Deductive Spatial Reasoning: A Computational and a Cognitive Perspective

Marco Ragni, Markus Knauff

In recent years a number of different cognitive theories that explain human reasoning processes have been developed. This paper focusses on two broadly discussed theories: the theory of mental logic and the theory of mental models. Both theories are presented and analyzed in formal terms and compared with respect to their explanatory and predictive power. Our claim is that only a precise computational model can classify reasoning problems and can help to formulate new predictions.

1 Introduction

It is remarkable how reasoning about relations is important in everyday life. A simple example is that if we know that Ann is taller than Beth and Beth is taller than Cath, we can easily infer that Ann must be taller than Cath. Formally, this kind of reasoning can be described by rules of transitivity. But what makes spatial reasoning cognitively difficult? To answer this question we need to know how humans represent and reason with such information. Some necessary notions are now introduced:

(I) The red car is to the left of the yellow car.
The yellow car is to the left of the orange car.
The yellow car is to the left of the green car.
The green car is to the left of the blue car.

Is the blue car (necessarily) to the right of the orange car?

The statements are called *premises*, the cars are the *terms*, and the question refers to a putative *conclusion*. A premise like the first one consists of (two) objects, and a (usually binary) relation like "to the left of". More precisely, the first object (red car) is the "to be localized object"(LO), which is placed according to its relation (left of) to the second object (yellow car[1]), which is the "reference object" (RO).

There are basically two main cognitive theories on how humans solve such problems: syntactic-based theories on the one hand and semantic-based theories on the other. For example, Rips [16] suggested that humans solve these problems by applying formal transitivity rules to the premises, whereas the mental model theory (MMT) proposed by Johnson-Laird and Byrne [7], suggests that people draw conclusions by constructing and inspecting a spatial array that represents the state of affairs described in the premises. Both theories can explain a number of psychological effects. But is it possible to distinguish both theories from a formal perspective? A necessary condition for such an analysis is that both theories are formalized. Although a number of psychological investigations have determined several effects, a formalization has not been proposed yet. The reason might lie in the difficulty to specify precisely the *exact* rules for the rule-based approach and the operations to manipulate mental models for the model-based approach, respectively. However, without having a formalized theory or computational model it is hard to test predictions or even try to falsify an approach. Of course, like in physics, a theory about natural phenomena can only be refuted based on experiments. But if the theory itself is underspecified, it might be possible that different experimenteers adapt the theory to suit the data in a contradictory way.

In the following methods from Artificial Intelligence are used to formalize the cognitive approaches and to develop predictions of the theory that can be tested with human subjects. We start with an analysis of the theory of mental logic and the theory of mental models for spatial reasoning. The latter theory is precised and integrated in a computational model for spatial reasoning based on Baddeley's working memory model. A discussion of some remaining questions concludes the paper.

2 Two Cognitive Theories

This section sheds some light on the two main cognitive theories about human deduction: the syntactic theory of mental logic and the semantic theory of mental models. Both theories are independent from any data-structure, i.e., they do not make any assumptions about how information is represented. Both theories are therefore limited in their explanatory power on how and why information necessary for reasoning gets lost.

2.1 Theory of Mental Models

According to the MMT, linguistic processes are relevant to transfer the information from the premises into a spatial array and back again, but the reasoning process itself completely relies on the model manipulation only ([6], pp. 434):

> Reasoners use the meanings of assertions together with general knowledge to construct mental models of the possibilities compatible with the premises.

A *mental model* is an internal representation of objects and relations in spatial working memory, which matches the state of affairs given in the premises. The semantic theory of mental models is based on the mathematical definition

[1] The objects are abbreviated by R, Y, O, G and B

Auszug aus: Künstliche Intelligenz, Heft 1/2008, ISSN 0933-1875, BöttcherIT Verlag, Bremen, **www.kuenstliche-intelligenz.de/order**

of deduction, i.e., a propositional statement φ is a consequence of a set of premises $\mathcal{P}$, written $\mathcal{P} \models \varphi$, if in each model $\mathcal{A}$ of $\mathcal{P}$, the conclusion φ is true. In other words, the conclusion is valid if there is no counter-example, i.e., a model where $\mathcal{P}$ holds but not φ. According to the mental model theory (MMT) the human reasoning process consists of three distinct phases: The *model generation phase*, in which a first model is constructed out of the premises, an *inspection phase*, in which the model is inspected to check if a putative conclusion is consistent with the current model, and, finally, the *validation phase* where alternative models are generated from the premises that may refute the putative conclusion. Since the exact relation between the orange, the blue, and the green car is not specified multiple-models are consistent:

