CogSci 2019

Proposing a Cognitive System for Universal Mental Spatial Transformations

Kai Preuss & Nele Russwinkel
Technische Universität Berlin, Department of Psychology and Ergonomics, Chair of Cognitive Modeling in Human-Machine Systems (KModyS)

Introduction

Our research presumes the existence of a dedicated, unified cognitive system for the simulation of spatial processes. While mental spatial processing modeling often uses imaginal processes as a substitute, research shows that a distinction between visual and spatial representations is possible (e.g. Knauff & Johnson-Laird, 2002). Other times, modeling concentrates on highly task-specific models with low general validity. Our proposed system seeks to offer a universal, unified and valid platform for these processes.

Our goal is to create an extension for the cognitive architecture ACT-R in the form of a biologically plausible module that enables modeling of mental spatial transformation. It is conceived as a slim alternative to a proposed module by Gunzelmann & Lyon (2007).

We propose the operations made by this module to depend on the complexity of mental transformation: the *extent* and the *amount* of necessary transformational steps.

While the extent has a direct influence on the required time for transforming a mental representation of an object, the amount of transformations already applied to this representation is included as a limiting factor - object representations are increasingly difficult to maintain with additional mental changes. This could explain certain reaction time effects between task difficulties in a mental folding task as being caused by cognitive re-encoding processes.

To gauge the necessity and functionality of cognitive models relying on such a module, a baseline model lacking such a module is presented and compared to data collected in a new mental folding experiment. The shortcomings of this model subsequently demonstrate the need of a system dedicated to mental transformations:
We believe a dedicated mental spatial transformation to be a significant addition for modeling spatial behavior, both in ACT-R and other cognitive architectures.

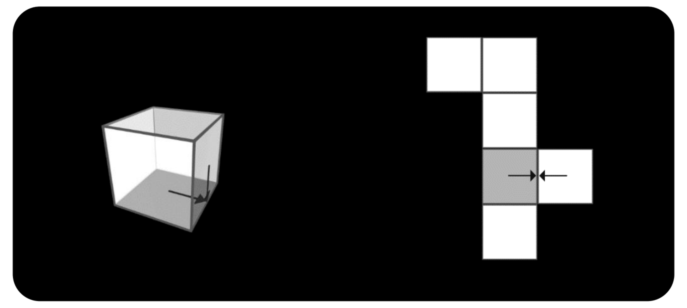
Keywords: Spatial Cognition; Cognitive Modeling; ACT-R; Mental Folding; Mental Transformation

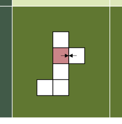
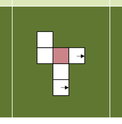
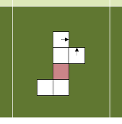
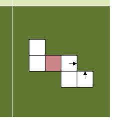
The Experiment

The experiment conducted was a cube folding task (originally by Shepard & Feng, 1972; in a variant by Wright et al., 2008): Participants were required to make an assessment whether two presented objects – one cube, one folding pattern – were equal.

Difficulty varied between 4 levels, each requiring a specific number of transformation steps to allow for a decision between the objects. 5 blocks were presented with 120 trials each, adding up to a total of 600 trials per participants.

This resulted in 40 valid datasets overall (20 male, 20 female) & balanced by ego- or allocentric orientation strategy (measured by pretest Reference Frame Proclivity Test, Goeke et al). Measured was behavioral data (such as reaction times or errors), as well as EEG and Eye tracking.



Difficulty Level	A	F	G	H
Squares Carried	None	2 + 1 + 1 = 4	3 + 1 + 1 = 5	3 + 2 + 1 = 6
Example Pattern				

Left: An example of a trial as used in the experiment.

