An ACT-R Approach to Reasoning about Spatial Relations with Preferred and Alternative Mental Models

Sven Brüssow (sven@cognition.uni-freiburg.de)[1]
Matthias Frorath (frorath@informatik.uni-freiburg.de)[1]
Marco Ragni (marco.ragni@cognition.uni-freiburg.de)[1]
Thomas Fangmeier (thomas.fangmeier@uniklinik-freiburg.de)[1, 2]
[1]Center for Cognitive Science, University of Freiburg
[2]Department for Psychiatry, University Clinic Freiburg

Abstract

A computational model of spatial relational reasoning implemented in the ACT-R cognitive architecture allows for the simulation of a wide range of behavioral data in the context of both determinate and indeterminate deductive spatial reasoning tasks. In that respect the presented study bridges the gap between results of previous work that investigated determinacy conditions separately. ACT-R's subsymbolic processing principles substantially contribute to the underlying theory of Preferred Mental Models as they add a powerful component making precise accuracy predictions possible. In addition, the data is informative about a possible strategy when the task is to judge if an externally presented spatial description matches a mental model that resulted from the current reasoning process.

Keywords: Spatial reasoning; mental models; ACT-R

Introduction

The objective of psychological theories to explain and predict experimental data is best realized by cognitive models implemented in cognitive architectures like ACT-R (Anderson et al., 2004; Anderson, 2007). Cognitive models help researchers develop intuitions about the cognitive demands of certain tasks, generate additional data that can be tested against human data, and eventually allow for theory revaluation. ACT-R is empirically grounded. Its explanatory power lies in the combination of discrete symbolic descriptions with constantly interfering non-discrete subsymbolic processes. Both concepts are needed for an implementation of a psychologically plausible theoretical account.

In the present work, we investigated the potential of ACT-R to substantiate the theoretical framework of the Preferred Mental Model Theory (PMMT). The PMMT describes the deduction process in the context of ambiguous descriptions and stands in the tradition of the classical Mental Model Theory originally introduced by Johnson-Laird (1980, 1983). Mental models are constructed from given spatial relational information that is typically given by a set of premises. The PMMT suggests distinct construction principles for determinate and indeterminate premises. If premises are determinate they allow for only one model (which we call UNI for *unique*). If premises are indeterminate they allow for multiple model derivations. In this case some mental models derived from these premises may be considered in the deduction process while others are neglected. The mental model that is preferred over alternatives will be referred to as the preferred mental model (PMM), the first alternative model is denoted by AM1, and the second by AM2.

The reasoning process is commonly divided into three distinct phases. First, in the *construction phase* the initial PMM is incrementally built up from the given premises. Second, during the *inspection phase* evidence shows that human reasoners try to use the spatial information encoded in the PMM to validate a putative conclusion (Knauff, Rauh, & Schlieder, 1995; Rauh, Hagen, Schlieder, Strube, & Knauff, 2000; Rauh et al., 2005; Jahn, Knauff, & Johnson-Laird, 2007); in accordance with the PMMT the PMM involves the lowest construction costs (Ragni, Knauff, & Nebel, 2005). Third, in the *variation phase*, a correct inference depends on the type of mental model taken into account; dependence solely upon the initial PMM can lead to counter-examples being missed. The respective inference process is formally described by discrete operations starting with incremental integration of premise terms into the PMM; followed by model-conclusion comparisons; and, if necessary, continued with PMM modifications to an alternative mental model (Ragni & Brüssow, 2011). Alternative models can be generated by applying additional variation processes. Hence, with respect to the indeterminate premises illustrated in Table 1, to validate the conclusion "is D to the left of B?" two transformations are necessary, provided that participants test the conclusion with the possible mental models in the predicted order of PMM, followed by AM1, and then AM2.

Although the model variation phase is important, little research has systematically investigated the underlying processes. Rauh et al. (2000) were the first to investigate errors of omission and commission in spatial relational reasoning with intervals. Their presented theory, however, is purely symbolic; it does not allow for memory effects or other subsymbolic effects. This is why we decided to implement the PMMT in ACT-R.