R Y O G B   R Y G O B   R Y G B O

The numbers of models is one effect responsible for human difficulties in reasoning [6]. A limitation of the MMT is that it is not able to explain a phenomenon encountered in multiple-model cases, namely that humans generally tend to construct a *preferred mental model* (PMM). This model is easier to construct, less complex, and easier to maintain in working memory than alternative models [8]. The determining factor in explaining human preferences is the principle of economicity [9]. In the model variation phase this PMM is varied to find alternative models [15]. This theory, however, has not been formalized yet and is therefore not fully specified in terms of operations necessary to process such problems as were described above. So it has been handled in a rather implicit and vague way.

2.2 Theory of Mental Logic

There are a number of different calculi for reasoning with spatial problems. Rips characterised the central idea of this approach in the following way ([16], p. 40):

> Reasoning consists in the application of mental inference rules to the premises and conclusion of an argument. The sequence of applied rules forms a mental proof or derivation of the conclusion from the premises, where these implicit proofs are analogous to the explicit proofs of elementary logic.

Van der Henst [19] e.g. proposes the set of rules of Figure 1, which are successively applied to the premises of a problem

1. $Left(x,y)\ \&\ Front(z,x)\ \rightarrow\ Left(z,y)$
2. $Left(x,y)\ \&\ Front(z,y)\ \rightarrow\ Left(x,z)$
3. $Left(x,y)\ \&\ Left(y,z)\ \rightarrow\ Left(x,z)$
4. $Left(x,y)\ \leftrightarrow\ Right(y,x)$
5. $(Left(y,x)\ \&\ Left(z,x))\ \rightarrow (Left(y,z)\ or\ Left(z,y))$
6. $(Left(y,z)\ or\ Left(z,y))\ \&\ Front(w,z)\ \rightarrow$ $(Left(y,w)\ or\ Left(w,y))$

Figure 1: Set of spatial inference rules from [19]

description. Assume we want to derive for problem (I) that the blue car is necessarily to the right of the red car. It is sufficient to apply the third and the fourth rule of Figure 1. That is, by

$$Left(R,Y)\ \&\ Left(Y,G)\ \rightarrow\ Left(R,G)$$

we derive that the red car is to the left of the green, by

$$Left(R,G)\ \&\ Left(G,B)\ \rightarrow\ Left(R,B)$$

we derive that the red car is to the left of the blue, and by

$$Left(R,B)\ \leftrightarrow\ Right(B,R)$$

that the blue car is to the right of the red car. In other words, a minimal solution needs three inference steps. Throughout the reasoning process a high number of relations has to be stored, namely all premises and all inferred relations. This can result in a high load on the working memory. It is remarkable that through the application of the rules in Figure 1 on a two-dimensional problem more than one base relation, namely right, left, front or behind, can hold for a tuple, i.e., it is possible that the same tuple holds that A is left of B and A is in front of B. So, by the application of the rules some kind of fuzziness appears.

All rule-based theories claim that the difficulty mainly depends on two things: On the number and kind of inference rules applied to derive a conclusion. If we assume that all the rules have the same difficulty, then the main difficulty depends on the number of rules being applied. Hence, the more inference steps are necessary, the more difficult is a conclusion [7, 19].

This idea implicitly assumes that humans use some kind of search procedure involving optimal search strategies. A search strategy is called optimal if it always finds the minimal solution for a problem. For instance breadth-search is an optimal search strategy where depth-first is not [17]. So the search strategy is an important factor in determining the complexity of a problem, but there is no data indicating which kind of search procedure humans apply. This point can be further specified: Assume we have a set of premise $\mathcal{P}$ and two conclusions φ_1 and φ_2, where the steps to infer φ_1 from $\mathcal{P}$ takes 3 steps and to infer φ_2 from $\mathcal{P}$ takes 4 steps. If it cannot be generally assumed that reasoners necessarily choose the minimal derivation steps to derive a putative conclusion, then $(\mathcal{P}, \varphi_1)$ is not always easier than $(\mathcal{P}, \varphi_2)$. Otherwise the reasoner would always have to know which inference has to be drawn next–a property normally assumed with nondeterministic rather than deterministic systems. From a formal perspective both theories–the theory of mental logic and of mental models–do have the same explanatory power.

