Chapter 1

Introduction

Human reasoning with spatial information often takes place in the presence of and in interaction with external representations of spatial information. For example, wayfinding and orienting may involve the use of maps; reasoning about spatial configurations may involve various kinds of diagrams; or architectural designing may require the skillful production and reading of hand-made sketches or plans. It is one of the characteristics of these settings that the involved mental reasoning processes are not independent of the specifics of the employed external representations: Bits of information may be read off from an external representation and come to be represented mentally, often in forms that retain certain structural aspects of the external representation. Conversely, results of mental reasoning may be externalized into a sketch or mental spatial representations may be mapped in aspects of content or structure against an external diagram.

Sequences of steps of reading off of information from external spatial representations, of mental reasoning in the presence of such representations and of (potentially) modifying an external representation based on the outcome of such reasoning have been variously described as being vital for good performance on spatial tasks in many areas, including diagrammatic reasoning or design. Vision is the main perceptual channel through which an external diagram or a map are apprehended. As visual perception and the mental faculties involved in spatial reasoning are both limited in capacity, information is read off from an external spatial representation in a sequential, piecemeal and selective manner rather than all at one time. It has been demonstrated that the sequence of visual inspection steps may be influenced by factors that lie in the structure, content and visual features of an external representation as well as by factors determined by mental reasoning, mental spatial representations, or memory.

Analyzing the inspection sequence may therefore help to assess information on involved mental reasoning processes and representations: In short, if one knows where a reasoner looks and when during the reasoning process, one may hypothesize aspects of his mental reasoning processes including those of structure and content of the involved mental representations. Conversely, one may try to explain why he looks when and where during the reasoning based on the external and his hypothesized mental state of the world.

The relating of mental and external representations and processes during spatial reasoning has, first, direct implications for our understanding of such reasoning during the presence of external representations and, more generally, of how the human mind operates in the world and on representations of it. It has, secondly, implications for human-computer interaction: if hypotheses about aspects of the mental state of affairs are available during human reasoning computer-based applications may exploit them for either the good or the bad of the reasoning outcome. For instance in tutoring scenarios, hints may be given at specific moments based on a human reasoner's hypothesized mental state to improve the reasoner's learning on how to solve certain types of problems. In multi-player games, a computer-generated adversary may exploit hypotheses about which problem aspects a human has or has not mentally represented, and in which ways, to generate actions which can best hurt the human's chances to win the game. In collaborative settings of human and computational spatial reasoning, a computational reasoner may use hypotheses about the structure of the human's mental representations or about his reasoning strategies to adapt its own actions to these representations or strategies so as to improve overall effectiveness and efficiency of the collaboration.

1.1 An Exemplary Reasoning Problem

Fig. 1.1 shows an example of a spatial reasoning problem which is based on an external diagrammatic representation. The problem consists of a diagrammatic and a sentential part. In the sentential part, actions are described that should be executed to modify the diagram and and to bring it into a different, distinct configuration.

Let us imagine that the problem is shown to a reasoner and that he is asked to try to solve it while looking at it without being able to physically modify the diagram (i.e., no physical changes to the diagram are permitted). Then, the reasoner will need to inspect the problem, match the configuration in the diagrammatic part against the description given in the sentential part,

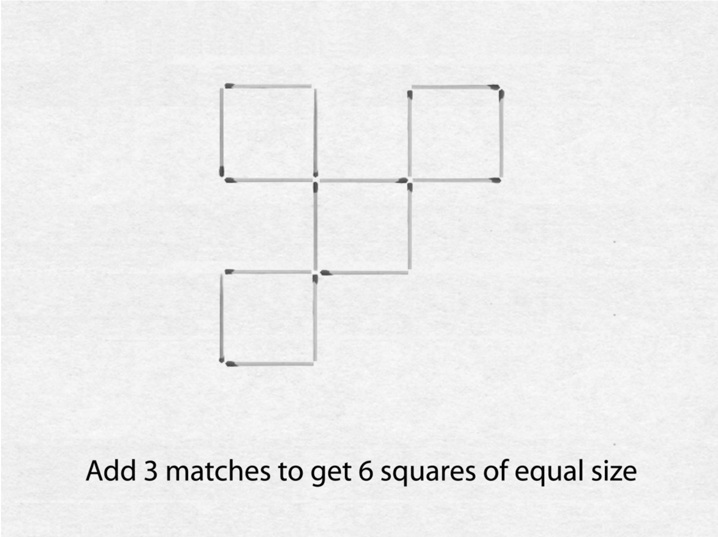

Figure 1.1: Matchstick problem solving – an exemplary reasoning problem: The diagrammatic configuration *(above)* needs to be changed to match the description given below (i.e., consist of six squares of equal size) by executing the permitted actions *(add 3 matches)*.

