Complexity in Spatial Reasoning

Marco Ragni (ragni@informatik.uni-freiburg.de)
Department of Computer Science, Georges-Köhler-Allee
79110 Freiburg, Germany

Thomas Fangmeier (thomas.fangmeier@cognition.uni-freiburg.de)
Center for Cognitive Science, Friedrichstraße 50
79098 Freiburg, Germany

Lara Webber (lara.webber@cognition.uni-freiburg.de)
Center for Cognitive Science, Friedrichstraße 50
79098 Freiburg, Germany

Markus Knauff (knauff@cognition.uni-freiburg.de)
Center for Cognitive Science, Friedrichstraße 50
79098 Freiburg, Germany

Abstract

We introduce a unified approach to account for the problems people have in spatial reasoning. This approach combines two theories: the *mental model theory* which aims to explain the deduction process, and the *relational complexity theory* which explains the processing complexity of the spatial relations needed in order to conceptualize the reasoning problem. We propose that a combination of these two theories can account for some of various errors found in spatial reasoning. We present two experiments in which we demonstrate that participants use the principle of *first free fit* to construct preferred mental models. We then formally implement these findings in the *Spatial Reasoning by Models* computational framework.

Keywords: Spatial reasoning; Mental Models; complexity; computational framework

Introduction

Everyday spatial reasoning is strongly connected to the extensive use of spatial relations which locate one object with respect to others. Examples of such relations include binary relations such as "to the left of", or "in front of", and even more complex relations like the ternary relation "in-between". A typical reasoning problem dealing with such spatial relations is to infer relations between objects from an incomplete description of a spatial configuration of objects. Hence, in the deductive reasoning process implicit relations between a series of objects are to be inferred from assertions describing the spatial configuration. An easy example is provided by the following problem:

The hammer is to the right of the pliers.
The screwdriver is to the left of the pliers.
The wrench is in front of the screwdriver.
The saw is in front of the pliers.

Which relation holds between the wrench and the saw?

The first four assertions are called *premises*, while the question refers to a possible *conclusion* that can be drawn from the premises. The *mental model theory (MMT)*, proposed by Johnson-Laird and Byrne (1991), suggests that people draw conclusions by constructing and inspecting a spatial array that represents the state of affairs described in the premises. This reasoning process consists of three distinct stages: comprehension, description, and validation (see next section for explanation). According to the MMT, linguistic processes are only relevant to transfer the information from the premises into a spatial array and back again, but the reasoning process itself relies only on non-linguistic processes.

A limitation of the MMT so far is that this theory does not explain the difficulty humans have with complex relations. The MMT explains the complexity of reasoning problems, but neglects the construction complexity of the models. This is, we believe, where the *relational complexity theory (RCT)* introduced by Halford (1993) comes into play. Different models representing a spatial description can be measured in terms of cognitive economicity (Halford, 1993; Goodwin & Johnson-Laird, 2005). But a weakness of the RCT is that it does not explain how the reasoning process itself works. In other words both theories, RCT and MMT are limited to some extent, or—more positively—can complement each other.

This paper suggests an integration of RCT and MMT by providing a formal framework which is based on a particular specification of the MMT, namely the theory of *preferred* *mental model (PMMT*, Knauff, Rauh, & Schlieder, 1995; Ragni, Knauff, & Nebel, 2005; Rauh, et al., 2005).

In the next section we will describe the MMT and RCT in order to provide a base for a formal definition of RCT in the spatial domain. We then report empirical findings in support of our hypotheses about relational complexity and the strategies used in construction of preferred mental models. In the last section we present a unifying approach which is able to account for these results in a theoretical framework.

Theoretical Approaches

According to the MMT (Johnson-Laird & Byrne, 1991) a spatial reasoning process can be divided into three distinct phases. In the following we adopt the notation of the phases from Knauff, Rauh, Schlieder, & Strube (1998): In the *construction phase*, reasoners construct a mental model that reflects the information from the premises. If new information is encountered during the reading of the premises it is immediately used in the construction of the model. During the *inspection phase*, this model is inspected to find new information that is not explicitly given in the premises. Finally, in the *variation phase* alternative models are constructed from the premises that refute this putative conclusion. However, some questions remain open, for example how is an initial model constructed, and what strategies are used in construction? How can it be explained that reasoners ignore some models and are not able to find counter-examples?