ACT-R is both a theory and architecture of cognition. It has successfully been used to simulate a wide range of cognitive tasks. Standing in the tradition of production systems, originally introduced by Newell and Simon (1972) as appropriate formalisms for describing human problem solving behavior, procedural knowledge is described by sets of production rules that operate on memory chunks that in turn represent declarative knowledge. Modularly organized, ACT-R provides distinct components each specialized for certain perceptual or cognitive tasks; environmental visual information is processed in the vision-module, internal goal formulation

Table 1: Determinate premises resulting in the model "ABCD" (top) and indeterminate premises resulting in one of the three mental models "ABCD," "ACBD," or "ACDB" (bottom). Model denomination refers to unique model (UNI), preferred mental model (PMM) or alternative mental models (AM1, AM2). $\overset{AB}{\rightarrow}$ represents the premise "A is to the *left* of B" and $\overset{AB}{\leftarrow}$ the premise "B is to the *right* of A." The remaining premises should be read accordingly. Arrows, in addition, indicate the order of term presentation. Double curve arrows above mental model representations mark those terms that need be transposed to transform the source model to the next alternative model.

P1	P2	P3	UNI	PMM	AM1	AM2
$\overset{AB}{\rightarrow}$	$\overset{BC}{\rightarrow}$	$\overset{CD}{\rightarrow}$	ABCD	∅	∅	∅
$\overset{AB}{\leftarrow}$	$\overset{BC}{\rightarrow}$	$\overset{CD}{\rightarrow}$				
$\overset{AB}{\rightarrow}$	$\overset{BC}{\leftarrow}$	$\overset{CD}{\rightarrow}$				
$\overset{AB}{\leftarrow}$	$\overset{BC}{\leftarrow}$	$\overset{CD}{\rightarrow}$				
$\overset{AB}{\rightarrow}$	$\overset{AC}{\rightarrow}$	$\overset{CD}{\rightarrow}$	∅	$A\overset{\leftrightarrow}{BC}D$	$AC\overset{\leftrightarrow}{BD}$	ACDB
$\overset{AB}{\leftarrow}$	$\overset{AC}{\rightarrow}$	$\overset{CD}{\rightarrow}$				
$\overset{AB}{\rightarrow}$	$\overset{AC}{\leftarrow}$	$\overset{CD}{\rightarrow}$				
$\overset{AB}{\leftarrow}$	$\overset{AC}{\leftarrow}$	$\overset{CD}{\rightarrow}$				

is constructed in the goal module, problem state information is processed by the imaginal module, the retrieval module accesses information stored in declarative memory, the manual module performs motor responses, and the procedural module represents the central executive controlling the activities of these modules. Each module is equipped with an interface called a buffer holding single chunks that are thus available across modules. In addition to the capacity restriction of just one chunk per buffer, productions process information in a strictly serial way; only one production may be active at a time. Note that this is not a general restriction of production systems. Modules, however, may be active in parallel; those chunks currently distributed across buffers are processed simultaneously.

A distinguishing feature of ACT-R is that the above described symbolic behavior is constantly directed by subsymbolic probabilistic processes. For example, the availability of chunks to the retrieval buffer depends on their level of activation, a dynamic numeric value that is computed for each chunk. Once a chunk has been created its initial activation decreases according to a fixed decay rate. Chunk activation, however, can also increase dependent upon (i) the number of positive retrievals in the past, (ii) spreading activation from other chunks in the buffers, and (iii) merging of identical chunks. Similarly, production selection depends on the utility value associated with each production; if multiple productions match simultaneously, their utility value controls ambiguity resolution.

Although the recent focus of research in the context of ACT-R as a theory has definitely been on grounding the architecture on a neurological basis, traditional wide range coverage of behavioral data predictions yet remains an inevitable prerequisite. Earlier versions of the presented model were reported previously (Ragni, Fangmeier, & Brüssow, 2010; Ragni & Brüssow, 2011). It was originally developed to test ACT-R's BOLD function (Anderson, 2007) with fMRI data obtained from a study by Fangmeier, Knauff, Ruff, and Sloutsky (2006) and tested with more sophisticated behavioral data obtained from a study by Ragni, Fangmeier, Webber, and Knauff (2007). Fangmeier et al. investigated phases of the reasoning process using only determinate tasks whereas the study of Ragni et al. reported data restricted to indeterminate tasks. Furthermore, to inhibit linguistic processes Fangmeier et al. presented premises and conclusions as abstract terms whereas Ragni et al. presented complete sentences. Hence, it remained desirable to test model predictions for both uniformly presented determinate and indeterminate material obtained from one and the same study.