establish where both differ, understand which actions are permitted to change the given configuration into the described one, build up a partial or complete mental representation of the problem and apply successive changes to the given configuration in order to find a sequence of changes that effectively solves the problem. All 'changes' to the configuration need to be made and maintained purely mentally as no physical alterations are allowed. Also, the reasoner needs to mentally track the changes which he made and check whether their sequence is still within what is permitted by the sentential description. Many different problem solving approaches can be imagined: For example, our human reasoner may apply sophisticated problem solving strategies such as heuristics derived from similar problems which he had solved in the past or he may simply alter the currently mentally held configuration at random and hope to find a solution through such procedure.

In terms of interaction of mental reasoning with diagrammatic content, the complete initial configuration may either come to be fully mentally represented at the beginning, leading to a completely mental problem solving process, or it may only be mentally held in part at any given moment. In the latter case, one may hypothesize that repeated checks are needed between what is mentally held and what is shown in the diagram and that those problem parts or aspects which are mentally represented will change over the course of the reasoning process. Which ever may be the case, one may surmise that once a potential solution has been found it may be compared to the goal's specification in the sentential description before an answer is given. At the end of a successful problem solving process, the reasoner should be able to provide a solution (i.e., in our example problem, to indicate where the 3 additional matches should be placed). In the example, several different solutions are possible; see Fig. 1.2 for details.

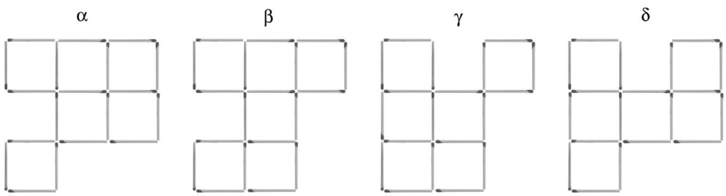

Figure 1.2: Four different solution models for the matchstick problem of Fig. 1.1

Let us now revisit our initial discussion of how a sequence of visual inspection steps of a diagrammatic problem may not only be influenced by factors that lie in the structure, content and visual features of an external representation but also by factors determined by mental reasoning processes and strategies, by mental spatial representations, or by memory. The four valid solutions to our example problem differ structurally in their external representations as diagrams; as each solution would have to be mentally constructed when given as an answer by a human reasoner one can assume that their corresponding mental representations would differ similarly. They may be the products of entirely different reasoning strategies or they may be produced by variants of the same strategy. The important idea is that differences in the solution's mental representations may show in the sequences of visual inspection steps by which a reasoner perceives the problems during reasoning. Should this in fact be the case then one would be able to derive an individual reasoner's constructed solution model from this sequence.

1.2 Themes of this Thesis

This thesis addresses spatial reasoning with problems in diagrammatic formats. Its first main focus is on spatial structures that play a role in binding mental and external representations of a problem during reasoning and that have profound influences on the reasoning's course and outcome. Its second main focus is on processes of visual attention and visual inspection of diagrams during spatial reasoning and on how the involved structures can help with interpreting observed shifts of visual attention in terms of problem solving processes and involved mental representations. To this end, this thesis suggests novel ways of investigating and computationally describing the theories, extends in theory and practice the methodological foundation for cognitively adequate human-computer interaction in spatial reasoning with diagrams and related areas, and finally describes why and how the proposed approaches possess concrete benefits for a number of application scenarios.

Furthermore, this contribution provides an extensive synopsis of the relevant literature from an interdisciplinary point of view as it gives an interpreted account from a combined computational, cognitive, psychological and design studies perspective. The discussion of the relevant phenomena aims at connecting mental representations and external diagrams in spatial reasoning through studies of eye movements. The combination of pairwise in-depth discussions of visual attention and spatial reasoning, of eye movements and visual attention, and of eye movements in spatial reasoning provide a novel, systematic account which was previously not available in this form. The account opens up a number of new research paths into a promisingly developing and inherently interdisciplinary field of spatial cognition that sits at the junction of diagrammatic reasoning, visual attention and human-computer collaboration.

