Abstract

This thesis presents a cognitive science contribution to the investigation of mental processing of knowledge about geographic spaces. In cognitive science, mental representations of spatial knowledge are metaphorically referred to as cognitive maps. However, investigations in cognitive psychology reveal that the cognitive map metaphor is inadequate, which makes necessary a more suitable conception of human geographic knowledge processing. In the present work, the issue is addressed from an artificial intelligence (AI) perspective. An experimental computational modeling approach for mental processing is presented. Results about human memory and visual mental imagery from cognitive psychology are combined with AI techniques of spatial and diagrammatic knowledge processing. The diagrammatic reasoning architecture MIRAGE is developed as a comprehensive conception of human geographic knowledge processing.

The work is based on the theses that (1) geographic knowledge representations in the human mind are constructed on demand, that (2) this construction is based on underdetermined knowledge from long-term memory, that (3) this knowledge is stored in a fragmentary, hierarchically structured form, and that (4) the resulting representation in working memory is a visual mental image.

MIRAGE is structured according to the psychological distinction of human memory in long-term memory and working memory. It uses topological, orientation, and shape information stored in spatial knowledge fragments in long-term memory. The construction of visual mental images in working memory is described starting from the retrieval of pieces of knowledge in long-term memory. A working memory representation is constructed on the basis of retrieved spatial knowledge fragments. Missing spatial information is complemented by default knowledge. The representation built up in working memory is used to construct a visual mental image in a visual buffer. This image is inspected to yield a spatial relation. Complex image construction strategies are developed that provide solutions to the problem of representing underdetermined spatial information in a quasi-pictorial representation.

The basic functionality of MIRAGE has been realized in a prototypical implementation to demonstrate the dynamic behavior of the model.

Through the integration of psychological results with AI techniques of visuospatial information processing in a common modeling conception, MIRAGE provides an essential contribution to the investigation of human spatial knowledge processing. This modeling conception forms the basis for ongoing discussions, for empirical investigations, and for future AI projects.

Acknowledgments

First of all I would like to thank my advisor Christian Freksa for his critical, imaginative, and friendly support. This work has been shaped in innumerable discussions with him. His comprehensive assistance helped me identify my research goals and find a way through the often pathless field of interdisciplinary research.

Special thanks also to Mary Hegarty for her constructive criticism and many valuable psychological clues. Likewise, I want to thank Stephanie Kelter who guided my basic psychological studies and accompanied my work in its early phases. I thank Dan Montello for commenting on intermediate stages of my work and for pointing to other forms of geographic knowledge than I focus on in my work.

This work has been done within the research project Spatial Structures in Aspect Maps (Fr 806/8, Spatial Cognition priority program, DFG). I thank the DFG for funding the project and for establishing the Spatial Cognition priority program, which provided a fruitful framework for spatial cognition research.

I thank the members of the Knowledge and Language Processing Group (WSV) and of the Cognitive Science Program (GrKK) in Hamburg for the pleasant work climate, for manifold discussions and constructive criticisms. Special thanks are to Ladina Tschander, Markus Guhe, Lars Kulik, Alex Klippel, Sven Bertel, Steffen Egner, Simone Pribbenow, Thora Tenbrink, and Heike Tappe, who provided valuable input to my work, helped me identify what I actually wanted to investigate, and provided essential emotional support.

I am also indebted to Christopher Habel, Markus Knauff, Paul Lee, Christoph Schlieder, Sabine Timpf, and Werner Kuhn for ideas, comments, and discussions.

Finally, I particularly thank Gabi Weise for her patience, for her loving support, and for being herself.

Table of Contents

List of Figures

List of Tables

1 Introduction

Consider the following questions:

- What is the shape of France?
- Where is the city of Paris located in France?
- Which boundaries of France are formed by coast lines?
- If you fly from Paris to London, which direction will you go?
- Which city is farther away from Paris: London or Madrid?

When we refer to our individual knowledge about geography we are able to recall an abundance of spatial information that enables us to envisage how things are like in the world. We can conceive of locations of geographic entities, we can decide about the relative position of one place with respect to another, or we can think of how a region is shaped and how it is related to its spatial vicinity.

The type of knowledge involved in answering the above questions is called *geographic knowledge*. The attempt to answer these questions requires *mental processing*, i.e., a certain cognitive effort is needed to find an appropriate answer.

1.1 Mental Processing of Geographic Knowledge

Geographic knowledge is knowledge about geographic spaces (Montello, 1993). Geographic spaces are much larger than the human body. Examples of geographic spaces are countries, states, or continents. Knowledge about geographic spaces is contrasted to knowledge about environmental spaces. Unlike knowledge about environmental spaces, geographic knowledge typically cannot be acquired by direct experience in a spatial environment. Instead, geographic knowledge is acquired using secondary sources of information like verbal descriptions or pictorial representations, for example maps (cf. Montello & Freundschuh, 1995).

1.1.1 Cognitive Maps

Spatial knowledge in mind is usually referred to metaphorically as *cognitive maps* in psychology, anthropology, and geography (Tolman, 1948; Downs & Stea,

1977; Hirtle & Heidorn, 1993). Initially meant as an analogy to external map-like representations, the metaphor became more and more understood in a figurative sense. The map metaphor taken literally suggests that mental representations of spatial environments are spatially coherent, picture-like internal representations which preserve spatial relationships between represented entities homogeneously both in scale and in resolution, and which can be inspected by mental processes analogous to visual perception processes on real external maps.[1]

Numerous empirical investigations in cognitive psychology have revealed that the map metaphor for mental representations of spatial knowledge must not be interpreted in a literal sense. Instead of being coherent, veridical, and complete, mental representations of spatial knowledge must be conceived as fragmentary, distorted, and incomplete (cf. Tversky, 1993).

To give a few examples,

- mental representations of distances and orientations between geographic locations show systematic distortions with respect to their actual values;
- shapes of geographic objects and angles between linear features are mentally modified to fit more ideal forms (i.e., lines are straightened, angles are idealized towards right angles);
- objects are displaced and rotated to form more schematic configurations;
- spatial aspects that form symmetric relations in the environment (e.g., the distance between objects) often result in asymmetric mental representations (i.e., the distance from a location A to another location B in its mental representation may systematically deviate from the distance from B to A);
- mental representations of spatial knowledge do not form a single homogeneous structure but are organized in a hierarchical manner (for instance, the spatial relationship between a building in a city A and a building in a city B is given by the spatial relation between the two cities).

Altogether, the findings about cognitive maps suggest that mental representations of geographic knowledge are not stored in a ready-made form. Rather, mental operations involving visual and spatial knowledge must be conceived as *construction processes*:

> "The view of memory for the visual world that the data seem to favor is a constructionist view, that representations of the visual world are constructed, and that systematic errors may be introduced in the construction of representations as well as in retrieval of information from them" (Tversky, 1992: 135; cf. Portugali, 1996a).[2]

---

[1] Today, the term 'cognitive map' is mostly used in a very general sense, just meaning 'mental representation of spatial knowledge' (Hirtle, 1998; see Section 2.1).

[2] Cognitive processes often construct mental representations that are partially based on facts that are known and partially based on inferred information. For example, trying to remember something that has been learnt often involves inference processes. These inference processes use facts that have been learnt and complement them by details that are plausible but not necessarily true (e.g. Bransford et al., 1972; Sulin & Dooling, 1974). So decisions are often based on facts that are likely to be true due to their plausibility

When we think of geographic facts represented in our minds we usually do not have the impression that our knowledge is deficient. The characteristic of our spatial knowledge as resulting from construction processes suggests that we are able to envisage many more pieces of spatial knowledge than we have explicitly represented in the mind. We can use knowledge that actually is stored in memory to construct mental representations that help us envisage the tentative spatial situation we are currently interested in.[3] I will illustrate this characteristic through the following example.

1.1.2 Mental Construction of Spatial Knowledge: An Example

In an empirical investigation of mental representations of spatial knowledge, Stevens and Coupe (1978) asked people to decide about the relative spatial orientation between pairs of well-known geographic locations. More specifically, the participants in this experiment were asked to decide about the cardinal direction of one city with respect to the other. For example, they were asked to indicate graphically the cardinal direction of San Diego (California) with respect to Reno (Nevada). Although the participants (undergraduate students of the University of California, San Diego) were familiar with the geography of the western United States they presumably had never explicitly elaborated on this question before.

Nevertheless, the participants were not uncomfortable answering the question and most of them indicated San Diego (California) to be farther west than Reno (Nevada). In fact, however, Reno is located farther west than San Diego (see Fig. 1.1). For my purposes it is interesting that the participants in this experiment did not know the requested answer explicitly from memory (otherwise they should have known the correct answer). Apparently, they used available pieces of knowledge stored in their minds (the relative orientation between the respective states) to envisage the spatial layout of the area under consideration. They grounded their decision on a mental construction, which they produced for answering the question they were asked. Thus, with respect to the question they were asked, their knowledge can be regarded as being *underdetermined*; by using their available knowledge of the possible geographic layout they were able to produce a mental representation to ground their decision on.

Stevens and Coupe (1978) explained their results by means of the hierarchical organization of long-term memory. The experiment revealed that the participants did not reproduce a mentally stored orientation between the two locations but instead derived it, presumably utilizing the relative orientation between the two states (California and Nevada) the cities are located in. The relative position of the two states is Nevada being farther east than California. It seems plausible that the

---

rather than on real knowledge. Also in visual object recognition, visual constructions in mind play a paramount role; this will be further reported in Section 2.3.