Accordingly, we present data and predictions based on a combination of both determinate and indeterminate tasks. In addition, after participants completed the reasoning task they were requested to recall their previously created mental models and then to decide whether a certain constellation of presented terms match. The data suggest that participants again first use the PMM rather than an alternative.

Method

Participants. Twenty-eight students (15 female, $M = 22.86$ years, $SD = 3.17$) participated in the study. Four participants were excluded from further analysis because they missed the chance level threshold of 21 correctly solved tasks on conclusion validation determined by a binomial test. Participants gave written consent and either received a small monetary reward or course credit for their participation.

Procedure, Materials, and Design. Participants were seated in front of a computer screen and first completed six training tasks. Using a within-participants design they then completed 48 balanced and randomized tasks chosen from a total of 288 tasks. There were 24 task sets each assigned only once. As is illustrated in Figure 1 trials exposed participants with (1) an initial premise processing, (2) a conclusion validation, and (3) a mental model validation phase.

Including instructions each session lasted approximately 45 to 60 minutes. Response times and answer correctness for both conclusion and mental model validation were logged.The material consisted of 144 determinate and 144 indeterminate tasks each presenting three consecutive premises (P1-P3) followed by one putative conclusion and a putative mental model. Premises and conclusions consisted of two terms. Terms were separated by a centered dot and each term either appeared to the left or to the right of it. The alignment of terms on the screen at the same time encoded the underly-

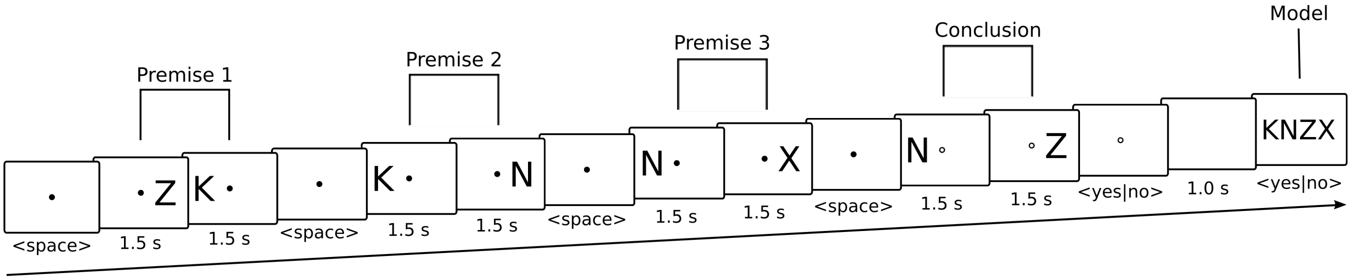


Figure 1: Sample trial. Spatial relations are indicated by their position, i.e. right aligned ‘Z’ in the first premise means ‘Z is to the right of.’ For each trial premise and conclusion terms were presented at least 1500 ms. To see the next pair of terms participants had to press the space bar in the premise processing phase. To complete the conclusion processing phase they had to press either the left or right arrow key representing ‘yes’ and ‘no’ respectively. Premise and conclusion presentation differed in the term separating center point. For the final model validation task participants had to respond using the arrow keys again.

ing spatial relation; if the first term appeared left aligned this indicated a “left-of” relation to the second term that consequently appeared right aligned as soon as the first was cleared from the screen and vice versa. Note that for illustratory purpose we use the terms “A,” “B,” “C,” and “D” whereas in the experiment, to avoid sequence effects or particular abbreviations that can be memorized, “K,” “N,” “X,” and “Z” were used (cf. Figure 1). Similarly, for each structurally equivalent premise across trials different terms were used resulting in 24 permutations of model terms. Table 1 shows premise combinations for determinate and indeterminate tasks. P1 and P2 differed in the presentation order of terms whereas P3 was the same in all trials.