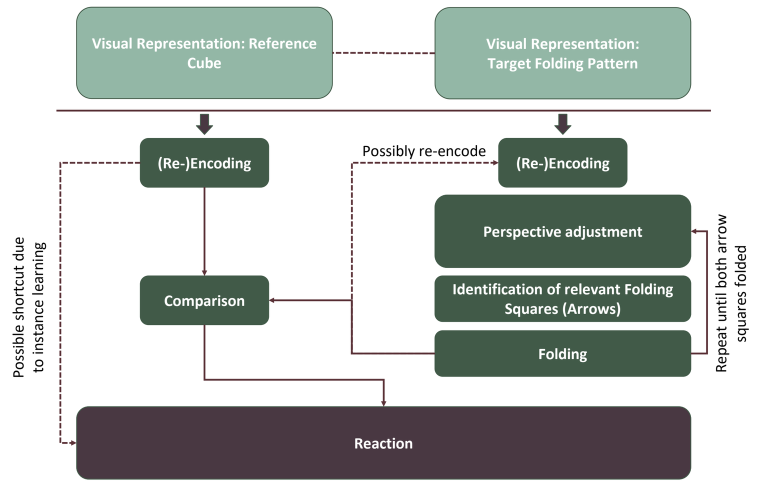
Right: Difficulty Levels used during the experiment, differing in the number of necessary transformation steps to assess the final arrow positions.

The Baseline Cognitive Model

As a benchmark for later improved models, an ACT-R model was created that uses default ACT-R modules and capabilities to simulate solving efforts during the mental folding task, without an additional spatial module. Being a basic reference point, it should prove inferior in both data fit and cognitive plausibility to an enhanced model, thus attesting for the need of a spatial module.

Spatial information about both folding patterns and cubes (such as cube face relations or perspective changes through folding) is available as world knowledge. During task solving, the model retrieves this information based on cues to the visual module, stores information about arrow positions on both cube and folding pattern and subsequently creates mental images of folding paths for each arrow on the folding pattern. Depending on this path, both the presumed eventual cube face of the arrow as well as its direction are now appropriately transformed by a memory retrieval for each successive path step. The resulting mental images for the correctly folded arrows are then compared to the actual arrow positions and directions on the reference cube.

To simulate learning behavior, a simple pattern recognition mechanism was introduced: positions of arrows and base of the folding pattern are remembered as "same" or "different" to the arrow positions on the cube. At the start of a trial, an attempt is made to retrieve a memory pertaining to that trial. If successful, a decision can be reached more quickly.



A process model, highlighting the main cognitive processes in solving the mental folding task. These processes are simulated in the cognitive model.

The Spatial Module

Our proposed spatial system is aspiring to allow for modeling of several paradigms in spatial transformative cognition. For this, three general goals were set:

- Sufficient explainability: the resulting ACT-R module must be able to explain known effects in spatial modeling, such as growing time cost through growing complexity or inter-individual differences
- Universal applicability: without over-fitting to a specific variant of mental spatial processing, several paradigms should be feasible with the module, while still maintaining falsifiability
- Process validity: incorporating cognitive plausibility in terms of limitations, time demands, inaccuracies and mistakes

In the proposal by Gunzelmann & Lyon (2007), the module's functionality partially overlaps with existing module functions, as its 4 buffers take on wider functionality – an episodic buffer works as a mediator for the declarative module, while the functionality of proposed egocentric and environmental buffers seems mostly identical. The functionality of default modules can, however, account for a large share of that of the proposed structures, putting their necessity into question.