[3] I will refer to this tentative spatial situation in the world as *spatial configuration*; so spatial configuration means the mental construction of what things are envisaged to be like in the world.

participants in the experiment did not know the relative positions of the two cities explicitly but that they used their knowledge of the relative positions of their containing states to decide for San Diego being farther west than Reno. Questions like "What is farther north: Madrid (Spain) or Washington (DC)?" or "... Seattle (USA) or Montreal (Canada)?" revealed comparable results. However, in the present context I will not elaborate on the hierarchical structure of spatial memory but on the constructive aspect of spatial knowledge.

Fig. 1.1. Map of California and Nevada exhibiting the relative positions of the cities of San Diego and Reno

The example given above shows that answers constructed on the basis of underdetermined knowledge can be misleading. This is due to the fact that a construction based on underdetermined knowledge made up of available pieces of knowledge only results in a *possible* spatial configuration; this configuration is not necessarily the only possible solution and therefore may be wrong. During the construction process, missing pieces of knowledge have to be substituted to overcome the underdeterminacy of the spatial knowledge used for the construction.

I will assume that for most tasks spatial reasoning about geographic spaces involves underdetermined knowledge and that the performed construction usually leads to satisfactory results. Clearly, an answer built on the basis of underdetermined knowledge can be the better the more pieces of information can specify the result and can be employed for the construction process to answer a given question.

1.2 Theses and Assumptions

The core theses of this work are the following:

1. Geographic knowledge representations in the mind result from *construction processes*. These representations are constructed on demand to elaborate on a certain spatial question.
2. The mental constructions are based on *underdetermined* geographic knowledge. Typically, not all pieces of knowledge needed for the construction task are available from memory.
3. The knowledge used for the construction is represented in long-term memory in a *fragmented* and *hierarchical* form. There is no ready-made, coherent representation of geographic knowledge. Rather, partial facts are structured in hierarchies and are retrieved prior to the construction process.
4. For the construction, a spatio-analogical representation format is used in working memory. A *mental image* representation is constructed and inspected to provide an answer to a current geographic question.

I will explain my theses in the following sections.

1.2.1 Knowledge Construction and Human Memory

As motivated above, geographic knowledge required in specific situations often cannot be retrieved from memory in a ready-made form. Instead, an appropriate mental representation must be constructed on the basis of pieces of knowledge available from memory when needed. Cognitive psychology distinguishes long-term memory and working memory depending on whether it holds information over an extended period of time or whether it represents information just for a limited period, respectively.[4] With respect to the mental phenomena described in this thesis, the geographic knowledge is represented in long-term memory, whereas the construction of a specific geographic representation is performed in working memory. To answer geographic questions the working memory representation is evaluated (cf. Fig. 1.2).

1.2.2 Characteristics of Geographic Knowledge

I have stated above that geographic knowledge is acquired by secondary sources of spatial knowledge like verbal descriptions or map-like representations. When considering how geographic knowledge is acquired from a secondary information source (say, from a geographic map) it becomes evident that we usually cannot acquire a complete and accurate representation of a given environment.

---

[4] Psychological theories of human memory will be reported in Section 2.2. For the time being, I will distinguish between long-term memory representations (which contain the geographic knowledge available in mind) and working memory that operates on the knowledge taken from long-term memory. Working memory is employed in building up the required representation needed to answer a given question.

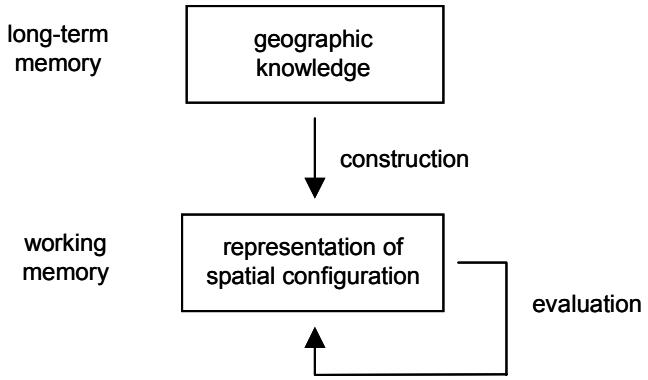

Fig. 1.2. Construction of working memory representation based on geographic knowledge from long-term memory

Two ways of gaining complete mental representations using a map are conceivable: a human map interpreter may extract and learn spatial relations encoded in the map explicitly, or he may try to memorize the map as a whole, i.e., in the form of a picture-like representation (like a mental photograph) for later use. In the first case every spatial information of potential interest would be mentally available for immediate recall, whereas in the second case the mental representation of the map could be used to extract the required relation upon request. However, both ways are not feasible:

1. Regarding the abundance of spatial relations that exist in a geographic map, it becomes evident that a map user acquiring geographic knowledge from a given map will not be able to mentally store every relation that can be read off the map.[5] That means that it cannot be assumed that every spatial relation in the map will be represented explicitly in mind.
2. On the other hand, it is not feasible to mentally store the map as a whole, either (like a photograph in the mind). Although this would mean that a mental representation equivalent to the external map containing all spatial knowledge implicitly for later use was available, the memorization of the whole map will fail because of the map's complexity.

Due to the representational requirements of human memory, a more efficient form of organization is necessary in mind. When geographic knowledge is acquired, specific spatial features are selected for mental representation when they seem most relevant for later use or when they are most salient in the acquisition process, whereas information of minor importance or salience is just ignored. Aspects of

[5] If we consider a single spatial aspect that may be represented in the form of a binary relation, for example spatial orientation between two locations, there are $\frac{1}{2}(n^2-n)$ relations, where n is the number of spatial locations in the map. So if the map contains 10 objects, this would result in 45 direction relations to be memorized; for 100 entities, it would give a total of 4950 relations, just to represent the orientation relations between any two objects!

minor importance are abstracted from, and the mental representation concentrates on what is most likely to be important for later use. As a consequence of this information selection and abstraction, the mental representation of a geographic environment usually is *incomplete* compared to what potentially might be represented about a geographic environment. The mental representation from a cognitive psychology perspective is viewed as a mapping of selective aspects of the real, or represented, world (in the example conveyed by the information contained in the map) to the internal, or representing world in the mind (Palmer, 1978; Rumelhart & Norman, 1985). I will refer to the characteristic of geographic knowledge of being incomplete as *scarce knowledge*.

Besides selecting information for being represented in mind, thus resulting in incomplete representations, mental representations also tend to have a *coarse* or *qualitative* characteristic. This characteristic means that usually there are no exact metric values of spatial relations between geographic entities represented in mind. Rather, mental representations tend to concentrate on semantically relevant distinctions between possible cases. Knowledge is usually represented in a schematic form that employs classes of relations rather than exact measurements. For example, we may have knowledge about the rough shape of France rather than about the exact course of its boundaries; we may know that Paris is centered in the northern section of France rather than knowing its exact geographic coordinates; or we may be able to say that London is located north-westerly of Paris, without being able to specify an exact angle.

Both characteristics of geographic knowledge, the property of being coarse together with the property of being scarce, may cause geographic knowledge to be *underdetermined* with respect to a piece of knowledge that is required in a given situation. I will assume that in specific situations most of the geographic knowledge in the mind can be of this type: although we may have acquired exact values for some spatial relations (e.g., distances between places; cf. Montello, 1998) or detailed knowledge about specific geographic environments, we encode most of our spatial knowledge in mind as coarse and scarce knowledge. However, the characteristic of being underdetermined also has advantages. It points to an efficient form of representation of geographic knowledge in the mind. In our mental representations we operate with spatial knowledge that enables us to envisage more geographic detail than is actually stored in memory.[6] Therefore I will refer to geographic knowledge being scarce and coarse as *lean geographic knowledge*. The leanness property allows for processing geographic knowledge on the basis of coarse and incomplete pieces of information.

[6] This feature of spatial knowledge processing in the mind is related to the psychological concept of *cognitive economy* (Collins & Quillian, 1969). According to the cognitive economy assumption, knowledge is stored in semantic hierarchies. Semantic properties that apply to many items are stored in superordinate locations within these hierarchies. Thus, redundancies in mental representations are avoided.

1.2.3 Spatial Knowledge Organization in Long-Term Memory

As discussed in Section 1.1.1, there is no single coherent representation structure for geographic knowledge in the mind. The cognitive map metaphor in the naive sense, i.e., as a strong analogy to external geographic maps, cannot provide an appropriate conception for mental representations of geographic knowledge. Instead of coherent representations in long-term memory we rather have to assume that our knowledge consists of fragmentary pieces of information. This fragmentary character allows for using the available lean knowledge in varying contexts in a flexible way.

I will assume *spatial knowledge fragments* as minimal units of lean knowledge stored in long-term memory. A spatial knowledge fragment contains a geographic property that holds for an entity or a geographic relation between two or more geographic objects.

An important question is, how the spatial knowledge fragments are organized with respect to each other: when we assume that spatial knowledge is fragmented into pieces, we must decide for an organizational structure between the pieces to allow for accessing them. As we know from empirical investigations in cognitive psychology, mental representations of spatial knowledge are often organized hierarchically (see for example the experiment by Stevens and Coupe (1978) reported in Section 1.1.2). I will assume that spatial knowledge fragments are organized hierarchically in long-term memory. This hierarchical structure is important for accessing the spatial knowledge fragments when they are retrieved from memory. The order in which fragments are retrieved depends (among other factors) on how they are structured through their hierarchical organization; the order of this retrieval, in turn, influences the way in which the working memory representation can be constructed.

Spatial knowledge fragments are used to construct working memory representations of geographic configurations, which I assume to be in the form of *mental images*. This conception will be introduced next.