Specifically, the key themes comprise the following: (a) Human spatial reasoning and problem solving with diagrammatic problems; (b) diagrammatic reasoning; (c) eye movements in spatial diagrammatic problem solving; (d) issues of control and foci of attention in diagrammatic problem solving; (e) mental models, mental model-based reasoning and computational models of cognition; (f) anticipatory cognition; (g) a systematic, integrated basis for cognitively adequate support in in spatial / diagrammatic problem solving, and implications for human-computer interaction; (h) spatial and diagrammatic reasoning in architectural design and spatial planning. Fig. 1.3 provides a topic map that depicts the most relevant topic interrelations to be discussed within the scope of this thesis.

The chief questions for which the current contribution provides answers are these: Can human spatial reasoning with diagrams be described through computational cognitive modeling such that the resulting models can be employed by an adaptive computational system for providing cognitively adequate assistance to the human reasoner in order to improve reasoning performance and efficiency? If so, how? And, how can these models then be usefully employed

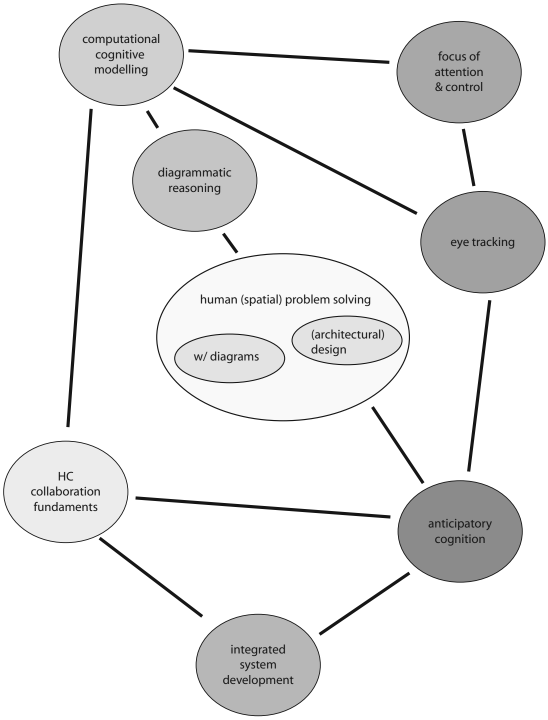

Figure 1.3: Main topics within the thesis and relevant discussed interrelations between topics.

for collaborative human-computer spatial reasoning or tutoring scenarios, such as in architectural design or problem solving?

1.2.1 Core Contributions

The core contributions of the present thesis are twofold: First, as described above, it systematically 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. It does so by thoroughly sorting and discussing existing work from a number of different fields and by interpreting this work for the dual purpose of increasing our knowledge about cognitive processes involved in spatial reasoning and of systematically applying this knowledge to human-computer collaborative settings.

The second field of core contributions of this thesis lies in model-based representations of attention for human-computer interaction. Specifically, the presented approach involves (a) proposing spatial structures and mental models as joint vehicles to interpreting attentional shifts as observed through eye movements in diagrammatic reasoning, (b) proposing and implementing a computational framework in which such observations can be made and accordingly interpreted in live reasoning scenarios, (c) proposing methods for how such framework can be employed in order to couple processes of human- and computer-based spatial reasoning more closely, thereby permitting more effective and efficient human-computer reasoning, and (d) detailing the approach's perspectives in the area of design and assessing its strengths and weaknesses with respect to design tool requirements.

1.3 Structure

This remainder of this thesis is grouped into two parts which largely correspond to the two main fields of contributions discussed above. Part I comprises Chapters 2, 3, 4, 5 and 6 and examines in various ways how mental representations and external diagrams can be related through eye movement research. In particular, Chapter 2 discusses how human (visual) perception and (spatial) cognition have been profoundly shaped by structural properties of the environment and how human cognition can be only understood given the presence of corresponding structures. The role of spatial analogies for mental formats of spatial knowledge representation and spatial reasoning are examined. Chapter 3 continues the consideration of analogical structures and processes in that it offers a detailed discussion of spatial reasoning with mental model, mental images and diagrams. As diagrams are visually apprehended, Chapter 4 examines the interrelationship between (selective) visual attention and eye movements. Chapter 5 provides a more technical interlude and discusses various approaches of measuring eye movements and of analyzing eye movement data. Chapter 6 focuses on discussing works on the role of eye movements for and induced by higher-level cognitive processes such as spatial reasoning and problem solving. The chapter concludes with an excursion on reasoning in collaborative setting, such as among humans or in mixed human-computer settings.