Our preferred mental model theory (PMMT) is an account based on the mental model theory and able to explain such findings (Knauff, et al., 1995; Ragni, et al., 2005; Rauh, et al., 2005). The term PMM refers to a phenomenon encountered during reasoning with multiple-model problems. In problems in which more than one model is consistent with the premises (so called indeterminate problems) reasoners often construct only one single model – the PMM. This model is easier to construct and to maintain in working memory compared with all other possible models (Knauff et al., 1998). In the model variation phase this PMM is *varied* to find alternative interpretations of the premises (e.g. Rauh et al., 2005). But *how* is a PMM constructed, and what strategies are used? We developed a computational model—*spatial reasoning by models (SRM)—that c*onsists of a spatial array and a focus (spatial working memory) and uses the PMM to explain empirical findings from human spatial reasoning (Ragni, et al., 2005).

One of the main results of the computational theory is the distinction of two insertion principles. Let us consider the following example:

B is to the right of A
C is to the right of A

Here the focus can insert in a spatial array the first object A, then move to the right of A and insert the object B, move back to A and then moves to the right to insert object C, it finds the cell occupied (by object B) so it moves to the right of B inserts object C, and makes an annotation on C to indicate indeterminacy. This gives us the first model, and this model is constructed according to a hypothetical principle we call *first free fit (fff).* This means that an object is inserted at the first free position. Alternatively, object C could be placed into the cell on the right of A. If this cell is occupied by another object (object B), this is then shifted to the next cell. This we call the *first fit (ff)* principle, and it gives us the second possible model. In other words the ff principle always inserts the object at the next position that fulfills the spatial relation specified in the premise. This sometimes means that other objects already in the model must be moved. In the following we report two experiments with human participants in which we tested the *first fit* and the *free first fit* principles against each other and then present a detailed theory of Relational Complexity and formalize it for spatial reasoning.

Empirical Data

In this section we report two experiments with humans to examine (i) which of the possible principles (*ff* versus *fff*) is more likely to be used, and (ii) how the level of complexity affects the PMM, and the corresponding accuracy and reaction times during verification.

We assume that participants construct models according to the *fff*-principal. Furthermore, we assume that the higher the complexity is the more difficult it is to validate the conclusion. This may result in longer latencies and more errors due to the higher number of operations that are required. In addition, we assume that the participant has only the hypothesized PPM in mind. If this is the case, then they should generate the PMM more often than the ¬PMM (Experiment 1), furthermore, relations (conclusions) that only hold in a ¬PMM should be rejected more often than relations that only hold in the PMM (Experiment 2). Note, that from a logical point of view a logically valid conclusion is only given if the relations hold in all possible models. The conclusion is logically invalid if it only holds in one but not in all possible models.

Experiment 1 – The fff Principle

In this experiment we investigate if participants adopt an *fff*-strategy when constructing a PMM.

Participants, Materials, Procedure, and Design. Twenty participants from the University of Princeton were shown eight different indeterminate problems with four premises. In order to avoid artefacts for a certain problem form we also used the horizontal mirror of each two-dimensional problem. The five terms (indicated in Table 1) were randomly replaced with the name of a fruit (lemon, orange, kiwi, peach, mango, and apple). We also included two determinate problems (one for each dimension) to avoid any preference for indeterminate problems.

Table 1 shows the different problems and the possible models for each. For the two-dimensional problems exist two possible models and for the one-dimensional problems three or five possible models. The PMM (assuming that the *fff*-principal is the one participants use) is always the first model in the table and is written in bold letters.

The premises were presented to the participants on a computer screen. Each premise was presented sequentially (in a self-paced manner), and remained on the screen until the presentation of the fourth (and final) premise. After the final button press all the premises were removed and the participants were then asked to draw the model on a sheet of paper. They were free to draw more than one model if they noticed that this was a possibility. However, they were neither instructed to draw more than one model, nor told that in some problems more than one model was possible.

Table 1: Premises and possible models for each problem: In Experiment 1 all depicted problems were used, in Experiment 2 just problems (a) and (b). The problems (a) and (b) were indeterminate 2-dimensional, (c) and (d) were 1-dimension, (e) and (f) were determinate 1- and 2-dimensional. The models indicated in bold letters are the hypothesized PMM for each problem. The mirrored 2-dimensional problems are not shown.