After the premises were displayed a putative conclusion was presented in the same way. Merely the term separator differed indicating that the presented terms now belong to the conclusion. Eventually, after participants responded to the conclusion a configuration of four terms was displayed on the screen representing a putative mental model. In both determinate and indeterminate tasks participants had to decide whether the presented configuration matched one of the mental models they had previously constructed.

Cognitive Model

In general, configurations of spatially related terms are represented as mental model chunks with position slots holding the respective terms. Hence, premises, conclusions, or entire mental models share the same structure. Of the ACT-R internal parameters only activation noise (0.31), retrieval threshold (-0.75), and latency factor (0.4) deviated from the default values but were held constant across runs.

The ACT-R model processes the first premise without encountering any indeterminacy and integrates premise terms directly into an empty mental model chunk. It then makes no further use of the information encoded in the first premise. This is inspired by the findings of Mani and Johnson-Laird (1982) that human reasoners, having processed the first premise, forget the corresponding information; there is no need to remember it because there are no alternative mental models possible that would later require a retrieval of this information to allow for a modification of the initial PMM. For successive premises, however, indeterminacy may occur due to multiple possible positions for term integration. Consequently, the ACT-R model creates extra representations for successive premises that allow for a later retrieval if indeterminacy occurs. The process of flagging a successive premise to make it available for later processing steps is referred to as “annotating a premise”. In particular, the ACT-R model flags the already inserted term after which indeterminacy occurred as the “reference object” term (RO) and the new term as the “to be located object” term (LO); the LO has to be inserted at the first free position rather than at the directly adjacent first fitting position that is already occupied by another term from a previously presented premise. Figure 2 illustrates the process of the construction of the initial PMM including annotation assignment.

If the PMM fails to validate the conclusion, the ACT-R model can now retrieve annotated premises and reuse them to modify the PMM to an alternative model as is illustrated in Figure 3. Annotation usage proceeds by transferring the current mental model chunk to the imaginal buffer and retrieving the previously annotated premise. An alternative term configuration can be obtained by moving the LO towards the RO, i.e. to the first fitting position. In a subsequent step the ACT-R model retrieves the conclusion chunk again and tests it in a second comparison step with the modified mental model. These steps are repeated until the presented conclusion agrees with a mental model causing the motor module to press the key for a positive response, or when it fails to retrieve further annotated premises resulting in a key press representing a negative response.

In the final mental model validation phase the ACT-R model first retrieves the initial UNI/PMM again and transfers it to the imaginal buffer. This transfer is necessary because, like the model construction phase illustrated in Figure 2, intermediate term retrievals repeatedly occupy the retrieval buffer. The reason for the explicit retrieval request for the UNI/PMM is that participants make fewer errors and need

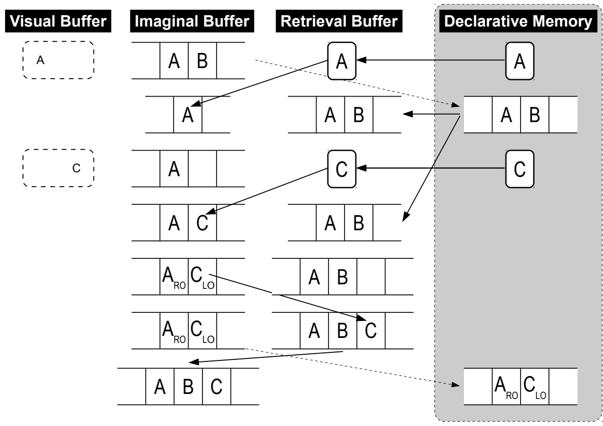


Figure 2: Preferred mental model (PMM) construction. This illustration starts with an incomplete mental model chunk in the imaginal buffer comprising just the information from the first premise. Objects processed by the visual module are linked to term chunks in declarative memory that the model needs to retrieve before it can integrate them into a mental model chunk. Note that the first premise is treated differently because the ACT-R model creates no explicit chunk for it; for simplicity, the corresponding processing steps are omitted here. Subscripts refer to the annotation flags "reference object" (RO) and "to be located object" (LO) necessary for identification in a potential later retrieval (cf. Figure 3).

