positive structurally similar bases than after they have only summarized them. The results of the current experiment suggest that this “inverted effect” is most likely associated with altered evaluation of the intentional inferences as a result of the analogical processing.

In Chapter 6 entitled “General discussion” (pages 125 - 140) Shahbazyan has summarized the main conclusions following from the experimental results, as well as considers some open questions, implications, limitations and directions for further research. The main conclusions are:

Support for the hypothesis that analogical inferences play a role in understanding others’ intentions. In a series of experiments it was demonstrated that analogical inferences, i.e. inferences that follow from an analogous prior situation, influence the interpretation of a character’s intention in an ambiguous target situation, but their impact interacts with three other factors: depth of processing, evaluation, and activated stereotypes.

Enhanced understanding of the mechanisms of analogy making. The results contribute to current analogy research by showing that analogical inferences based on relational similarity play a role in understanding new situations, even without explicit instructions. The results of the experiments provide support for the idea put forward by Kokinov and Petrov (2000) that once people encounter a given situation, this situation activates both concrete and abstract prior knowledge. Last, the presented results bear relevance to the role of the process of evaluation and element alignment in analogical processing.

Linking analogy research and the traditional studies on social cognition. There is a general agreement among theorists in social cognition that understanding mental states in ambiguous situations requires filling-in of the missing information from prior knowledge.

Providing a valuable alternative to the existing approaches. The results, obtained in this research, cannot be accounted by Simulation theories and Theory theories, as well as by Direct perception theories.

Pages 141 - 158 are a list of references which contains 179 articles cited in this Ph. D.

Pages 159 – 162 contain Stories used in the experiments in English.

Pages 163 – 167 contain Stories used in the experiments in Bulgarian.

I have also some critical remarks. First, at some places the text is too wordy and this makes it reading difficult. Second, by one hand, the chapters 1, 2, 3 and particularly chapter 4

(pages 13 – 63, i. e. 50 pages) are a review of the theories concerned with understanding others' intentions and the results interpreted in the light of these theories. By the other hand, chapter 5 (pages 64 – 124, i. e. 60 pages) is a description of the experiments and the results obtained. In my opinion such a review is too extensive in comparison with the description of the experiments and the results. The balance between the different parts of the Ph. D. thesis would be better if the information presented in chapters 1, 2, 3 and 4 was presented in a more concentrated manner. Third, as it was already noted, the results of all experiments support the suggestion that analogy making plays role in understanding others' intentions. However, it might be assumed that human thinking is much more sophisticated process to be limited to some form of analogy making only in understanding others' intentions. Might be suggested that depending on the conditions, situations and tasks, our brain applies not one, but a number of algorithms in understanding intentions of others. This was briefly commented by the author (page 135), but the question remains open. Thus, Shahbazyan's research might be interpreted as an attempt to reveal one side only of the process of understanding others' intentions.

A few minor errors were also found, most of which involve typographical mistakes that are easily corrected.

In conclusion, despite the critical remarks, my opinion about the thesis is definitely positive. The rationale of the study is sound and interesting experiments are carried out and unambiguous results are obtained. Luiza Shahbazyan has published four papers on the basis of his work, thus providing international visibility of the study.

The thesis of Luiza Shahbazyan definitely meets the requirements of the Cognitive Science Program for a Ph. D. degree.

31.05.2015

Assoc. Professor Dimitar Mitov
Sofia University "St. Kliment Ohridski"