Possible factors influencing the time and viability of a transformation have been suggested:

- Shepard & Metzler (1971) found a linear relationship between degree of discrepancy between objects and reaction time
- Shepard & Feng (1972) reported unsolvability of harder difficulty levels (requiring more than 6 transformative operations), suggesting an upper limit to mental transformation capacity
- Lotz & Russwinkel (2016) suggested a decay factor for spatial representations, with demanding tasks requiring additional visual or memorial re-encoding
- Neely & Heath (2010) propose a factor for the complexity of a transformation

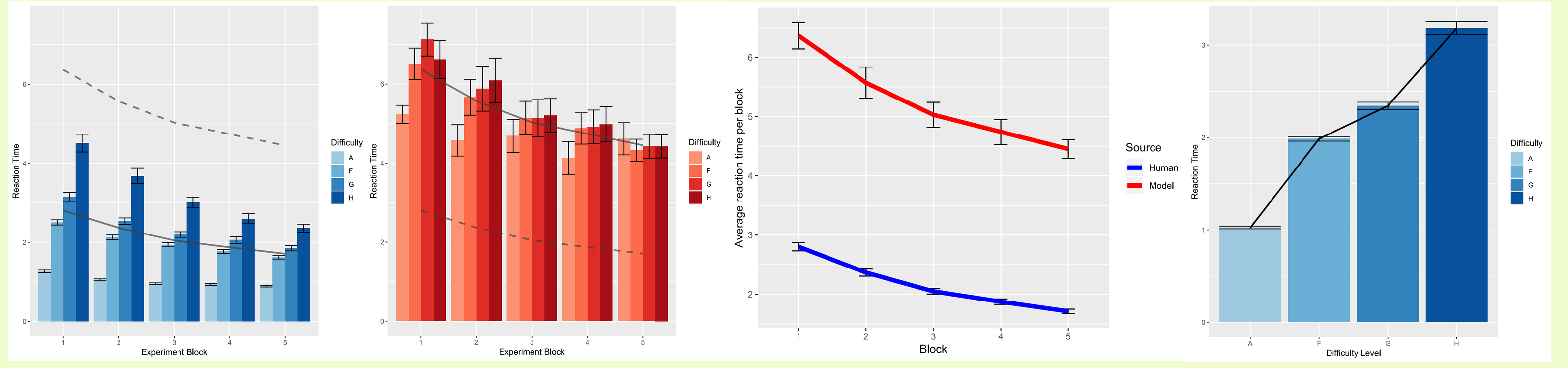
Our approach suggests two available buffers – a spatial storage buffer and a spatial command buffer. Based on the literature mentioned above, our module computes the predicted time for an operation based on the amount of change necessary, mediated by the number of transformative steps taken so far. Additionally, its actions are limited by a maximum number of sequential transformations, after which re-encoding is required. So far, its abilities include translation, rotation, scaling and comparison of objects in (up to) three-dimensional space. Objects are represented as 3D point clouds with added semantic information. Inter-individual differences could be partially modelled by changing parameters for latency and upper limit of transformations, respectively.

Results

Reaction times: significant main & interaction effects for the factors Difficulty and Block*: reaction time increased with higher difficulty, decreased over the course of the experiment and the decrease was more pronounced for more difficult trials. Faster reaction times for egocentric over allocentric solvers.

Baseline Model: comparison with behavioral data revealed a highly significant correlation with reaction times per experiment block**: instance learning seems to be vital for spatial pattern memorization.

* Difficulty Level: $F(1.74,67.96) = 282.86$, $p < .001$; Experiment Block: $F(2.16,84.22) = 144.17$, $p < .001$; Interaction: $F(5.94,231.84) = 12.57$, $p < .001$
** $r > .999$, $p < .001$



- *Far Left:* Average participant reaction times in seconds per Difficulty Level and Experiment Block. Black solid line denotes the learning effect, averaged over levels of difficulty. Dashed line shows model learning effect for comparison.
- *Center Left:* Model reaction times in seconds per Difficulty Level and Experiment Block. Black solid line denotes learning, averaged over levels. Dashed line shows human learning for comparison.
- *Center Right:* Comparison of human RT (blue) with model RT (red), averaged over level and compared over Experiment Block
- *Far Right:* Mean effect of difficulty level on human RT. The sums of squares carried necessary during folding are 0 in Level *A*, 4 in Level *F*, 5 in Level *G* and 6 in Level *H*.