1.2.4 Visual Mental Images and Diagrammatic Reasoning

When a working memory representation of a geographic configuration is constructed from spatial knowledge fragments from long-term memory, several knowledge fragments are combined in a single representation structure. Although the representation in mind is not map-like in a literal sense, I assume that the constructed mental representation is in a spatio-analogical format. The theoretical framework for this spatio-analogical working memory representation is given by the conception of *visual mental imagery*. *Visual mental images* (or just *mental images*) are spatio-analogical representations in working memory constructed from knowledge in long-term memory. Mental images are inspected by mental processes that compute the requested result.

There is neuropsychological evidence that the same neural systems are involved in mental reasoning about spatial configurations as in the visual comprehension of external scenes (Kosslyn, 1987; Kosslyn & Shin, 1994). For

example, in thinking about geographic configurations the same neural structures are involved as in studying external geographic maps. So mental imagery is conceived as the internal (mental) counterpart of visually perceiving and mentally processing of external objects (cf. Fig. 1.3).[7] Besides in processing genuine visual content, mental images are also involved in understanding written or spoken texts, when spatial states of affairs are mentally represented and processed. These spatio-analogical *situation models* are built up in memory to represent the situational content (not the verbatim text structure) of the text that has been understood (e.g., Tversky et al., 1994; Kaup et al., 1999).

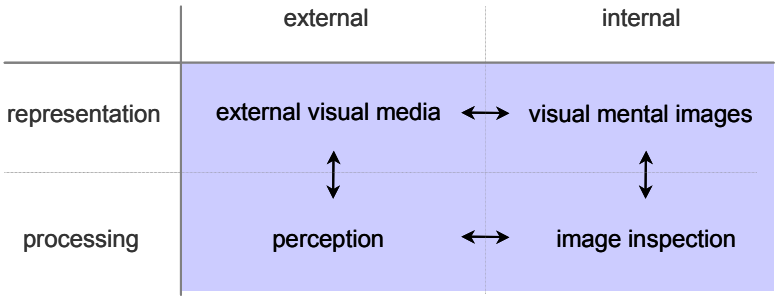

Fig. 1.3. The relationship between external visual media (e.g. maps) and (internal) visual mental images. Neuropsychological results reveal that the same neural systems are involved in perceiving external and in processing internal image structures

Processing geographic information in the mind by constructing visual mental images on the basis of knowledge retrieved from long-term memory is an efficient way of operating with geographic knowledge. The following features of mental image construction in reasoning about geographic configurations can be identified:

- Mental images are customized representations of geographic configurations. They are constructed to meet the specific requirements of the task to be solved in a given context. For example, they focus on the relevant geographic entities, they can be constructed at an adequate scale and resolution, or they may contain exactly those spatial properties and relations that are of interest at the required level of detail.
- Lean geographic knowledge as represented in long-term memory can be used in mental images in a flexible manner. Spatial knowledge fragments can be successively integrated into the mental image in working memory. Since human working memory is restricted in capacity it is necessary to select the appropri-

[7] However, it is also possible to perform purely propositional reasoning processes on the basis of external visual media: specific spatial relations can be extracted from the external representation and can then be used to mentally compute further spatial relations without employing visual mental images.

ate pieces of knowledge to be integrated in the image. As a consequence of the successiveness of the image construction process the working memory representation can be used already in early stages during the construction process. So when a fast albeit coarse answer to a spatial question is required, an image inspection can be performed before the image is completely constructed. This *anytime characteristic* accounts for the fact that cognitive systems often act under resource restrictions. Time constraints, for instance, can force fast decisions that are based on sub-optimal solutions of a problem. The successive construction of mental images allows for their modification and refinement to obtain successively more accurate results using *image modification* processes (see Section 2.3).

- The main purpose of the working memory representation is to convert lean geographic knowledge from long-term memory into an image that is specific enough to answer a given question. For this purpose it is usually necessary to compensate for missing information not contained in long-term memory. So default knowledge is used to complement coarse and scarce knowledge available from long-term memory during the construction process. Default knowledge is general knowledge that is available from other situations; it is not reliable with respect to the specific given situation. As a consequence of the image construction successively performed and the employment of default knowledge, at any stage the working memory representation consists of reliable information from long-term memory and unreliable default components supplied for the construction process. When the image construction process continues, the amount of default knowledge may be reduced when it can be substituted by more reliable pieces of knowledge retrieved from long-term memory.
- In mental images, geographic knowledge from different information sources and of different modalities is combined in a common representation format. In artificial intelligence and in cognitive psychology it is often distinguished between *propositional* and *pictorial* types of knowledge (Paivio, 1971; Larkin & Simon, 1987). Both types of knowledge are employed in mental image constructions. By combining these two forms of representation in the common representation format of the mental image, the mutual advantages of both representational types can be exploited (Freksa et al., 1999).[8]
- Both types, propositional and pictorial representations, can be used in mental images to *make knowledge explicit* that is only implicitly contained in the long-term memory representation. So by constructing the mental image representation, spatial relations that derive from explicitly represented relationships become evident, and pictorially represented spatial properties of geographic entities underlying the image construction can be made explorable with respect to other geographic entities. This characteristic of image-like representations to

[8] However, it is also conceivable to combine mental image representations with independent propositional representations; i.e., propositional information are not integrated in the mental image, but are used to augment the image in the sense of a hybrid mental representation.

explicitly exhibit spatial information implicitly contained in a given body of knowledge is related to the core idea of diagrammatic reasoning (DR) in artificial intelligence (Koedinger, 1992; Glasgow et al., 1995).

In this thesis, spatial and diagrammatic reasoning techniques will be used to model the construction of visual mental images in human working memory.

In diagrammatic reasoning spatio-analogical representation structures are used to support the reasoning processes through the medium's spatial structure. The spatial structure of the medium can be used to restrict the relationships between the entities that are represented in this medium. As a consequence, inference processes may be reduced to mapping a spatial scenario into the diagrammatic representation medium and to reading off the resulting configuration.

With respect to the task of modeling the construction of geographic knowledge representations in working memory, this reasoning process is performed in the visual mental image representation. The mental image is constructed partially by knowledge explicitly retrieved from long-term memory and partially by default knowledge that complements information not represented in long-term memory. Finally, the desired spatial relation is read off the image by image inspection processes.

Depending on how many pieces of knowledge suitable for the construction of a mental image for answering a given question are available from long-term memory, a variety of cases can be distinguished under the perspective of the mental image construction process (cf. Fig. 1.4):

1. When the knowledge required to answer a given question is explicitly represented in memory, it simply can be retrieved. With respect to the above example (cf. Section 1.1.2), this trivial case corresponds to explicitly knowing the relative location of Reno with respect to San Diego, because, for example, the information has been investigated earlier and can be recalled.
2. If the required knowledge is not explicitly represented in long-term memory, an attempt is made to construct the answer in working memory. Depending on the pieces of knowledge available in long-term memory, three alternative situations are conceivable:
   - 2a. First, the pieces of knowledge stored in long-term memory may be sufficient to perform the mental image construction; in this case the mental image can be constructed and the answer can be read off the image.
   - 2b. Second, in the case of lean knowledge in long-term memory, default knowledge is employed to construct the mental image.
   - 2c. Third, it can be the case that pieces of knowledge retrieved from long-term memory are in conflict with each other; in this case, a mental image construction may fail unless the conflicts can be resolved.

For the purposes of this thesis, the last two cases 2b and 2c are the most interesting ones. With respect to diagrammatic reasoning they entail the questions how to visualize spatial facts that are underdetermined in the sense that they do not enforce a unique visualization, and how to deal with spatial facts that (seem to) have no diagrammatic counterpart at all. These problems are further investigated in Chapter 3 and Chapter 4, respectively.

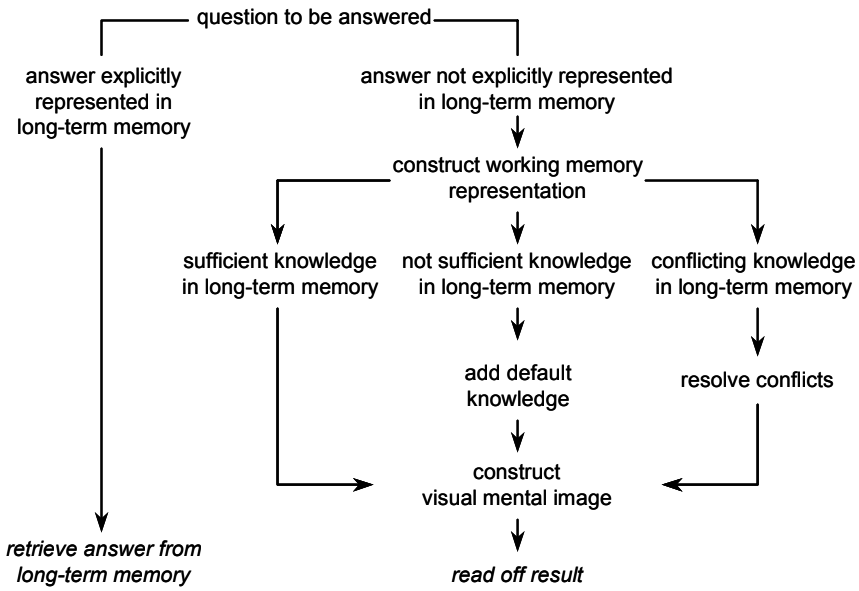

Fig. 1.4. Possible cases in the construction of geographic knowledge representations in working memory depending on the availability of suitable knowledge from long-term memory

1.3 Research Questions and Goals

So far, I have motivated that geographic knowledge represented in mind is employed in constructing mental representations to envisage a configuration of geographic entities in the world. In this thesis I will elaborate on the question how this mental construction process can be explained from an artificial intelligence point of view.