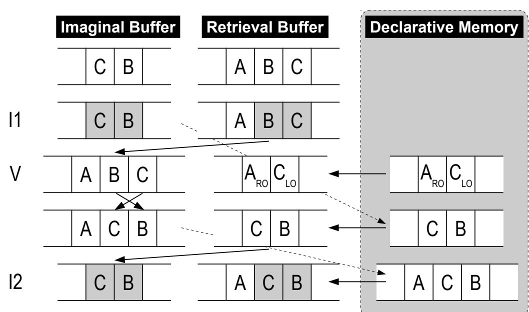


Figure 3: Annotation usage. Labels refer to mental model inspection (I2, I2) and variation (V). Comparisons across buffers are highlighted in gray. Note that this simplified illustration starts with a conclusion in the imaginal and a preferred mental model (PMM) in the retrieval buffer. If in I1 the comparison of conclusion and PMM fails, the PMM is transferred to the imaginal buffer, the conclusion is released to declarative memory, and an annotated premise is retrieved allowing for PMM variation. Eventually, the conclusion is retrieved again allowing for further comparisons in I2.

less time when no competing alternative model has to be used for validating the previous conclusion (cf. Figure 4b).

An important difference to the premise processing phase is that the ACT-R model skips the creation of an additional mental model chunk; instead, it compares the terms in the retrieval buffer with the UNI/PMM in the imaginal buffer directly. Similarly, if in an indeterminate case the PMM fails to match, no modification takes place; instead, a retrieval request for an alternative model follows and eventually the comparison phase restarts. The motivation for this omission of additional mental model creation is that the respective terms are continuously displayed on the screen thus functioning as external memory; any memorization process involving the creation of internal representations would be redundant. In the premise processing phase, however, internal representations are necessary because no complete mental model is presented but has to be derived in multiple steps. Similarly, in the conclusion validation phase no external mental model representation is available. Without generating and keeping internal representations conclusion validation would, therefore, be impossible.

Results and Discussion

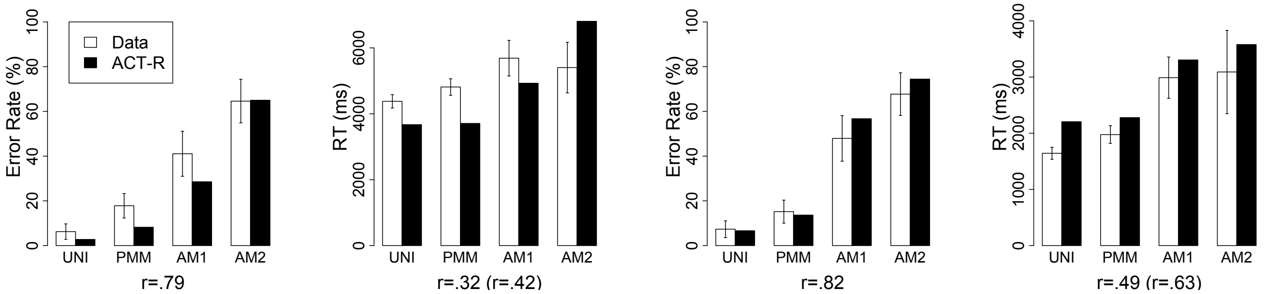
The results support the predictions of the PMMT regarding the preference for the PMM and the increasing reasoning difficulty with the transformation distance of alternative models (cf. Table 1). Table 2 shows the mean error rates (in %) and response times (in milliseconds) for both conclusion and mental model validation.

Table 2: Human data. Mean error rates (in %) and response times (in milliseconds) for both conclusion and mental model validation.

	Conclusion validation		Mental model validation	
Type	Error rate	RT	Error rate	RT
UNI	6.25	4379	7.33	1643
PMM	17.80	4813	15.18	1976
AM1	41.05	5688	47.92	2990
AM2	64.58	5402	67.71	3090

Pearson's correlation was used to test whether human data correlated with model predictions. Correlations were computed on a by-task aggregate level for those tasks with an expected correct response resulting in 120 data points each for human and model means. The rationale behind this is that to correctly reject an invalid conclusion reasoners cannot prematurely terminate the validation process as they would have to test all possible mental models; premature termination is only possible with a correct mental model. Hence, if the task was to reject an invalid conclusion for both data and model predictions we expected no differences across conditions. Please also note that for mental model validation only those cases were subject to analysis in which the externally presented model matched the model that previously validated the conclusion.

