

---

## **Spatial Structures and Visual Attention in Diagrammatic Reasoning**

Doctoral thesis in artificial intelligence, 2009 (Dr. rer. nat., *summa cum laude*)  
by Sven Bertel

Cognitive Systems Group, Department of Informatics, University of Bremen  
Contact: bertel@informatik.uni-bremen.de

Supervised by Prof. Christian Freksa, Ph.D.

### **Extended Abstract**

The thesis addresses questions of relating diagrams and mental representations in diagrammatic reasoning through eye movement research. It proposes model-based representations of attention for improved human-computer cooperation. In particular, the thesis proposes in theory and details in practice a computational framework for live capture and analysis of eye movement data in diagrammatic reasoning scenarios. The analysis includes an on-line generation of hypotheses about spatial mental representations currently held by human reasoners. The generated hypotheses may be employed to guide reactive and proactive behavior in semiautomated reasoning systems more (cognitively) adequately than previously possible, for example, in live human-computer collaboration or tutoring settings. Among other fields of application, the framework may be used to externally influence mental visuo-spatial reasoning in selective ways through administering specific patterns of sensory cues. Additionally, the presented theoretical and practical approaches may significantly contribute to the development of novel techniques and research methodologies aimed at better understanding human visuo-spatial reasoning and problem solving.

Key themes of the thesis are spatial reasoning with problems in diagrammatic formats; the binding of mental and external representations through common spatial structuring mechanisms; visual attention and visual diagram inspection during spatial reasoning. Central aims of the thesis are to extend the current theoretical and conceptual basis for understanding human reasoning with diagrams, to propose novel practical approaches towards increasing such understanding, and to provide novel theoretical and practical foundations for more cognitively adequate joint human-computer reasoning with diagrams.

---

**Background.** Human reasoning about spatial environments, spatial configurations as well as design problems often involves external diagrammatic representations such as sketches, drawings, or maps. Frequently, a good use of diagrams is prerequisite to performing well in these tasks. This is for at least two groups of reasons: First, organizational (i.e., spatial) phenomena of our physical environment have deeply shaped the organization and functioning of human perception and cognition in general. Also, many cognitive functions regularly employ and rely upon specific external spatial structures, such as in interpreting perceptual stimuli or offloading computation and knowledge to the environment. Mental reasoning is dependent on specifics of employed external representations and external structural properties often influence mental spatial representations. In the case of diagrammatic reasoning, such influence is manifest in an overlap of involved mental representations and processes with those involved in visual perception as well as in mental spatial reasoning.

Second, employing diagrams holds additional, specific benefits for spatial reasoning and problem solving: Representational properties of diagrams (e.g., concerning spatial indexing of information or representational specificity) lead to favoring certain (perceptual) inferences and structural interpretations over others. This effectively introduces heuristics which help reduce inferential and representational complexities. Re-representation sequences that consist of repeatedly externalizing visuo-spatial knowledge in diagrammatic formats and then reading off information from diagrams are in fact a powerful driving force in many visuo-spatial reasoning and problem solving processes; they have been extensively studied, for example, in architectural design research.

Diagrammatic representations are usually apprehended through vision. Mental processing capacities in visual perception and spatial reasoning are limited and we find that information is in fact read off in sequential, piecemeal and highly selective manners. The sequence may be influenced by external factors such as structure, content, or visual features as well as by mental factors such as reasoning strategies, currently held mental representations, or task. Various, partially contradictory evidence exists concerning interdependencies of eye movements and overt as well as covert shifts of visual attention; still it seems warranted to assume that an analysis of the visual inspection sequence during diagrammatic reasoning may often yield much useful information on involved mental reasoning processes and representations. Two assumptions are central: Differences in problem solving performance and strategies can be seen in differences of eye movement patterns. Conversely, hypotheses on a reasoner's current mental representations can help explain his or her visual inspection sequences during reasoning.

---

Many spatial configuration and design problems show great potential for partial automatization as various (but by no means all) problem solving steps can be formally specified and outsourced to computational problem solvers. Realizing this potential for partial automatization crucially hinges on the ability to establish a functional and effective sharing of reasoning between human cognitive and computer-based processes. We argue that, for successful performance in many shared reasoning tasks (e.g., in design), it is crucial that the automated reasoning processes dynamically and closely adapt to reasoning steps initiated by the human reasoner. Such dynamic and close adaptation can only be achieved through systems that rely on monitoring relevant human behavior (e.g., eye movements in diagram inspection or drawing acts during sketching) and analyze the registered data on-line with respect to the human's current cognitive states, reasoning processes and strategies.