1.3.1 Research Questions

The research questions pursued address representation structures, processes operating on these structures, and the control of the system's overall behavior:

1. How can we describe the representation structures holding both the underlying geographic knowledge represented in mind and the mentally constructed geographic knowledge representation produced on the basis of the underlying knowledge?
2. What kinds of processes operate on these representations, i.e., which processes access the mental knowledge, and how can we understand the processes that

construct mental representations of geographic knowledge to envisage a geographic configuration?

3. How can we model the dynamic behavior of the system, i.e., how can we describe the control structures that drive the processes mentioned above to achieve the intended results?

These three central questions entail a number of sub-questions:

- How is the geographic knowledge in memory characterized that is used for constructing mental representations?
- How is this knowledge organized and accessed?
- How is it used to envisage a geographic configuration?
- How does the construction of the geographic configuration proceed?
- What kinds of knowledge representation structures serve the intended purposes of the construction process?
- How is missing knowledge dealt with in the mental construction process?

1.3.2 Goals

The main goal of this project is to build a *computational model* [9] that accounts for the construction of geographic knowledge representations in the mind on the basis of lean knowledge. For this purpose, the phenomena and principles of human geographic knowledge processing will be analyzed and described.

The computational model will integrate facts acquired from empirical studies, from metaphorical conceptions about human spatial knowledge processing, as well as from existing models, both conceptual and implemented. These facts will be implemented in a model that exhibits the characteristics of its components as well as the dynamic interaction between its parts. So the thesis aims at integrating phenomena in a processing model which provides a generalized conception of ideas and findings hitherto regarded only in isolation.

The resulting structure and its dynamic behavior will be evaluated with respect to the preconditions underlying the model's development. This reflection on the modeling process and its results will reveal further issues to be investigated and will generate new questions for further research.

The work reported in this thesis is a contribution from artificial intelligence (AI) to answering questions on mental processing of geographic knowledge. AI methods are employed in a system that helps explain human intelligence in terms of computational structures and processes. This thesis will also provide further ideas for the application of cognitive principles in technical AI systems.

The work reported here can be of interest to cognitive scientists investigating mental spatial knowledge processing. The computational model will raise questions regarding mental operations based on lean spatial knowledge that can be investigated empirically by cognitive psychologists. The results of empirical studies can help refine the proposed model.

---

[9] The approach of experimental computational modeling will be elaborated in Section 1.4.

Computer scientists are interested in employing cognitive principles of spatial knowledge processing in technical AI systems. For example, the communication with and control of autonomous agents navigating in real or virtual environments may require a cognitively adequate system for dealing with scarce and coarse pieces of spatial information. These pieces of information may stem from input of the system's sensors or from instructions given by a human instructor in natural language or in other qualitative forms like diagrammatic representations (Freksa et al., 2000a; b).

Geographic information scientists are interested in human strategies for envisaging geographic configurations for the application in geographic information systems (GISs). The development of geographic information systems that operate in such a way that they are compatible with human ways of dealing with geographic knowledge is a core goal in cognitive GIS research: "To be effective, GIS will need interfaces compatible with the way human beings absorb, represent, and use spatial information" (Taylor & Tversky, 1995: 235). But not only with respect to the design of user interfaces for GISs human processing of geographic knowledge is interesting: "We believe that several interesting and important questions for GIS research and design are related to issues of human perception and cognition of space and spatial information" (Montello & Freundschuh, 1995: 169-70; cf. Freksa & Barkowsky, 1996).

1.4 Approach: Experimental Computational Modeling

The approach taken in this thesis is *experimental computational modeling* of cognitive structures and processes. It is based on the general assumption that cognitive phenomena can be investigated, explained, and reconstructed as computational processes operating on discrete data structures. To employ this approach, the phenomena to be examined have to be conceptually analyzed to a level of detail that allows for defining data structures and processes that can be described in a programming system and implemented on a digital computer.

1.4.1 Computational Cognition

Since the late sixties of the twentieth century the dominating paradigm for investigating cognitive phenomena has been based on the assumption that cognition is *information processing*. This *information processing approach* has replaced the behavioristic research paradigm, which was predominant in psychology in the first half of the century. One of the leading theories in the information processing approach as the basis of intelligent behavior is the *physical symbol system hypothesis* (Newell & Simon, 1972; Newell, 1980).[10]

[10] The other major approach is given by the connectionist paradigm. In contrast to manipulating symbols as basic entities of cognitive processes, connectionists proceed on the assumption that intelligent behavior emerges from the interaction of highly interconnected simple (neuron-like) processors in neural networks.

The physical symbol system hypothesis claims that cognitive processes can be conceived as transformations of formal symbol structures. Symbol structures are composed from elementary symbols according to syntactic rules. Since these elementary symbols correspond to entities in the world, they form the semantic basis of a physical symbol system. Intelligent behavior is assumed to result from generating, manipulating, and evaluating symbol structures (Posner, 1989; Osherson, 1990; Strube et al., 1996).

Whereas behavioristic approaches aimed at explaining intelligent behavior merely on the basis of stimulus-response relationships, the central idea in the symbol processing paradigm is that knowledge is *represented* within the cognitive system. *Knowledge representations* – formed by physical symbol structures – are operated on by cognitive processes. Interestingly, one of the pioneering results that forced giving up the behavioristic paradigm was gained by investigating spatial orientation skills in rats: the experiments by Tolman (1948) lead to the representational conception of the *cognitive map* as indispensable precondition for performing spatial tasks.

The research paradigm of symbol processing has evolved in tight correlation with the development of digital computers. Digital computers seemed to provide a promising metaphor for information processing in the mind. On the other hand, the new perspective on intelligent behavior promoted ample interest in computer programs that perform tasks that require intelligence when performed by natural cognitive systems. However, recognizing the complexity of the investigation of intelligent behavior lead to the identification of *cognitive science* as an interdisciplinary research program.

The central idea of *cognitive science* is that the cooperating disciplines computing science, cognitive psychology, linguistics, and philosophy (together with others like geography, architecture, or semiotics) each provide valid methods for investigating the phenomena of cognition. By uniting all of them in a common research paradigm, shortcomings of one discipline may be overcome by others. With respect to this perspective the method of *computational modeling* can be considered one of the main techniques for bringing together methods from different disciplines.

Besides providing a qualified means for conveying ideas and their solutions to other researchers, building computational models reveals a couple of significant advantages in contrast to other scientific techniques, like purely verbally or mathematically described models. First of all, as being based on a computational algorithm, there is one unique way in which the model can be interpreted, both by the computer and by human researchers that use the model in scientific discourse. This unique way of interpretation provides a clear description of the issue to be investigated and helps avoid misunderstandings.

Second, as a result of the need to provide an algorithm for the computational model, computational models reveal shortcomings in any theory as conceptional gaps, missing components, or yet unconsidered cases that manifest in the implementation and become obvious in the running model. As such, computational

models can serve as a motor for further iteration cycles between the disciplines cooperating in cognitive science.

Third, the fact that the dynamic behavior of a computational model running on a digital computer can be observed and assessed is a major advantage compared to models that are only statically described. In cognitive science models, dynamically operating natural systems are described; the models can be observed in their dynamic behavior, and they can be assessed with respect to what they intend to describe.

1.4.2 Building Computational Models

The method of computational modeling has become widely accepted within cognitive science as it allows for the interdisciplinary interaction between the cooperating disciplines. In particular, computational modeling can be used for a cyclic interaction between the scientific disciplines. As sources for the construction of a cognitive model can be identified (see Fig. 1.5):

- empirical findings from cognitive psychology;
- theoretical considerations, e.g., from computer science and philosophy;
- results gained from preceding computational models; and
- intuitions, provided for example by metaphorical conceptions.

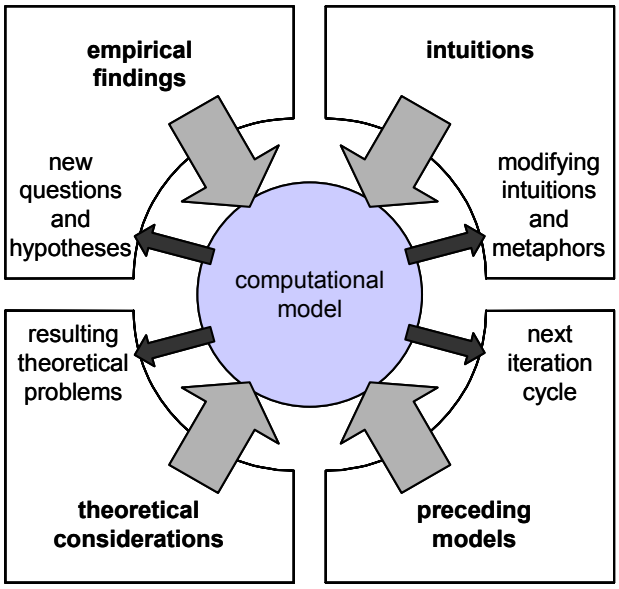

Fig. 1.5. Sources for the construction of computational models

Conversely, implemented computational models can have an impact on further research in each of their contributing disciplines:

- *Cognitive psychology* examines human or animal behavior by performing *empirical investigations*. Based on hypotheses about the question to be pursued, data is collected and evaluated to verify or refute the respective hypotheses. The results can be used for building a computational model, which can be tested and modified to fit the empirical observations. Several empirically investigated details can be integrated within a single model, which may reveal conceptual gaps in the underlying assumptions. The operation of the implemented computational model in turn can be used to stimulate further questions and hypotheses to be investigated empirically by cognitive psychology in a subsequent iteration phase.
- As implemented computational models are realized within the field of cognitively motivated *artificial intelligence*, they rely on *theoretical results* from computer science. These results comprise the characteristics of the modeling instruments, i.e., the data structures and the processes operating on them. The emphasis is on the computational properties of the algorithms and on the question which properties of artificial information structures can be exploited best for obtaining the desired behavior. The attempt of constructing a computational model raises further questions which have to be theoretically tackled in computer science.
- Often, computational models are constructed as an extension of already existing implementations. These implementations can be *preceding modeling attempts* for the same or a related issue, or they can be purely technically motivated implementations that are adopted for the task to be performed. Every computational model, on the other hand can be used for further developments, both in a cognitive modeling or in a technical respect (for example in user interface design).
- Although mentioned last, at the very beginning of a cognitive modeling attempt there is often an *intuitive motivation* inspiring the model. This intuitive motivation may be provided by a *metaphorical conception* about the cognitive phenomenon under consideration. Using a metaphor means that a notion that is well understood is adopted from another scientific or common sense area to conceive of the cognitive phenomenon to be explained by the model (cf. Section 2.1). An implemented model can serve for further refining a metaphorical conception, or it can be transferred to another domain to be investigated.