2.3 An Analysis

A first experiment of analyzing differences between the mental logic approach and the mental model approach investigated determinate and indeterminate problems (cf. Figure 2). If we apply the rules 1-4 in Figure 1 to the determinate and indeterminate problem description in Figure 2, the conclusion can be derived in an equal number of steps. Therefore, from a pure rule-based perspective, both problems have the same difficulty. Nonetheless, empirical studies show that determinate problems are significantly easier to solve than indeterminate problems [7]. Van der Henst [19] argues for extending the classical set of rules by the rules 5 and 6. But these rules are in some sense artificial and not necessary to derive the conclusion. Thus it seems to be a post-hoc alteration of the specified set of rules to explain the data [3]. The mental model theory explains this phenomenon by regarding the model relation. To check if the putative conclusion

Auszug aus: Künstliche Intelligenz, Heft 1/2008, ISSN 0933-1875, BöttcherIT Verlag, Bremen, **www.kuenstliche-intelligenz.de/order**

(II) A is to the right of B.
C is to the left of B.
D is in front of C.
E is in front of A.

(III) A is to the right of B.
C is to the left of A.
D is in front of C.
E is in front of A.

Which relation holds between D and E?

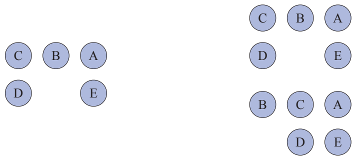


Figure 2: A determinate (II) and indeterminate (III) problem [7]

"E is to the left of D" can be derived, all consistent models have to be checked. Since in the one-model case there is only one model consistent, the effort to check the conclusion is smaller than checking two models.

Another argument in favor of the mental model theory is that recent experiments have shown that humans generally tend to construct a PMM in multiple model cases [8] in reasoning with Allen's Calculus and that they vary this model according to the principle of local transformation [15]. Such preference strategies are not limited to one calculus alone, they can also hold for other calculi [12]. For instance, reasoners use for solid objects the fff-Principle, i.e., if they have to insert an object C to the left of B in a model A B, they put this object to the left of A. A general principle to explain such effects is the number of operations necessary.

In another experiment [11] participants received indeterminate premises (III) and questions of different relational complexity [4] of the form: Is D as near to C as E is to A? The participants had to decide whether the query was consistent or not. The results show that binary relations are easier to process than ternary and quartenary relations. But what is more remarkable is that participants received a premise description with left, right, and front, but were able to solve a question with the relation *near*. In other words, participants received a relational description that was different to the relations they were asked for. It is not only remarkable that they had to translate one kind of relation into another, but also that they were actually able to give an answer that required positional information. Typical transitivity rules maintain the relation of the premises, i.e., the relation of the condition is used in the consequence (cf. Figure 1). Instead of using classical rules other logical systems can be used to describe "nearness". It is, however, much simpler to explain such phenomena with a model-based approach.

Finally, the high number of rules that have to be specified for each relation and the number of relations that have to be stored in the working memory is remarkable. It is possible to prove that a relational system consisting of two natural relations like *adjacent left* and *distant left* on discrete structures is inherently incomplete, i.e., there is an infinite number of rules necessary to prove all valid conclusions [10]. This is different in reasoning with models, which are *integrated representations* of the relational information, i.e., the relations have not to be stored but are implicitly represented by the position of objects. Computational models are needed to specify operations and the memory used to solve tasks in order to analyze these theories.

3 The Computational Model

Human reasoning strongly depends on working memory. Baddeley's working memory model (WMM) [1], assumes a central executive (CE), which is responsible for monitoring and coordinating the operations of two subsystems: the *phonological loop* (PL) and the *visuo-spatial sketchpad* (VSSP). The first subsystem, the PL, stores information in a language-based form. The second subsystem, the VSSP, is indepedent from the first in terms of memory limits, stores visual and spatial information. Both subsystems are controlled by a CE which is able to store and manipulate information. Central questions for a combination of the preferred mental model theory and WMM are: In which subsystem and how does the reasoning take place? What are limits of the subsystems and the control process? Since the deduction process in relational spatial reasoning uses mental models [2] these can be located in the WMM in the VSSP, where the construction and manipulation of the mental models by a special device, which is called focus, takes place.