Problem		***Possible models***
(a) A is to the left of B. C is to the right of A. D is behind C. E is behind A.	**(1)**	**E D** **A B C**
	(2)	E D A C B
(b) A is to the left of B. C is to the right of A. D is behind C. E is behind A.	**(1)**	**E** **A B C** **D**
	(2)	E A C B D
(c) B is to the right of A. C is to the right of B. D is to the right of B. E is to the right of C.	**(1)**	**A B C D E**
	(2)	A B D C E
	(3)	A B C E D
(d) B is to the right of A. C is to the right of B. D is to the right of B. E is to the right of B.	**(1)**	**A B C D E**
	(2)	A B C E D
	(3)	A B D E C
	(4)	A B D C E
	(5)	A B E D C
(e) A is to the left of B. C is to the right of B. D is behind C. E is behind A.	(1)	E D A B C
(f) B is to the right of A. C is to the right of B. D is to the right of C. E is to the right of D.	(1)	A B C D E

Results and Discussion. Only two participants (10%) produced more than one model. Each separate problem (problems: U or Z shape, inverse shape, 2- and 1-dimensional, 2, 3, or 5 possible models, indeterminate or determinate) was drawn correct and above chance (Binomial test: $\geq$ 75% correct, $p \leq 0.002$). Altogether, 83% of the drawn problems were correct ($p \leq 0.001$). In addition, we found that 78% of all drawings were of the PMM, while only 22% of all drawings represented a *¬PMM* (Binomial test: 78%, $p = \leq 0.001$).

As we hypothesized participants showed a preference for PMM. In the majority of cases participants' inserted new terms in the manner described by the fff-principle. Even though there are more than two models possible, participants mostly drew only one model, and in the majority of cases this was the PMM.

Experiment 2 – RC and PMM

In this experiment we test if relations that only hold in a ¬PMM are rejected more often than relations that only hold in the PMM. We also examine the effect of relational complexity on the variation process.

Participants, Materials, Procedure, and Design. Twenty-one participants from the University of Princeton were shown 18 two-dimensional indeterminate problems, each with two possible models (see Table 1). The procedure for the presentation of the premises followed the same format as Experiment 1. After deletion of the premises a set of relations was presented on the screen and the participants were asked if these relations were valid. We varied the complexity and the validity of the relations.

The complexity of the relations was counterbalanced across the problems: one third contained binary relations: "Is C near to B and B near to C?" the second third ternary: "Is C as near to B as B is near to D?", and the last third quaternary relations: "Is A as near to D as C is near to E?" For the purpose of this experiment, the relation "near to" means if the term is in direct contact with the other term, regardless of the dimension (horizontal, vertical, diagonal). If one imagines a grid with the term in the center then this means the term that is "near to" it can be in one of eight possible positions.

We use three different types of relations that differed with respect to validity (see Table 2). One of the offered relations was the logically correct solution, because it holds in all models. The other two offered relations could either hold only in the PMM or only in the ¬PMM, but not in both.

For every complexity level there were two problems with the same premises but with relations that were only valid in the PMM (YN) or the ¬PMM (NY).

We were interested in the answers for the offered logically correct relations (conclusions in the strong logical sense, YY) and in the proportions of acceptance and rejection of relations that were only correct in one of the possible models (PMM or ¬PMM). If the participants have only the PMM in mind then they should answer "yes" to the relation which is correct only in the preferred model (YN) and "no" for the one which correct only in the non-preferred model (NY). We expected the opposite for participants who have only the ¬PMM (NY) in mind. Premise reading times, the accuracy and reaction time during the verification of the offered models were recorded.

Table 2: Example for the validity of the offered relations with a ternary complexity (holds for both models: YY, only hold for PMM: YN, or only holds for ¬PMM: NY).

Possible models	*Offered relations*	*Valid for PMM/ ¬PMM*
PMM **E D** **A B C** ¬PMM E D A C B	Is C near to B as B near to D?	Yes/Yes (YY)
	Is E near to B as B near to C?	Yes/No (YN)
	Is E near to C as C near to B?	No/Yes (NY)

Results and Discussion. As expected we found higher reading times for the second premise due to the introduction of the indeterminacy (t test: premise 2 > premise 1, $t(303) = 4.45$, $p \leq 0.001$ and premise 2 > premise 3, $t(303) = 5.40$, $p \leq 0.001$, premise 1 vs. premise 3, not significant). This is in line with previous results in the literature (cf. Carreiras & Santamaria, 1997).

Overall we found that 69% of all responses were correct (Binomial Test: 69% correct answers, $p \leq 0.001$), which is more than one would expect by chance.

As the level of complexity increased so did the errors (ternary > binary: $Z = -3.0$, $p = 0.003$, quaternary > binary: $Z = -2.413$, $p = 0.016$), as did the corresponding reaction times ([ternary + quaternary] > binary: $t(20) = 2.45$, $p = 0.024$).