Figures 4a and 4b compare data from 24 participants with predictions based on 2400 model runs. Each run processes one of the 24 task sets; hence, with each task set consisting of 48 tasks this resulted in a total of $2400 \times 48 = 115.200$ simulated experimental trials.



(a) Conclusion validation

(b) Mental model validation

Figure 4: Average error rates and response times for conclusion and mental model validation. Conclusion validation either requires a mental model constructed on the basis of determinate premises (UNI) or indeterminate premises (PMM, AM1, AM2). Correlations were computed on a by-task aggregate level only for those tasks with an expected correct response resulting in 120 data points each for human and model means. Response time correlations were computed with and without taking the AM2 condition into account. For the latter, the corresponding coefficient is given in brackets. Error bars show 95% confidence intervals. All correlations were significant at a $p < .001$ level.

Error rate correlations between data and ACT-R predictions showed a significant effect for both conclusion, $r = .79, p < .001$ (cf. Figure 4a) and mental model validation, $r = .82, p < .001$ (cf. Figure 4b).

Response time correlations between data and ACT-R predictions showed a significant effect for both conclusion, $r = .51, p < .001$ (cf. Figure 4a) and mental model validation, $r = .33, p < .001$ (cf. Figure 4b). In the AM2 condition, however, response times for conclusion validation contradict the theoretical assumptions of the PMMT as there is an unexpected decrease in the average response time in comparison to the AM1 condition. However, an error rate of 64.58% results in only few contributing participants when only correct responses are taken into account. Furthermore, a lower response time in this context suggests that participants used the AM1 for conclusion validation rather than the predicted later stage AM2. As a last resort they may even have guessed. A similar explanation may hold for the average response time for mental model validation and a corresponding error rate of 67.71% being comparably high. Correlations, hence, were computed again without taking those tasks requiring an AM2 for conclusion validation into account based on 96 data points each for human and model means. Correlations then improved and showed a significant effect for both conclusion, $r = .61, p < .001$, and mental model validation, $r = .42, p < .001$.

When processing determinate premises with only one resulting mental model (UNI) participants were faster and made fewer errors than in the remaining conditions. When processing indeterminate premises a well-established preference effect could be reported for situations where only the preferred mental model (PMM) was necessary to validate a conclusion; participants were faster and made fewer errors than if an alternative model (AM1 or AM2) was required. They performed poorer, however, than in the UNI condition. According to the PMMT this effect can be explained by additional processes that annotate indeterminate premises allowing for later PMM modifications to an alternative mental model. From an ACT-R theoretical perspective, because these processes require additional time, additional activation decay of the PMM also results; consequently, its retrieval at conclusion validation takes more time and is less reliable. Finally, a "distance effect" for AM1 and AM2 could be reported; the more modifications to the initial PMM were necessary, the less likely they were to be correctly accepted or rejected.

With respect to the mental model validation task human data suggest that differences in response times depend on the number of mental models created in the conclusion validation phase. In the single model condition (UNI/PMM) response times were lower than in the multiple model condition (AM1/AM2). No explicit differences between AM1 and AM2 emerged. For the AM2 condition this suggests that after the PMM reasoners tried only one of the two alternatives.

There are, however, slightly higher predicted response times for mental model validation in the AM2 condition (cf. Figure 4b) that could not be reported for the human data. From an ACT-R perspective there is a clear explanation: As suggested, the mental model validation process starts with the PMM generated in the premise processing phase. Accordingly, to successfully validate a conclusion in the AM1 condition, apart from the PMM chunk, the ACT-R model needed to create an additional chunk for the AM1. In the AM2 condition, however, it had to create chunks for both, the AM1 and the AM2. Hence, even if the same number of mental models are recalled in both conditions AM1 and AM2 —as the response times in the human data suggest—in the AM2 condition the PMM experienced more activation decay because more time has passed since its last usage. Consequently, the mapping of activation to retrieval time results in higher values for the AM2 than for the AM1 condition.