Several cognitive models have been developed to model diverse aspects of deductive relational reasoning. One of the best developed systems is a model by Schlieder and Berendt for simulating the empirical results found in an experiment which was conducted to analyze the three-term series problems in Allen's Interval Calculus [18]. Although this model can explain how humans construct a PMM by reasoning with intervals, it cannot be applied to reasoning with solid objects. This was the main motivation for developing a new model, able to parse problems consisting of the most parsimonious relations right, left, front and behind, to construct, to inspect and vary the PMM [13].

However, no existing computational model for relational reasoning integrates an algorithmic approach as well as a working memory model. This is the starting point for our analysis. The CROS-Model (Cognitive Relational Operating System) which formalizes the WMM and PMMT consists of: A conceptualization of the WMM (with subsystems), a *manipulation device* for the mental models, a (relational) language describing object positions, and a *semantic interpreter*, interpreting the language.

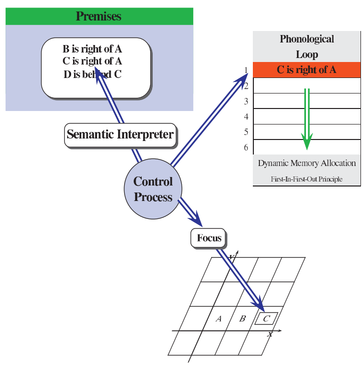


Figure 3: The CROS-Model

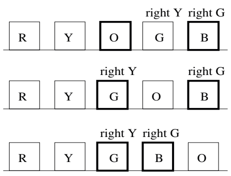
The VSSP is a spatial array (SA) of two-dimensional grids, called *layer*, in which mentals models are generated and manipulated by a device called *focus*. The focus can perform a small number of operations like moving, reading, and inserting. E.g. for 'A left B' and 'C right D', there are two possible submodels, each placed in its own layer, so that submodel AB would be in the first and CD in the second layer. Formally, the CROS is a 6-tuple (I, SI, A, F, PL, C) with I the input device, SI the semantic interpreter, A a spatial array, that contains the layers (submodels), F the focus, manipulating the spatial array, PL, a memory for storing verbal information, and C a control process.

When processing natural language strings, the meaning of the input has to be interpreted. In linguistics, as well as in psychology, the existence of a *semantic interpreter* (SI) is assumed, which maps syntactically analyzed texts to the formal representation in our model. A discussion of the SI, however, would go beyond the scope of the paper. Problems related to the ambiguity of spatial relations are not accounted for. The CROS interprets the string "A is left of B" as: both objects are in the same line and A is to the left of B. The relations "right", "front", and "behind" are equivalently defined. If indeterminacy occurs, information about alternative models must be stored. Since a mental model is only a representation, such information must be held in another subsystem. The appropriate memory system in the WMM for this kind of propositional information is the PL, which is consistent with neuropsychological evidence [8]. The PL uses a dynamic memory allocation system, which allows the modeling of activated objects.

Since both systems, the SA and the PL, are only memory systems and the focus manipulates only the SA, a *control process*, which manages the CROS, is needed, that manages the subsystems and controls the focus operations on the SA. The control process has a limited instruction set. Several instructions directly control read/insert/move operations of the focus, statements to branch or loop the control flow, and simple test instructions [14]. With this set of instructions, algorithms for all three deduction phases can be defined and different insertion strategies can be compared. The premises are read and interpreted iterativeley by the SI, and the control process immediately inserts the new encountered information into the model by moving the focus within the SA and adding indeterminacy information to the PL. The focus has the ability to create new layers for premises that cannot be constructed into one layer.

Now four types of premises must be distinguished: (1) the first premise, (2) premises in which one object from the preceding premise appears and a new object which must be inserted in the array, (3) the type of premises, where no objects of the previous premises appear, e.g., D r C, A r B, B r C. And (4) premises in which two formally separate models are connected.