**Approach.** This thesis addresses fundamental aspects of how such shared human-computer reasoning can be realized. To this end, it presents original theoretical, empirical and computational research into the relevant cognitive, perceptual and computational processes and representations. For the case of human reasoning with diagrams, it analyzes the specifics of fundamental reasoning and communication asymmetries that exist in shared human-computer reasoning and it identifies sets of involved representational and procedural key factors and bottlenecks; it proposes the notion of anticipatory cognitive computing in which predictive cognitive processing models can inform the computational part in the reasoning duet about the human's current cognitive state, such as regarding mental problem solving properties (e.g., on preferences in mental model construction or a directing of attentional focus). This thesis proposes a methodological framework for live anticipatory cognitive computing that is directed at enabling the computational reasoning partner to better adapt to current mental processes and states of the human partner, and thus to allow for more efficient and effective human-computer collaboration to emerge, leading in turn to a more efficient and effective solving of spatial configuration and design problems. Imminent implications are discussed for design research and practice as well as for other related fields.

Specifically, this thesis targets the interrelations between visual fixations, the visual focus of attention in diagram-based reasoning and foci in higher-level reasoning (e.g., in spatial problem solving). In doing so, it breaks ground for a new interdisciplinary field of spatial cognition research that sits at the junction of diagrammatic reasoning, visual attention and human-computer collaboration. The thesis chooses a model- and spatial structure-based approach for

---

describing human spatial problem solving and it capitalizes on human individual and general preferences in model construction and selection; it proposes a combination of spatial mental model- and focus-based interaction schemes and instantiates the schemes in a decision-model in which information of spatio-temporal distribution of visual attention triggers model-based selection processes. A graduated, hybrid network model is proposed and implemented that represents dynamic flows of attention-induced activation over different problem parts. On a fundamental level, the thesis addresses issues of how the preferential construction of spatial mental models can be stabilized or destabilized by actions of the computational reasoning partner and on how and when visual fixation data can be of help for assessing human cognitive states in diagram-based spatial problem solving.

The thesis discusses a series of conducted empirical pilot studies on eye movement patterns during diagrammatic problem solving, namely with matchstick configuration puzzles such as this one:

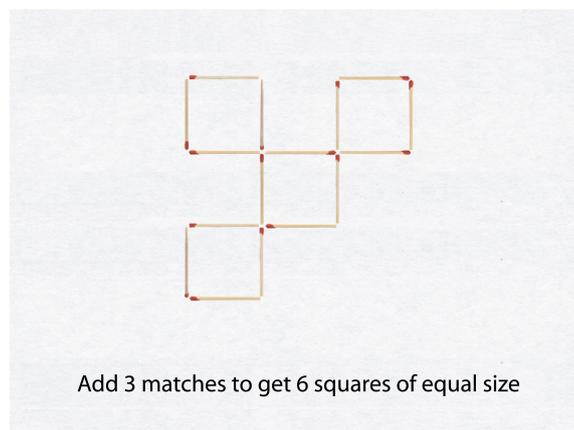


Fig. 1: Matchstick problem solving – an exemplary reasoning problem: The diagrammatic configuration (*above*) needs to be changed through the described actions (*add 3 matches*) so that it will match the qualification below (i.e., consist of six squares of equal size).

Results from the studies are analyzed with respect to processes during mental construction of solutions as well as to the constructed models. The analysis points, among other issues, to eye movement patterns which correspond to functional phases during problem solving (such as relating specific diagrammatic with specific propositional parts), to eye movement patterns that differentiate different solution models (e.g., by binding partial solutions into a global one) and to patterns that reflect individual problem solving styles.

---

Finally, the thesis presents the design and implementation of an integrated demonstrator system for the proposed methodological framework. The characteristic traits of model- and focus-based computing and its effects are demonstrated for an exemplary class of diagrammatic configuration problems and based on the capturing and live analysis of eye fixations and other data.

The thesis takes an inherently interdisciplinary stance in the presented discussions and proposed approaches. It draws upon and contributes to the sets of methods available in artificial intelligence and the cognitive sciences. It has additional ramifications for conceptualizing cognition-oriented human-computer interaction in general, and for constructing human-computer interfaces for design and configuration in particular.