The most important finding, however, is that it was more difficult to accept relations that only hold in the ¬PMM as compared to those that hold in the PMM (Binomial Test: 46% acceptance, $p = 0.371$; Binomial Test: 79% rejection, $p \leq 0.001$). This pattern of acceptance/rejection rates was evident across all three of the complexity levels (binary: 74%, $p = 0.003$; ternary: 79%, $p = \leq 0.001$; quaternary: 86%, $p = \leq 0.001$).

In the final analysis we are only interested in the answer pairs in which the participants gave a YN or NY answer. These answers are logically speaking invalid, however, they provide insight into which of the two models (PMM or ¬PMM) the reasoner had in mind (see Table 2). Here we found 77% of the responses were consistant with the PMM and only 23% of the responses with consistant with the ¬PMM (Binomial Test $p \leq 0.001$). The reaction times for the preferred and non-preferred relations do not differ (PMM: mean 14427 ms, SD 10276; ¬PMM: M 14193 ms, SD 9313; t test: $t(56) = 0.135$, $p = 0.893$).

Furthermore, we found significant differences between the answer combinations in the binary and the quaternary poblems (Binomial Test binary: 81% YN, $p = 0.007$; quaternary: 76% YN, $p = 0.049$) and a trend toward significance in the ternary problems (Binomial Test: 74% YN, $p = 0.064$).

Several outcomes resulted from this experiment. First, the complexity led to higher processing efforts with the ternary and quaternary relations in comparison to the binary relations. Second, it seems to be easier to reject an offered relation which holds in the ¬PMM than to reject one that holds in the PMM. This was the case for each complexity level. Third, in order to verify the consistency of the models that the participants had in mind only the answer combinations YN and NY are relevant. The comparison between these combinations showed a clear preference for the PMM.

A Unifying Approach

We claim that complexity in spatial reasoning stems from two different sources, namely the complexity to construct or to investigate a model (*model complexity*), and the complexity to deduce from a given set of premises a conclusion (*deduction complexity*). In the following we show that RCT provides an explanation for model complexity since it links the *fff* principle to the principle of economicity. For example consider the following problem:

A is to the left of B.
C is to the right of A.

Obviously, there exist two possible models satisfying these premises, namely

A B C (PMM)   A C B (¬PMM)

In terms of the PMMT, the second model is the result of the ff strategy, while the fff strategy yields the first model. The fact that the first model is the preferred one (see empirical findings in the previous section) can be explained by the RCT, since the first model is computationally cheaper than the second one, that means, the generation process of the PMM A B C must be computationally cheaper than the construction of the ¬PMM A C B. This is because for the *fff*-strategy only the binary relation "C is to the right of B" has to be processed, whereas for the *ff*-strategy a ternary relation like "C is in-between A and B" or "C is to the right of A and to the left of B" is needed. In other words, during the construction of the PMM, the indeterminate premise "C is to the right of A", can be replaced by the binary relation "C is to the right of B", while for the other model the premise have to be replaced by a ternary or two binary relations. In this sense the RCT theory predicts the *fff*-

strategy, which is empirically confirmed in the first experiment.

The RCT theory is not only sufficient to explain the model construction phase as conceived in the MMT, but also contributes to the explanation of the complexity in the model inspection phase. According to the RCT, complexity in this phase arises from the arity and segmentation of the considered relations. If we consider the RCT approach in terms of the SRM framework (Ragni, et al., 2005): the model is constructed as a spatial array, in which the objects are inserted. The focus is the central device for manipulating objects in the array. These focus operations (scanning a cell, inserting objects into a cell, moving objects in the array, etc.) all have the same cost.

The relational complexity is reflected through the different problems the focus has to perform, for example the binary relation "C is near B?" the focus has to test the conclusion by scanning the adjacent cells around B (or C). Ternary relations such as "Is C as near to B as B is near to D" are more difficult since in this case the operations the focus has to perform consists of three sub-processes: the focus has to figure out the distance between C to B as well as the distance between B and D, and finally has to compare these distances. Our empirical findings are well reflected in this computational model. Moreover, through this computational model we are also able to explain how complex relations are decomposed into binary relations like "right", "left", "front", "behind", and how the focus builds the array.

General Discussion

In this paper we propose that by combining MMT and RCT we can account for phenomena found in the spatial reasoning literature, empirically and computationally. Our aim was to investigate the influence of RC on the MMT. We demonstrated how the construction principles for the PMM can be explained by a combination of RCT and MMT, and how this influences the model inspection phase. In addition these findings fit into our formal framework (the SRM).