In this thesis, all four sources of computational models will be considered. Metaphorical conceptions of spatial knowledge representation and processing, and empirical findings from cognitive psychology (especially about processing spatial information in mental images) are most important. Based on the analysis of existing results, an artificial intelligence model will be constructed, implemented, and assessed with respect to the research questions. The issue of mentally reconstructing geographic knowledge in working memory will be addressed from an architect's point of view (cf. Sloman, 1994; Braitenberg, 1984). By analyzing and designing representation and processing schemes for geographic knowledge, an

artificial system will be synthesized. This artificial system can be used as a computer simulation for empirically studying geographic knowledge processing.

For this purpose a *metadescription* of the model to be designed will be provided:

> "In order for a computer program to serve as an embodiment of a theory, we must have a 'metadescription' of the program. This description should state the important principles of the theory and describe the corresponding features of the program" (Kosslyn, 1980: 138).

The purpose of this metadescription is to bridge the gap between the underlying theoretical assumptions and the model itself. Since the model embodies the theories it is based on, I will first document the theoretical principles and define their correspondence to the model's components. This method enables to distinguish between parts or aspects of the model that are genuine modeling components and those which are needed to assemble the model as a whole.

1.4.3 Modeling as Experimental Approach

I call the approach *experimental* computational modeling as the observation of the running model allows for experimenting under varying conditions. These experimental facilities enable critical reflections about the model's preconditions as well as about the computer implementation in a similar way as experimenting with human participants. Alternative design decisions can be tested to elaborate, to extend, and to refine the model. In comparison to an exclusively theoretical explanation of cognitive phenomena, computational modeling provides a concrete realization of a dynamic system. As mentioned above, this method forces to completely specify every component of the model up to the degree necessary for computer implementation.

This advantage of computational modeling of being very specific can be a severe drawback. All aspects of the model have to be described in detail for the sake of the model's completeness, even those aspects which are not of central modeling interest. Consequently, with respect to the cognitive adequacy of the model, it has to be specified which components of the model cover literal modeling aspects, and which are built for reasons of completing the model as a whole.

Besides this problem, a computational model is always open to criticism regarding the modeling decisions taken by the designer. A computational model that is determined by data structures and computational processes allows for many degrees of freedom regarding the specification of the structures and the processes, as well as their mutual interaction. For example, it has to be decided whether a specific aspect is to be explicitly encoded in a data structure or whether some process generates this aspect on demand from some implicit information. As a consequence, two systems designed in different ways may show the same operational behavior. So from a phenomenological point of view they would be indistinguishable (cf. Anderson, 1978; Pylyshyn, 1990).

Usually, there is no definite reason for particular modeling decisions, as the observed phenomena do not determine the internal structure of a system (Anderson, 1978). Nevertheless, a computational model provides a concrete embodiment of scientific conceptions that formerly existed as a bunch of – frequently disconnected – theoretical descriptions that each accounted for a different phenomenon. So a computational model is a specifically instantiated form of a scientific conception, and it provides a new basis for further discussions and explorations of a cognitive phenomenon.

1.5 Organization of this Thesis

This thesis is structured as follows. Chapter 2 gives an overview of pertinent research in spatial cognition, both from cognitive psychology and artificial intelligence. First, I review psychological results concerning mental representations of spatial knowledge. I use the diverse metaphors and conceptions coined to characterize spatial knowledge representation in the human mind as a guideline. The Sections 2.2 and 2.3 deal with psychological theories of human memory and the model conception of visual mental image processing, respectively. I describe the mutual interactions among the proposed functional subsystems in mental image construction, and present theoretical and implemented modeling approaches. The Sections 2.4 and 2.5 are devoted to the artificial intelligence sub-disciplines of qualitative spatial reasoning (QSR) and diagrammatic reasoning (DR), respectively. The chapter ends with a summary that exhibits the crucial points for this work.

In the third chapter I develop the model MIRAGE[11]. MIRAGE describes geographic knowledge processing in mental images. Pieces of geographic knowledge stored in long-term memory are used for the construction of quasi-pictorial representations (mental images) in working memory. The representations constructed in this way are explored and further refined when necessary. First, the characteristic properties of the model are described in detail. The representational and processing characteristics of the model's components are specified, both for the long-term memory and the working memory subsystems. In Section 3.2 the model is outlined and its overall characteristics are discussed. Section 3.3 defines the object types that represent geographic entities in MIRAGE and the relation that can hold between them. The next Section 3.4 explains the model's components in detail. Their representational characteristics together with the functional relationships between them are developed. The operation of the model is illustrated using the exemplary scenario presented in Section 1.1.2.

Chapter 4 focuses on the construction of visual mental images in more detail. Using a more demanding scenario, in Section 4.1 I demonstrate possible complications in the construction of spatial mental images based on the model developed in Chapter 3. Section 4.2 contrasts perspectives on diagrammatic representation and processing of underdetermined spatial knowledge from an AI and

[11] MIRAGE stands for Mental Images in Reasoning About Geographic Entities.

a cognitive psychology perspective. I elaborate on the consequences for the construction of mental images in Section 4.3. The options for reducing spatial constraints in the representation, of varying the completion with default knowledge, and the interpretation of qualitative spatial relations are explained. In Section 4.4 MIRAGE is refined to allow for more elaborate image construction facilities to overcome the complications dealt with in this chapter. The problem of unstable images caused by spatial conflicts is described first. Then, the image construction strategies of omission of facts, of revision of relational completion, as well as the revision of the image specification are developed.

Chapter 5 describes the prototypical implementation of the MIRAGE model. To demonstrate the dynamic interaction of the model components a subset of MIRAGE's functionality has been implemented in Common Lisp. The implementation is based on the system SIMSIS. SIMSIS has been designed for interpreting and constructing map-like representations of spatial knowledge. First, the computational tools employed in the modeling task are described in Section 5.1. The SIMSIS system is explained together with the theoretical concept of *aspect maps*. Aspect maps form the conceptual basis for the construction and interpretation of map-like spatial knowledge representations applied in this thesis. Section 5.2 documents how MIRAGE has been realized in Common Lisp. I develop and explain the structure of the implemented system. The operation and behavior of the model is documented in Section 5.3 using the exemplary scenario discussed in Chapter 3.

Chapter 6 provides the conclusion of this thesis and an outlook on future work. Section 6.1 summarizes the results of this thesis. In Section 6.2 the results are reflected with respect to the theses and research questions formulated in Chapter 1. The model is discussed in the context of the existing work reported in Chapter 2. In the outlook I identify open issues for further investigations. Possible extensions of MIRAGE are sketched out first. I then discuss questions to be investigated empirically to provide further criteria for the specification of the model. Finally, I point to possible application perspectives of my approach in artificial intelligence.

2 State of the Art

In this chapter I review research in cognitive psychology and in artificial intelligence (AI) that pertains to spatial and pictorial knowledge processing. First, I report the different conceptions that have been proposed for mental representation and processing of spatial knowledge. Second, I report theories about human memory in general and about the conception of visual mental imagery in particular. Third, I review the AI fields of qualitative spatial reasoning (QSR) and diagrammatic reasoning (DR).

2.1 Spatial Knowledge Conceptions: Cognitive Maps and other Metaphors

Spatial knowledge is conceived in a variety of metaphors. I review these metaphors for two reasons: first, metaphors induce theoretical approaches, they help building models, and they guide the development of hypotheses for empirical investigations. It seems sensible to look for the most adequate metaphor when either of these issues is pursued. Second, different metaphors reflect different aspects of mental spatial knowledge representation. On the basis of the *cognitive map* metaphor (Tolman, 1948; Kaplan, 1973; Downs & Stea, 1977), further metaphors have been proposed. Considering further results of empirical investigations, each new metaphor aims at overcoming inadequacies of its predecessors. Reviewing the different metaphors reveals what is actually known about mental spatial knowledge representations.

Scientific developments are often driven by metaphors that provide a new insight into an area of phenomena. Metaphors convey analogies to fields that are better understood and whose understanding is expected to be carried over to the new problem in certain respects. Moreover, as a means for facilitating communication among people, metaphors are used to convey difficult theoretical concepts by pointing to analogous situations that are more easily understood (Kuhn, 1993). Metaphors promote theoretical advances, and they help to develop theoretical models of observable phenomena (Hirtle, 1998). However, metaphors also

may lead to wrong conclusions when aspects that do not apply to the phenomena to be explained are derived from the analogical conception used (Lakoff & Johnson, 1980; Johnson, 1987; Lakoff, 1987).