The functionality of the CROS is now demonstrated by processing problem I. The construction process starts with the first premise and an empty layer. First the RO is placed, then the focus moves in the direction of the relation and places the LO to the next free cell. In our example Y is inserted first, the focus moves to the left and inserts R. The algorithm checks to which type each new premise belongs and then inserts the object(s) according to the specific case. For premises of type 2 only one object has to be inserted and if it cannot be placed as a direct neighbor, the model structure is indeterminate, and therefore, the control process annotates the object by inserting the relational information as a proposition into the PL, and the focus places the present object according to the fff-principle. For premises of type 3, where neither of the two objects are contained in the model, a new layer is generated, both objects will be placed in this new layer and are treated in the same way as a premise of type 1. If both objects are contained in different layers (type 4), both layers have to be merged according to the relation of the premise. If a counter-example exists, it is a model containing the additional knowledge. The second and third premises are of type 2 because Y is already in the model, so O and G are inserted to the right of Y according to the fff-principle. Because G cannot be placed adjacent to Y, it is annotated with 'right Y'. The next premise, which is to be processed, is also of type 2 and object B, that is not in the model, is inserted directly to the right of G. But because G is annotated, B has to be annotated too. Now the construction phase is complete and the resulting model is shown in the first line of Figure 4. Now that the construction of the model is finished, the *inspection phase* searches for new information that was not directly specified in the premises. The focus moves to the first given object (RO) and from there it inspects the model according to the relation in order to find the second object (LO). The *model variation* comes into play if a conclusion must be verified or if additional knowledge of two already contained objects has to be processed during the model construction process. The focus starts in the variation process with the PMM and varies the model by local transformations to generate a counter-example to the putative conclusion. The variation process starts from the generated PMM (in which the putative conclusion holds). The algorithm checks if one of the objects in the conclusion is annotated. Annotations on objects specify the positional relation to the reference objects, which we refer to as *anchor*. If an annotation include both objects of the putative conclusion then the putative conclusion holds. The same argument holds if none of the conclusions' objects are annotated because then the positions of the objects are fixed. If there is an annotation on one object (and not to the other), as in the example conclusion 'B is to the right of O' (see Figure 4), the only object of the conclusion which has to be moved is B and not O. This



The first line contains the generated PMM. The annotated objects can be varied, where the bold marked objects are to be varied to check the conclusion.

Figure 4: The variation process of the CROS for problem (I)

comes along with the use of annotations, i.e., an annotation is only created for indeterminate object positions. If the object which is to be moved has an anchor, it may be necessary to move the anchor first. A simple example can illustrate this process: B cannot be moved because G, the anchor of B, is a direct neighbor of B. Thus, the algorithm first exchanges the anchor to the left of O. Now the counter-example is generated by exchanging B beyond O and so B is left of O, so false is returned. If both objects are annotated, then first the LO of the putative conclusion is exchanged. The LO is moved into the direction of the RO until its anchor is reached. If this results in the generation of an inconsistent model, the algorithm stops and returns false. It is possible that the anchor object is in-between the LO and the RO, and that results in

Auszug aus: Künstliche Intelligenz, Heft 1/2008, ISSN 0933-1875, BöttcherIT Verlag, Bremen, **www.kuenstliche-intelligenz.de/order**

the exchange of the LO until it reaches its anchor. Then the anchor object is recursively exchanged towards the RO. If no further exchanges to the RO are possible, the exchange process starts to exchange the RO into the direction of the LO.

4 Conclusion

A common proverb says that chess is the drosophila of Artificial Intelligence, referring to the fruit-fly who is the favorite test bed for genetic theories. Analogically speaking, spatio-temporal reasoning is the drosophila of human relational reasoning: Arbitrary relational premises with transitive relations (nicer, smarter, etc.) imply an ordering relation and can therefore be interpreted by spatial representations. This and its importance in everyday life is the reason why spatio-temporal reasoning has been quite extensively studied in the field of deductive reasoning.

The preferred mental model theory as well as Baddeley's WMM are able to explain several empirical results in spatial reasoning. But both theories have neither been brought together nor been formalized. Since human reasoning is based on mental operations as well as on mental structure, only a cognitive as well as formal model, comprising both aspects, is able to explain intra- and inter-individual differences. This is the motivation for our investigation and formalization. The resulting model, the CROS, is able to cover a wide span of effects on experimental results from the literature [11–14]. An even more important confirmation of the CROS-model is that several CROS-predictions like the insertion principle (fff) and the continuous transformation process have been also empirically validated.

Our approach has something in common with a very early approach in cognitive science: Hunter [5] claimed that reasoning difficulty strongly depends on operations on the form of premises. Even if his approach was limited – he had the right intuition: the number of mental operations is the determining factor in explaining reasoning difficulty.