Conclusion

Our starting point was the question of how human reasoners process determinate and indeterminate spatial descriptions and a corresponding cognitive model introduced previously (Ragni et al., 2010; Ragni & Brüssow, 2011). The advantage of the present study is that the stimulus material comprised simple visual presentations of term objects rather than complex linguistic presentation such as sentences. The rationale behind this is that interfering linguistic processes could thus be inhibited. In this respect, the present study bridges a gap between the fMRI study by Fangmeier et al. (2006) and Ragni et al. (2007): Fangmeier et al. investigated only determinate problems but presented the material in a plain and functional way, whereas Ragni et al. covered a wide range of indeterminate problems but presented premise information verbally.

Furthermore, we systematically investigated the processes of mental model variation—the transformation from an initial PMM to an alternative model—that so far has received only little attention. In that context the presented ACT-R model is both an algorithmic foundation of the PMMT and a theory explaining behavioral data obtained from experiments using either linguistic or pictorial representation of premise information. To further establish the model and to investigate the persistence of the outcome of the reasoning process, it was extended and tested with a wider range of tasks consistently presented non-linguistically. It gives detailed insights into why certain tasks are computationally more demanding than others and, apart from being informative about the reasoning process itself, it simulates an additional mental model validation task that is in line with the experimental results indicating a primacy effect for the PMM; the data suggest that reasoners start comparisons by recalling the PMM rather than an alternative. Future work will focus on this particular point as the question of how reasoners proceed, naturally, has implications for the ACT-R model itself. Currently, if the PMM fails to match an alternative model is requested. Restarting the variation phase by retrieving annotated premises, however, would also be a possible strategy.

Acknowledgments

This research was supported by the DFG (German National Research Foundation) in the Transregional Collaborative Research Center, SFB/TR 8 within project R8-[CSPACE]. The authors are grateful to Tobias Sonntag, Philip Stahl, and Andreas Klein for assistance in participants acquisition and experiment implementation.

References

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York: Oxford University Press.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036-1060.

Fangmeier, T., Knauff, M., Ruff, C. C., & Sloutsky, V. M. (2006). fMRI evidence for a three-stage model of deductive reasoning. *Journal of Cognitive Neuroscience*, *18*, 320-334.

Jahn, G., Knauff, M., & Johnson-Laird, P. N. (2007). Preferred mental models in reasoning about spatial relations. *Memory & Cognition*, *35*, 2075-2087.

Johnson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science*, *4*, 71-115.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

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 17th annual conference of the cognitive science society* (p. 200-205). Mahwah, NJ: Erlbaum.

Mani, K., & Johnson-Laird, P. N. (1982). The mental representation of spatial descriptions. *Memory & Cognition*, *10*, 181-187.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.

Ragni, M., & Brüssow, S. (2011). Human spatial relational reasoning: Processing demands, representations, and cognitive model. In W. Burgard & D. Roth (Eds.), *Proceedings of the 25th AAAI conference on artificial intelligence.* San Francisco, CA: AAAI Press.

Ragni, M., Fangmeier, T., & Brüssow, S. (2010). Deductive spatial reasoning: From neurological evidence to a cognitive model. In D. D. Salvucci & G. Gunzelmann (Eds.), *Proceedings of the 10th international conference on cognitive modeling* (p. 193-198). Philadelphia, PA: Drexel University.

Ragni, M., Fangmeier, T., Webber, L., & Knauff, M. (2007). Preferred mental models: How and why they are so important in human reasoning with spatial relations. In C. Freksa, M. Knauff, B. Krieg-Brückner, B. Nebel, & T. Barkowsky (Eds.), *Spatial cognition V: Reasoning, action, interaction* (pp. 175–190). Berlin: Springer.

Ragni, M., Knauff, M., & Nebel, B. (2005). A computational model for spatial reasoning with mental models. In B. Bara, B. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th annual conference of the cognitive science society* (p. 1064-70). 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*, 239-269.

Rauh, R., Hagen, C., Schlieder, C., Strube, G., & Knauff, M. (2000). Searching for alternatives in spatial reasoning: Local transformations and beyond. In L. R. Gleitman & A. K. Joshi (Eds.), *Proceedings of the 22nd annual conference of the cognitive science society* (p. 871-876). Mahwah, NJ: Erlbaum.