Part II of this thesis comprises Chapters 7, 8, 9, 10 and 11 and focuses chiefly on model-based representations of attention for human-computer collaboration. In detail, Chapter 7 presents a function-oriented interpretation of mental theories, models and images, suggests spatial structures and graduated levels of problem abstraction for analyzing visual attention data during diagrammatic reasoning, and presents and discusses selected findings from a series of empirical investigations into model-based reasoning with diagrams. Chapter 8 presents a computational system for live model-based eye tracking, examines selected system components and discusses theoretical and practical aims during the system's development for cognition research and application scenarios. Chapter 9 presents various implications that arise for human-computer interaction from combined model- and attention-based real-time approaches. Chapter 10 provides detailed arguments that relate design problem solving to the classes of diagrammatic problems in focus in preceding chapters and makes the case that the theoretical and practical approach of this thesis provide novel and significant contributions to the development of cognitively adequate joint human-computer design systems. Finally, Chapter 11 offers a summary and conclusions of the presented work and offers an outlook on future research paths for which the present thesis provides conceptual and practical foundations.

Part I

Connecting Mental Representations and External Diagrams Through Eye Movement Research

The chapters in the first part of this thesis will in particular collect, sort, discuss and evaluate evidence for the following claims: *(a)* Human perceptual and cognitive systems capitalize on the spatial structure that abounds in the world. The systems construct reality by anticipating and interfacing with environmental properties and outcomes of environmental processes. *(b)* The systems capitalize on environmental structures to the extent that spatial organization permeates all levels of mental processing – from neural and cortical to psychological and cognitive architectural levels, from the processing of sensory input to higher-level cognitive functions, including problem solving and reasoning, and from modal mental representations to abstract, intermodal ones. Multiple and distributed mental representations of space exist. *(b/2)* As a corollary, human cognition only makes sense when it is situated in the environment in, for and through which it has evolved. This has implication for the study of cognition. *(c)* Human spatial reasoning involves processes which operate on spatial mental models and visual mental images. The particularities of these formats have profound structural and procedural implications for human spatial reasoning. *(d)* Diagrams are important for a variety of cognitive tasks. Diagram-based tasks offer remarkable specific possibilities for the study, disturbance and support of mental spatial reasoning activities. *(e)* Attentional focus in visuo-spatial reasoning is governed both by bottom-up and top-down processes. The top-down processes can be modulated by the task at hand, goals and memory contents, among other factors. *(f)* Under normal conditions, attentional shifts and eye movements are synchronized and attentional and visual foci coincide on a common visual target. *(g)* Methods and techniques of measuring and analyzing eye movements have finally reached a stage on which it has become technically practicable to start employing them for human-computer interaction. *(h)* Eye movements depend on and reflect the cognitive task at hand, often also during problem solving. They reflect mental reasoning strategies and aspects of constructed mental representations. The associated oculomotor patterns are inherently part of spatial memory and are functionally relevant for mental spatial reasoning. *(i)* Cognitive modeling offers the necessary tools and methods to study human cognition in general and spatial reasoning in particular. The field has progressed to the point where it can inform on currently held mental spatial representations. Based on such information, human-computer interaction becomes a possibility that selectively and effectively leads human spatial reasoning. *(j)* A number of classes of relevant spatial, diagrammatic reasoning tasks exist that can benefit from human-computer interaction with such properties.

Chapter 2

Mental Representations of Space and Mental Spatial Reasoning

To begin with, it seems important to point to the perceptual appreciation of the environment as the ultimate foundation of all mental theory construction. As Minsky puts it in *The Society of Mind* on *the shape of space* (1985, chapter 11, p. 110):

> "The brain is imprisoned inside the skull, a silent, dark, and motionless place; how can it learn what's it like outside? The surface of the brain itself has not the slightest sense of touch; it has no skin with which to feel; it is only *connected* to skin. Nor can a brain see, for it has no eyes; it is only *connected* to eyes. The only paths from the world to the brain are bundles of nerves like those that come in from the eyes, ears, and skin. How do the signals that come through those nerves give rise to our sense of "being in" the outside world? The answer is that this sense is a complicated illusion. We never actually make any *direct* contact with the outside world. Instead, we work with models of the world that we build inside our brains." [original emphases]