In research about mental representation and processing of spatial knowledge, the theoretical development has been promoted and limited by cartographic metaphors, first of all by the *cognitive map* metaphor in its different forms, extensions, and caveats (Hirtle, 1998: 264). However, the cognitive map metaphor is still being used and further elaborated for the sake of new empirical findings. Besides the cognitive map metaphor, a bunch of different notions for mental representations of spatial knowledge have been used. Most of them describe very similar aspects of mental spatial knowledge representation; they only differ with respect to the respective researchers' intentions, or they depend on the scientific discipline they stem from (e.g., geography, city design, architecture, psychology, anthropology, etc.).

Among the different metaphors for spatial mental knowledge processing are *spatial images* (Lynch, 1960), *spatial schemata* (Lee, 1968), *environmental images* (Appleyard, 1970), *mental maps* (Gould & White, 1986), *cognitive atlases* (Kuipers, 1982; Hirtle, 1998), *spatial mental models* (Tversky, 1991; Taylor & Tversky, 1992), *cognitive collages* (Tversky, 1993), *inter-representation networks* (Portugali, 1996b), and *geographic information systems* (Hirtle, 1998). Montello and Freundschuh (1995) present a list of common cognitive and spatial terms (both adjectives and nouns) that allow for a combination of more than 200 metaphorical notions for mental spatial knowledge representations.

These attempts to get to more and more appropriate metaphors to describe mental spatial knowledge representation are motivated by two considerations: first, the initial cognitive map metaphor is to be modified gradually to make it compatible with empirical findings of psychological experiments; second, for developing new designs for further empirical investigations, more and more sophisticated conceptual frameworks are needed.

Since the cognitive map metaphor basically supports the idea of mental *representation* as a precondition of computational cognition (cf. Section 1.4.1), and since it allows for modification and extension in a flexible way, it still can be accepted as general research paradigm for the investigation of spatial mental knowledge processing (cf. Kuhn, 1963).

2.1.1 Cognitive Maps

The development of the conception of spatial knowledge as *cognitive maps* in psychology is closely related to the research paradigm of *mental representation* together with the manipulation of discrete symbol systems (cf. Section 1.4.1). Prior to the representational paradigm, i.e. in the behavioristic period in the first half of the twentieth century, the only valid way of investigating spatial skills in humans and in animals was by examining correspondences between perception stimuli and the resulting behavior. The pioneering work by Tolman (1948) who tested rats in mazes, however, has shown that the pure investigation of stimulus-

response relationships is not sufficient for explaining observable spatial skills. The existence of representations in mind that internally reflect the external world could no longer be ignored.

The most general meaning of the notion of *cognitive maps* or *cognitive mapping* is "a construct which encompasses those cognitive processes which enable people to acquire, code, store, recall, and manipulate information about the nature of their spatial environment" (Downs & Stea, 1973: xiv).[12]

The *cognitive map* metaphor expresses two main criteria:

1. It claims that a *mental representation* exists in mind. Spatial behavior is mediated by an internal representation structure that corresponds to the external environment in crucial respects.
2. The cognitive map is similar to external spatial knowledge representations like geographic maps in some respects. Spatial tasks are performed on the internal map in a similar way as on an external map in the given situation.

Regarding the second criterion, however, it has become obvious that the map metaphor should not be likened to external geographic maps in a strong sense. Instead, "... the term *cognitive map* is itself conceptually neutral and is taken by most researchers as a generic term to denote the internal representation of space, regardless of the actual form of the representation" (Hirtle, 1998: 264). What are the properties of this internal representation and how can we better grasp its function?

One of the earliest theoretical investigations using the *cognitive map* metaphor was the work by Lynch (1960) who investigated mental representations of city environments. Lynch had his participants draw sketch maps reflecting people's assumptions of their urban environments. The purpose of his investigations was to test the *legibility* of different cities, meaning the city design's capability to be understood and mentally represented by the people who live there. This mental conception – Lynch (1960) uses the term *(city) image* – can be structured by five main elements that build people's cognitive maps: *landmarks*, *paths*, *nodes*, *districts*, and *edges*. To order them by the characteristic of their mathematical dimension, nodes and landmarks are zero-dimensional (punctual), paths and edges

[12] A notion that is often mentioned in the discussion about alternatives for cognitive maps is the notion of *imaginary maps* (Trowbridge, 1913). Trowbridge uses this term to designate either the mental orientation system with respect to a given starting point in a restricted familiar region (domi-centric orientation), or a method of orientation on the basis of cardinal directions centered in the orienting person (ego-centric orientation). This conception seems to presuppose something like a mental representation that is used for performing orientation tasks in a given environment. However, although more than three decades prior to Tolman's (1948) experiments, Trowbridge (1913) uses the term *imaginary map* only for describing erroneous mental reproductions of cardinal directions that deviate from the actual cardinal directions north, west, south, and east. So the term *imaginary map* is only used as a metaphor for *distorted* knowledge about directions. A person producing correct cardinal directions (at any point she stays), according to Trowbridge, has the 'correct' or 'real map' in her mind.

one-dimensional[13], and areas two-dimensional. While *paths* and *nodes* form elements used to navigate on within a city, *landmarks*, *edges*, and *districts* are used for structuring the environment and providing clues for orientation.

The distinction of mental spatial elements according to their dimensionality has become the basis for nearly all subsequent investigations. The order in which people acquire knowledge about their environments has led to the classification of spatial knowledge into *landmark knowledge*, *route knowledge*, and *survey knowledge* (Hart & Moore, 1973; Siegel & White, 1975). The first that is learned about an unknown environment is the existence of a couple of *landmarks*: distinct places or significant objects that can be used for rudimentary orientation and navigation purposes. When mental representations of a given environment are further refined, landmarks are connected by *routes* that can be used as distinct paths to navigate through the environment. Route knowledge, however, does not yet entail an overall understanding of the environment. Only when the distinct routes are integrated into a coherent representation of the two-dimensional structure, the most elaborate form of spatial knowledge – *survey* or *configurational knowledge* – is reached. Like a geographic map, survey knowledge provides a general frame of reference for spatial operations.

The analogy between mental representations of spatial knowledge and external geographic maps has induced an abundance of empirical investigations that showed in which respects the map metaphor may be misleading. The insight gained by these experiments led to a number of extensions and modifications of the map metaphor. In the following section, these extensions and alternative metaphors proposed to replace the cognitive map conception will be reviewed.

2.1.2 Rubber Sheet Maps, Cognitive Atlases, Collages, and Geographic Information Systems

The earliest comprehensive criticism of the map metaphor is found in the investigation by Lynch (1960). Lynch stated that the mental representation of the spatial environment is "not a precise, miniaturized model of reality, reduced in scale and consistently abstracted" (Lynch, 1960: 87). Rather, the mental representation severely deviates from reality, it eliminates existing entities or adds entities not found in the world, it induces distortions and artificial structures. Nevertheless, mental representations typically preserve an important spatial aspect: the *topological relationships* between entities in the world. Invariance in topological relationships ensures that the overall (two-dimensional) structure of the represented environment is maintained. For instance, containment relationships or the order of neighboring areas are preserved, whereas shapes, distances, and directions can be distorted to an arbitrary extent.

---

[13] Since *edges* form linear boundaries between different areas, and since they are not used for navigating *on* but rather *across* them, they also may be expressed by the two-dimensional element of *districts*; districts decompose a city in several subareas.

To describe this phenomenon, Lynch uses a vivid metaphor: a *rubber sheet*. "It was as if the map were drawn on an infinitely flexible rubber sheet; directions were twisted, distances stretched or compressed, large forms so changed ... as to be at first unrecognizable. But the sequence was usually correct ..." (Lynch, 1960: 87).

Lynch also stated that mental representations of spatial environments do not exhibit a monolithic structure:

> "Rather than a single comprehensive image for the entire environment, there seemed to be sets of images, which more or less overlapped and interrelated. They were typically arranged in a series of levels, roughly by the scale of area involved ..." (Lynch, 1960: 85f.).

With this characterization he already anticipated the cognitive atlas metaphor, first used by Kuipers (1982). Kuipers (1982) suggests to consider mental representations of spatial knowledge as something like a *cognitive atlas*. Mental representations of geographic knowledge split into many partial representations that are often only loosely linked to each other. Every representation structure may have different representational properties. For example, they can substantially differ in scale and resolution, just like in a conventional atlas. Processing knowledge that has to be taken from different partial mental representations requires higher effort than using spatial information from just one representation structure.

In conventional atlases, maps are organized according to hierarchical principles. In a similar way, mental spatial representations use hierarchies to organize partial structures. Theories about mental representations of spatial knowledge can be distinguished according to whether they account for mental hierarchical principles or not (McNamara et al., 1989). Non-hierarchical theories assume a holistic, map-like structure (e.g. Thorndyke, 1981), whereas in hierarchical theories, memory is organized in nested levels of detail (e.g. Stevens & Coupe, 1978; Hirtle & Jonides, 1985; McNamara, 1986). Hierarchical structures in human memory are induced by explicit organization principles (e.g. political boundaries, Stevens & Coupe, 1978), by implicit principles (e.g. non-spatial information, Hirtle & Jonides, 1985), or they are induced idiosyncratically, when no external principles are provided (e.g. McNamara et al., 1989). Moreover, hierarchical structures need not be complete, i.e., partial hierarchies are possible as well (e.g. McNamara, 1991).