There are a number of approaches analyzing relational complexity in relation to capacity limitations. Halford et al., (1994; 1998) devised a method for representing relational structure within a connectionist framework, although we have adopted the same principles we have presented a more symbolic implementation. An implication of this formalization was the possibility to distinguish between relations by the number of models they imply: determinate relations such as "to the right of" which leads to one model, compared with a relation such as "next to" which implies multiple models. Using this approach we can interpret and simulate findings in the literature on how people process indeterminate relations and preference of relations (cf. Jahn, Knauff, & Johnson-Laird, 2005).

The main aim of the experiments was to identify the strategies humans use when reasoning about spatial relations. In Experiment 1 we corroborated the fff-principle, for both one and two dimensional problems. Participants tended to draw the PMM based on this principle.

Experiment 2 revealed a number of results: First, for the model inspection phase, our findings suggest that ternary and quaternary relations are more difficult: both resulted in more errors and longer latencies. This result extends the findings of Goodwin and Johnson-Laird, (2005) who found people had difficulty in constructing a mental model when the premises consisted of more complex relations. Their findings can also be modeled in the SRM framework. Second, if the participant has the PMM in mind then it is easier to reject a putative conclusion which was only valid for the ¬PMM, than it is to accept the conclusion which was only valid for the PMM. This is a strong support for the PMMT and its computational model. The SRM framework, a computational model based on the PMMT has been presented, analyzed, and compared to other computational models in (Ragni et al., 2005). An implementation of the SRM in ACT-R had been presented in (Boeddinghaus, Ragni, Knauff, & Nebel, 2006).

Acknowledgments

This research was supported by grants to MK from the DFG (German National Research Foundation) in the Transregional Collaborative Research Center, SFB/TR 8 project. The authors would like to thank Stefan Wölfl, Bernhard Nebel, Geoff Goodwin and Phil Johnson-Laird, and three anonymous reviewers for comments and suggestions.

References

Boeddinghaus, J., Ragni, M., Knauff, M., & Nebel, B. (2006). Simulating spatial reasoning using ACT-R. In *Proceedings of the Seventh International Conference on Cognitive Modeling* (pp. 62-67). Trieste, Italy.

Carreiras, M., & Santamaría, C. (1997). Reasoning about relations: Spatial and nonspatial problems. *Thinking and Reasoning, 3,* 191-208.

Halford, G. S. (1993). *Children's understanding: The development of mental models.* Hillsdale, NJ: Erlbaum.

Halford, G. S., Wilson, W. H., Guo, J., Gayler, R. W., Wiles, J., & Stewart, J. E. M. (1994). *Connectionist implications for processing capacity limitations in analogies.* In K. J. Holyoak & J. Barnden (Eds.), Advances in Connectionist and Neural Network theory, chap. 7, pp. 363--415. Norwood, NJ: Ablex.

Halford, G. S., Wilson, W., & Phillips, W. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental and cognitive psychology. *Behavioral Brain Sciences*, *21*, 803-831.

Goodwin, G. P., & Johnson-Laird, P. N. (2005). Reasoning about relations. *Psychological Review*, *112,* 468-493.

Jahn, G., Knauff, M., & Johnson-Laird, P. N. (2005). Preferred Mental Models in Spatial Reasoning. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.) *Proceedings of the 27th Annual Cognitive Science Conference.* Mahwah, NJ: Erlbaum.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Hove, UK: Erlbaum.

Knauff, M., Rauh, R., & Schlieder, C. (1995). Preferred mental models in qualitative spatial reasoning: A cognitive assessment of Allen's calculus. In *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* (pp. 200-205). Mahwah, NJ: Erlbaum.

Knauff, M., Rauh, R., Schlieder, C., & Strube, G. (1998). Continuity effect and figural bias in spatial relational inference. In *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 573-578). Mahwah, NJ: Erlbaum.

Ragni, M., Knauff, M., & Nebel, B. (2005). A Computational Model for Spatial Reasoning with Mental Models. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.) *Proceedings of the 27th Annual Cognitive Science Conference.* Mahwah, NJ: Erlbaum.

Rauh, R., Hagen, C., Knauff, M., Kuß, T., Schlieder, C., & Strube, G. (2005). Preferred and alternative mental models in spatial reasoning. *Spatial Cognition and Computation, 5*, 239-69.