2.1 Structures Within and Without

Understood in this vein, perception is a necessary basis for cognition to occur. Perception is cognition's incoming channel on the state of the world 'outside' and as such it has profound influence on how cognitive processes operate. In his introduction to *Autopoiesis and Cognition*, Humberto Maturana suggests that, instead of asking how an organism can obtain the information that it needs from the environment, one should rather ask "How does it happen that the organism has the structure that permits it to operate adequately in the medium in which it exists?" (Maturana & Varela, 1980, , p. xvi). It is through some specific organization of perceptual processing that structures in the physical environment are appreciated and that they, indirectly, become effective in the structuring of cognitive processing (such as through evolutionary processes, through learning, etc.). Perceptual properties influence cognition in structural and functional matters (see Fig. 2.1 for an illustration of the postulated interrelations).

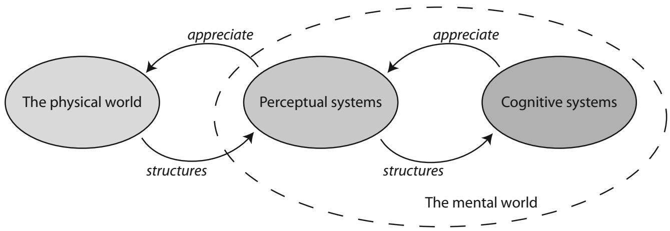

Figure 2.1: Structural phenomena in the physical world influence the structures of perceptual systems which in turn structurally influence cognitive processing. We will see in later parts of this section in which respects this conceptualization may be too simplistic.

On the one hand, perceptual properties determine which aspects of the world are available for cognitive processing and how they are available. For example, much of the functioning of the visual system relies on the ability to detect changes in perceptual stimuli, such as a change in contrast or in orientation of contrast. One of the causes for this property lies in how visual stimuli are successively grouped and processed along the visual pathways: Receptive fields of neurons on a retinal level can detect spots of light; their outputs feed into cells in the lateral geniculate nucleus where a concentric center-surround organization can discriminate edges from non-edges; outputs of the LGN cells are again combined by simple cells in primary visual cortex to signal the presence of an edge with a certain spatial orientation, and so on to the detection of increasingly complex features (e.g. Hubel & Wiesel, 1977). In some sense, perceptual levels work as filters that pass on and pre-structure certain information to cognitive levels, such as that about oriented edges in a visual scene, while not passing on other information. One can argue that the adequate functioning of the filters depends, among other factors, on the presence of fitting structures in the environment and that the filtering is ultimately learned individually or evolutionarily from stimuli produced by such structures, as shown for the former by visual deprivation studies (e.g. Hirsh & Spinelli, 1970). Many and diverse descriptions of perceptual functioning as filtering mechanisms exist (e.g., for computational descriptions in visual perception, see Adelson & Bergen, 1985 for work on motion, Jones & Malik, 1992 for stereo vision, or Malik & Perona, 1990, for texture discrimination). The more classical perspectives on selective attention as filters in perception and cognition have been provided by Broadbent (1958), Treisman (1964a), and others. I should like to point out, that the discussion of perceptual filtering mechanisms on this and the next few pages focuses on structural issues of perception and on structurally-induced filters rather than on the relation between attention and filtering. We will turn to attentional aspects of filter theories in much more detail later on (see Sect. 4.1).

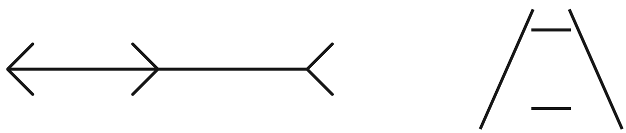

Figure 2.2: Constructing realities I: examples of visual illusions with distortions of magnitude. *Left:* A Müller-Lyer illusion (after Müller-Lyer, 1889). Which of the two parts of the horizontal line is longer? Although both parts are actually of the same lengths the right-hand part appears to be longer. *Right:* A Ponzo illusion (after Ponzo, 1913). The two horizontal lines have the same lengths, however, the upper line appears to be longer.

On the other hand, there is much evidence that effects of perceptual mechanisms go well beyond what can be caused by a passive selection and structuring of information. Visual illusions are good examples of systematic distortions that are introduced by mental processing on perceptual and cognitive levels, resulting in mental representations of the (structures of the) world 'outside' that are not necessarily veridical (e.g., in the sense of being metrically correct). In the following, we will use some examples of visual illusions to make a more general argument about how mental processes exploit information about external structures for their functioning and about how they, ultimately, come to rely on the structures' presence.