The different partial representations in a conventional atlas use different reference systems in which geographic information is organized. Similarly, different spatial reference systems are also used in the mind. These mental reference systems are either induced by external organization principles (e.g. geo-referenced), or they are mentally generated through the relations of the represented entities with respect to each other. From the laws of Gestalt psychology it is known that spatial proximity between objects supports mental grouping and re-orientation of entities with respect to each other (Gogel, 1978; Rock & Palmer, 1990). In a series of experiments Tversky (1981) showed that people tend to mentally move and rotate geographic entities such that they form more ideal configurations. For

example, North and South America are displaced in east-west direction in such a way that they are vertically aligned with each other. In a similar way the American continents are horizontally aligned with Europe and Africa. With respect to external rectangular reference systems (in analogy to the sheet of paper a map is printed on) extended geographic entities are rotated to fit a more ideal orientation. For example, people tend to rotate the African continent and South America in such a way that they appear in a more upright orientation (Tversky, 1981). Direction information is often idealized according to external rectangular reference systems (e.g. 4-sector conceptualization, Huttenlocher et al., 1991).

Similarly, superordinate organization structures can serve as reference systems. As reported above, people distort spatial locations (e.g. cities) when superordinate organization structures (e.g. states) impose orientation biases. Consequently, when superordinate structures (e.g. continents) are moved or rotated, the locations of subordinate entities (e.g. cities) are distorted accordingly. For example, Rome, Italy is farther north than Washington D.C. In a mental representation, Washington may be farther north than Rome due to the mental alignment of North America with Europe (Tversky, 1981).

The *cognitive collage* metaphor has been proposed by Tversky (1993) as another metaphor describing distortions in mental representations of spatial knowledge. The collage metaphor accounts for hierarchies, for contradictions, as well as for multimodal knowledge sources. The idea is that mental representations of spatial knowledge do not resemble a coherent map or an ordered system of partial representations like an atlas. Rather, mental representations should be conceived as collage-like conglomerates of heterogeneous pieces of spatial information.

For example, knowledge about distances between places may be mentally represented in an asymmetrical way. When salient reference points are involved (e.g. well-known places, cities, etc.) the mentally represented distance *from* a reference point *to* another location may deviate from the represented distance in the opposite direction (Sadalla et al., 1980). Distance estimates can also be distorted by a person's perspective on a set of information. Holyoak and Mah (1982) found that distances between two locations are estimated smaller when these locations are farther away. Differing levels of accuracy in partial representations of an environment (e.g. metric vs. topological representations) can result in inconsistencies, for example when local and global representations of orientation information are combined (Moar & Bower, 1983).

In opposition to the cognitive collage metaphor, Tversky (1993) uses the notion of *spatial mental models* when a person has detailed and accurate knowledge about a spatial area of manageable size (see Section 2.1.3).

The cognitive atlas metaphor has been adopted and further extended by Hirtle: "By the use of this term, I mean the complex information set of spatial, visual, and declarative knowledge that is typically found in an atlas" (Hirtle, 1998: 265). Hirtle (1998) uses this extension of the cognitive atlas metaphor to propose the *geographic information system* (GIS) metaphor for mental representations of spatial knowledge. GISs can be regarded as the digital form of traditional atlases.

By the GIS metaphor, Hirtle (1998) emphasizes specific technical representation and processing characteristics found in GISs that can be applied to the explanation of mental representations of spatial knowledge. These characteristics comprise the use of raster vs. vector-based representations in GISs (cf. Couclelis, 1992), changes of accuracy when changing representational formats, issues of scale and resolution, techniques of data overlays used in GIS representations to account for partial representations that can be refined on demand, as well as the application of the problem of data integrity in GISs to representations of spatial knowledge in the human mind.

Golledge (1992) has pointed out common characteristics of acquiring, storing, and processing geographic knowledge in GISs and in the human mind. Peterson (1995) suggests the term *human geographic information system* (HGIS) in contrast to technical GIS systems to point to the specific cognitive characteristics involved.

2.1.3 Spatial Mental Models

Johnson-Laird (1983) proposed the notion of *mental models* to describe mental reasoning processes that require the integration of a set of given premises into a common representational framework for performing a reasoning task. He models the mental representation by a structure that exhibits representational characteristics analogical to the structure of the represented domain. In many cases this analogical characteristic of the representation refers to the ordering structure between the represented entities.

Another conception of mental models has been developed at the same time by Gentner and Stevens (1983) for referring to mental representations of common sense knowledge about physical properties of the world. Applications for this notion of mental models can be found, for example, in mental knowledge representations about the dynamic properties of liquids or the laws of mechanics (cf. the ideas of *naive physics*, Hayes, 1978; 1985).

Different authors use the notion of *spatial mental models* in different ways. First, 'spatial mental model' just means 'mental representation of spatial knowledge' in a general sense. Second, the term describes mental representations of spatial situations for mental inference processes in the sense of Johnson-Laird (1983; 1990). Examples are found in (Knauff et al., 1995) or (Schlieder & Berendt, 1998). Third, the term 'spatial mental model' is used synonymously with *situation model* in text understanding and discourse processing (e.g. Tversky, 1991; Kaup et al., 1999). Tversky (1993) uses the notion of spatial mental model in a fourth sense, namely in contrast to *cognitive collages* (see previous section). Spatial mental models according to Tversky (1993) are more or less accurate qualitative representations of simple or well-learned environments. Unlike cognitive collages they do not contain metrical relationships but form a coherent

albeit coarse representation of a spatial environment.[14] Examples for all types of mental models are presented in (Rickheit & Habel, 1999).

All variants of the notion of spatial mental models exhibit the constructional aspect of the representation in mind. Spatial premises are integrated in a common representational format for spatial reasoning, successively presented narrative data are integrated to form a unified representation of the tentative spatial situation, or loosely represented pieces of information are integrated into a consistent mental representation.

Kosslyn (1994a: 324) has proposed that mental models dealing with visual or spatial information are realized by *visual mental images* in the human mind. This may even be true for abstract, non-spatial information that can be mapped to a spatial structure, i.e., when a spatial analogy can be used in the mind. Mental images are evoked and operated on in working memory (Baddeley, 1986). Mental images and working memory are treated in Section 2.3 and in Section 2.2, respectively.

To prevent confusion I will avoid the term (spatial mental) 'model' as a denotation of a mental representation in this thesis. In the following, 'model' will refer to the computational model that I develop.

2.1.4 Other Conceptions

Finally, I will point to two conceptions for mental representations of spatial knowledge that have been developed outside the tradition of the cognitive map metaphor: *image schemata* and *inter-representation networks* (IRN).

Image schemata (Johnson, 1987; Lakoff, 1987) have been used to describe spatial relationships between objects according to basic mental categories. The idea of image schemata is to express elementary relationships between objects in space. These elementary relationships are characterized as immediately accessible through perception in a spatial environment. Some of them are very similar to the spatial concepts proposed by Lynch (1960). Examples are the *container* schema (for something that contains something else), or the *path*, *surface*, and *contact* schemata for spatial entities. The spatial concepts formulated by image schemata can be mapped to the aspects of spatial knowledge formalized in qualitative spatial reasoning (cf. Section 2.4). Image schemata have been applied in the description of human navigation task performance (Raubal, 1997), for the description of users' concepts in human-computer interaction (Kuhn & Frank, 1991), as well as in the design of geographic information systems (Frank & Raubal, 1999).

Inter-representation networks (IRN) emphasize the relationship between spatial knowledge about the external world represented internally in the mind, and spatial knowledge embodied in the environment, i.e. externally reflecting mental representations (Portugali, 1996b). IRN postulate the existence of complex and dynamic interactions between the mind and the world, which are used in performing

[14] When spatial mental models of this fourth type are acquired by spatial descriptions, this fourth type of spatial mental models may coincide with the third one.

spatial tasks. The core idea of IRN is that investigations of spatial cognition skills have to focus on these interactions. From an AI point of view this perspective is related to *situated cognition* approaches that avoid explicit forms of internal representation. The world is regarded as the most appropriate model of itself, which can be accessed and utilized by an agent operating in the world through immediate sensor-motor interaction (e.g. Brooks, 1991).

2.2 Human Memory

This section reviews general psychological conceptions of human working memory and long-term memory structures. More specific memory aspects like learning issues, temporal characteristics of retrieval and decay of memory contents, or interference of memory contents that are not relevant for the present work are not reviewed.

As in the investigation of spatial knowledge representation and processing (cf. Section 2.1), the development of models for human memory in general has been driven by metaphorical conceptions. Interestingly, nearly all metaphors for human memory structures are spatial (even those not related to the computer metaphor of the human mind, like the ancient wax tablet metaphor, the house metaphor, or the birdbrain analogy). Roediger (1980) gives an overview of the memory metaphors that have been used so far.

"In thinking ... of mind, we usually resort to a metaphor of an actual physical space, with memories and ideas as objects in space" (Roediger, 1980: 232). This spatial memory metaphor has two core implications: first, memory contents are treated as discrete entities (like real-world objects) located in some place within the mind's space. Second, for retrieving information stored in mind it is necessary to perform *search operations* (like in finding physical objects in real space). Analogies that stem from the computer analogy of the human mind often exhibit spatial organization structures that support retrieval processes, for example by spreading activation processes (Meyer & Schvaneveldt, 1971; Ratcliff & McKoon, 1981).

It has proven useful and plausible to model human memory by several separate sub-stores instead of by one single system. These subsystems have been distinguished in *primary* and *secondary memory* (Waugh & Norman, 1965), in *short-term* and *long-term stores* (Glanzer & Cunitz, 1966), or in *sensory register*, *immediate* or *short-term storage*, and *long-term storage* (Atkinson & Shiffrin, 1968).