Open Questions & Outlook

The proposed system is still in an exploratory stage, so several design decision are still subject to change.

While point clouds are an effective way to represent an object in three-dimensional space, its surface is difficult to incorporate – textures or colors could be added as semantic information, but might not be integrated into the actual 3D object itself in this way.

While basic support for differentiation of inter-individual spatial capabilities is offered, further research aims to explore functional differences in, for instance, reference frame proclivity. However, brain imaging results might still provide evidence of spatial processing that is incompatible with the presented structure. This could mean evidence for a more elaborate system akin to Gunzelmann & Lyon (2007), or even negate the need for a dedicated spatial module altogether.

For classic experimental paradigms in spatial transformation, our module could improve efficiency, accuracy and validity of future cognitive models. As 3D information can now be processed by ACT-R, it offers capabilities beyond the traditional application of interacting with screen surfaces and could thus be used for product development, testing ergonomics and usability of prototypes.

Other spatial paradigms, like navigation tasks or syllogistic reasoning, could also profit from a dedicated spatial module. Additionally, its abilities in point cloud perception lend themselves to pattern recognition in general, extending ACT-R with potential pattern learning capabilities.

ACT-R in a nutshell

Used to create psychologically plausible models of human behavior in specific situations by simulating the interaction of modules, each representing a cognitive system (visual, motoric, imaginal etc.). Modules can be controlled by addressing their buffers, which act as command interfaces.

Each action of a module has a time cost associated to it, enabling ACT-R to closely predict reaction times or error rates, as well as learning behavior. In recent years, it was also used to predict fMRI results (Borst & Anderson, 2015).

End result: Validation and/or prediction of human data

References:

Borst, J. P., & Anderson, J. R. (2015). Using the ACT-R Cognitive Architecture in Combination With fMRI Data. In *An Introduction to Model-Based Cognitive Neuroscience* (pp. 339–352). New York, NY: Springer New York.

Gunzelmann, G., & Lyon, D. R. (2007). Mechanisms for Human Spatial Competence. In T. Barkowsky, M. Knauff, G. Ligozat, & D. R. Montello (Eds.), *Spatial Cognition V Reasoning, Action, Interaction* (pp. 288–307). Springer.

Knauff, M., & Johnson-Laird, P. (2002). Visual imagery can impede reasoning. *Memory & cognition*, *30*(3), 363–371.

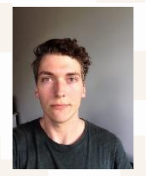
Lotz, A., & Russwinkel, N. (2016). Modelling different strategies in mental rotation. In T. Barkowsky, Z. F. Llansola, H. Schultheis, & J. van de Ven (Eds.), *Proceedings of the 13th biannual conference of the german cognitive science society* (pp. 35–38). Bremen.

Neely, K. A., & Heath, M. (2010). Visuomotor mental rotation: Reaction time is determined by the complexity of the sensorimotor transformations mediating the response. *Brain Research*, *1366*, 129–140.

Shepard, R., & Metzler, J. (1971). Mental Rotation of Three-Dimensional Objects. *Science*, *171*, 701–703.

Shepard, R. N., & Feng, C. (1972). A chronometric study of mental paper folding. Cognitive Psychology, 3(2), 228–243.

Wright, R., Thompson, W. L., Ganis, G., Newcombe, N. S. & Kosslyn, S. M. (2008). Training generalized spatial skills. Psychonomic Bulletin & Review, 15(4), 763–771.


Kai Preuss
Email: preuss@tu-berlin.de
Phone: +49 (0)30 314 27821


Nele Russwinkel
Email: nele.russwinkel@tu-berlin.de
Phone: +49 (0)30 314 72408

Cognitive Modeling in dynamic Human-Machine Systems

Office MAR 3-2
Marchstraße 23
10587 Berlin

KModyS

Technische Universität Berlin

www.kmodys.tu-berlin.de