Acknowledgments
This work was partially supported by the Deutsche Forschungsgemeinschaft (DFG) as part of the SFB/TR 8 Spatial Cognition. We would like to thank Bernhard Nebel, Felix Steffenhagen, Stefan Wölfl and Andreas Bittner.

References

[1] Baddeley, A. D. *Essentials of human memory*. Psychology Press, East Sussex, England, 1999.

[2] Byrne, R. M., and Johnson-Laird, P. N. Spatial reasoning. *Journal of Memory & Language 28*, 5 (1989).

[3] Goodwin, G. P., and Johnson-Laird, P. N. Reasoning about relations. *Psychological Review 112*, 2 (2005).

[4] Halford, G. S., Wilson, W. H., and Phillips, S. Processing capacity defined by relational complexity: implications for comparative, developmental, and cognitive psychology. *Behavioural Brain Science 21* (1998).

[5] Hunter, I. M. L. The solving of three term series problems. *British Journal of Psychology 48* (1957), 286–298.

[6] Johnson-Laird, P. N. Mental models and deduction. *Trends in Cognitive Sciences 5*, 10 (2001).

[7] Johnson-Laird, P. N., and Byrne, R. M. J. *Deduction*. Erlbaum, Hove (UK), 1991.

[8] Knauff, M. Deduktion und logisches Denken. In *Denken und Problemlösen. Enzyklopädie der Psychologie*, J. Funke, Ed. Hogrefe, Göttingen, 2006.

[9] Manktelow, K. *Reasoning and Thinking*. Hove: Psychology Press, 1999.

[10] Ragni, M. An arrangement calculus, its complexity and algorithmic properties. In *KI 2003: Advances in AI* (2003), A. Günter, R. Kruse, and B. Neumann, Eds., Springer LNAI, pp. 580 – 590.

[11] Ragni, M., Fangmeier, T., Webber, L., and Knauff, M. Complexity in spatial reasoning. In *Proc. of the 28th Cognitive Science Conference* (2006), Erlbaum.

[12] Ragni, M., Fangmeier, T., Webber, L., and Knauff, M. Preferred mental models: How and why they are so important in human spatial reasoning. In *Spatial Cognition V.* (2006), C. Freksa, M. Knauff, B. Krieg-Brückner, B. Nebel, and T. Barkowsky, Eds., Springer LNCS.

[13] Ragni, M., Knauff, M., and Nebel, B. A computational model for spatial reasoning with mental models. In *Proc. of the 27th Cognitive Science Conference* (2005), B. Bara, L. Barsalou, and M. Bucciarelli, Eds., Erlbaum.

[14] Ragni, M., and Steffenhagen, F. A cognitive computational model for spatial reasoning. In *AAAI Spring Symposium 2007* (2007).

[15] Rauh, R., Hagen, C., Knauff, M., Kuss, T., Schlieder, C., and Strube, G. Preferred and Alternative Mental Models in Spatial Reasoning. *Spatial Cognition and Computation 5* (2005).

[16] Rips, L. J. *The Psychology of Proof: Deductive Reasoning in Human Thinking*. MIT Press, 1994.

[17] Russell, S., and Norvig, P. *Artificial Intelligence: A Modern Approach*, 2nd edition ed. Prentice-Hall, 2003.

[18] Schlieder, C., and Berendt, B. Mental model construction in spatial reasoning: A comparison of two computational theories. In *Mind modelling: A cognitive science approach to reasoning*, Schmid, Krems, and Wysotzki, Eds. Pabst, 1998, pp. 133–162.

[19] Van der Henst, J. Mental model theory versus the inference rule approach in relational reasoning. *Thinking and Reasoning 8* (2002), 193–205.

Contact
Dipl.-Math. Marco Ragni
Georges-Koehler-Allee, 79102 Freiburg
Phone: +49 (0)761-203 8227
Email: ragni@informatik.uni-freiburg.de

Marco Ragni is a research assistant at the Department of Foundations of Artificial Intelligence at the University of Freiburg. He received his diploma in mathematics with a minor in computer science from the University of Freiburg. His research focuses on formal and cognitive aspects of qualitative reasoning, spatial planning, human deduction, and logic.

Markus Knauff is Professor of Experimental Cognitive Psychology at the University of Giessen. He received his PhD from the University of Freiburg. His main research interests are the cortical, computational, and cognitive correlates of complex thoughts. In this field, he is specifically interested in the connection between formal logic, mind, and brain.