The idea of the latter distinction described as the *modal model* by Atkinson and Shiffrin (1968) is that a limited amount of information provided by the senses is first stored in a sensory register. From there, information can be shifted into short-term memory using attention processes. In short-term memory, information is kept for some time before it is either transferred to long-term memory, or just dropped to make room for new pieces of information. The transfer to long-term memory is performed when appropriate *rehearsal* processing causes the content to be memorized.

There is evidence that not only the mere rehearsal of short-term memory content is responsible for transferring information to long-term memory, but also the way it is dealt with. The *levels of processing* model (Craik & Lockhart, 1972) accounts for the 'depth' of processing of information held in short-term memory: storing information in long-term memory is much more efficient when information is semantically linked to states of affairs already represented in the mind. For this reason, the levels of processing model is not consistent with the conception of short-term memory as a mere storage device prior to memorization. Instead, it focuses on the *type* of processes that are performed with the pieces of information to be represented in long-term memory (Gazzaniga et al., 1998). Consequently, the distinction in short-term memory and long-term memory has been replaced by the notion of *working memory* as differentiation against *long-term memory* (Baddeley, 1986). The working memory and long-term memory conceptions will be described in the following two subsections.

2.2.1 Working Memory

According to Baddeley (1986) the characteristics of the working memory model account for the realization of the shortcomings of the distinction between short-term memory and long-term memory used up to that point.[15] The working memory is a processing system with limited capacity that is used for processing information that either stems directly from the senses (i.e., by the way of sensory memory) or that has been retrieved from long-term memory. Moreover, there are (at least) two subsystems together with a common controlling instance instead of just one single working memory unit. These subsystems are the *articulatory loop* for auditory information, and the *visuo-spatial scratchpad* for visual and/or spatial information.[16] Logie (1995) further subdivides the visuo-spatial working memory component to account for differences between dealing with spatial and with visual mental tasks.

The working memory subsystems are coordinated by a central component, the *central executive* (Fig. 2.1). The central executive mediates between the two working memory subsystems and the long-term memory. As all working memory subsystems coordinated by the central executive retain information only for a short period of time (e.g., about 1.5 seconds for auditory information) periodical *rehearsal* is required to maintain a persisting representation in working memory. These rehearsal processes are also controlled by the central executive. Moreover, the central executive is used to translate information between the different modalities when required, i.e., it can evoke visual representations in the visuo-spatial scratchpad using auditory information and vice versa.

---

[15] The notion of working memory comprises a variety of complex human memory characteristics, which led to several different theoretical frameworks. A comprehensive overview together with a comparative discussion is found in (Miyake & Shah, 1999).

[16] Sometimes also called visuo-spatial *sketchpad*.

Besides the fact that it accounts for different codes for auditory and visuo-spatial information, the advantage of the working memory conception is that is fills the gap between short-term memory structures and long-term memory in a bidirectional way. Information does not need not be processed by one of the subsystems prior to being stored in long-term memory. Rather, the subsystems are used for processing different types of knowledge in working memory and keeping them vivid.

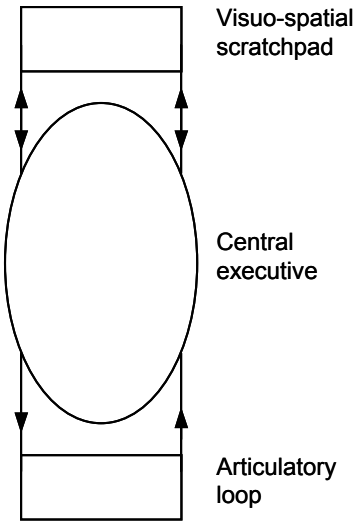

Fig. 2.1. The working memory model according to Baddeley (1986: 71)

2.2.2 Long-Term Memory

Like in working memory, different subsystems have been distinguished in long-term memory. One distinction has been made between *explicit* and *implicit* memory. This explicit – implicit distinction can be related to *declarative* and *non-declarative* types of knowledge in long-term memory in the following way (cf. Gazzaniga et al., 1998):

- Explicit (or declarative) memory comprises knowledge about facts (or *world knowledge*) and knowledge about events (or *personal knowledge*). The former, also called *semantic memory* is considered independent of the owner and the circumstances under which it has been acquired; the latter, also called *episodic memory* is closely related to the individual person and her experiences. Both types of knowledge may overlap, i.e., known facts can be complemented by remembering how they have been acquired. Common to both types of explicit long-term memory is that they are consciously accessible to their owners.

- In contrast, implicit (or non-declarative) memory content comprises all kinds of memory unconscious to the person who has them. Among those kinds of memory are procedural knowledge (i.e., knowing how to do something), habits, or knowledge related to perceptual priming. Common to all types of implicit memory is that they do not require intentional or conscious recollection of experiences prior to their use (Schacter, 1987).

When knowledge is retrieved from long-term memory, it is assumed that *spreading activation* processes are involved both for implicit and explicit memory. This means that if one conceives of memory content as some kind of network of intertwined facts, retrieving one fact will cause the activation and retrieval of a related fact next. Memory networks have been conceived as organized according to conceptual hierarchies (Collins & Quillian, 1969) or according to individual strengths of association between the represented concepts (Collins & Loftus, 1975). Activation spreading effects can be demonstrated in empirical investigations using techniques of associative priming (Meyer & Schvaneveldt, 1971; Ratcliff & McKoon, 1981).

2.2.3 Interacting Memory Systems in Mental Imagery

Before going into detail on mental processing of visuo-spatial information (*mental imagery*) in the next section, I will demonstrate how the different memory subsystems (like short-term memory, long-term memory, working memory, or visuo-spatial scratchpad) can be integrated with each other. This conceptual consideration is a necessary precondition for designing the intended processing model (cf. Chapter 3).

Mental imagery is performed in working memory in the sense of Baddeley (1986). Working memory can be conceived as relying on short-term memory[17] as well as on long-term memory (Kosslyn 1994a: 324). Furthermore, control processes are required for the interaction of the different subsystems.

Visual mental images are evoked in the visual buffer (a transient perceptual representation structure for visual information, see Section 2.3). This short-term memory structure requires periodical maintenance to keep the represented information. With respect to the working memory model of Baddeley (1986), the visual buffer can be related to the visuo-spatial scratchpad; image maintenance is performed by processes that can be related to the central executive in Baddeley's model.

Both the information that is used to construct visual mental images and the data needed for maintaining them is taken from activated structures in long-term memory (provided by specific long-term memory subsystems, cf. Section 2.3). As both long-term memory and short-term memory is involved in visual mental imagery,

[17] This use of the term 'short-term memory' must not be confused with the early notion related to the modal model (Atkinson & Shiffrin, 1968). I follow Kosslyn (1994a) who uses 'short-term memory' to characterize memory subsystems that require periodical maintenance to prevent their content from fading. So 'short-term memory' differentiates against long-term memory structures that usually do not need maintenance.

we have to distinguish between activated and non-activated structures in long-term memory. Activated long-term memory structures contain information that is actually used in the construction and maintenance of a visual mental image contained in the visual buffer, whereas non-activated long-term memory structures are not (yet) involved in current mental imagery processes.

The central executive subsystem in the working memory model according to Baddeley (1986) can be related to specific control processes that mediate between long-term and short-term memory (i.e., processes that activate long-term memory information and maintain images in the visual buffer). Due to capacity restrictions in the visual buffer, information must be swapped between long-term and short-term memory structures during a reasoning task that involves visual mental images (Kosslyn, 1994a: 324). The relations between the different notions and memory subsystems described so far are depicted in Fig. 2.2.

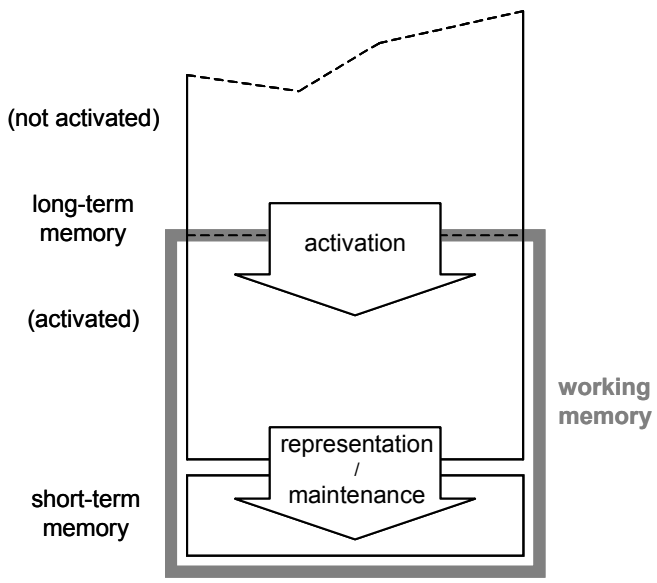

Fig. 2.2. The relations between long-term memory, short-term memory, and working memory

2.3 Mental Imagery

Mental imagery can be defined as "the mental invention or recreation of an experience that in at least some respects resembles the experience of actually perceiving an object or an event, either in conjunction with, or in absence of, direct

sensory stimulation" (Finke, 1989: 2). This characterization points to the crucial aspects relevant for the theory of visual mental images:

- Mental images are *constructed* in the mind. Besides *image construction*, there are also processes of *image inspection*, which make use of the mental image, and of *image modification*, which are used to conceive of alternatives in already constructed and inspected images.
- The characterization of mental images as an *experience*[18] points to the question whether they are a form of mental representation in their own right, or whether they are only epiphenomenal, i.e. an illusion that is produced by some other, non-pictorial mental representation. This question has been discussed extensively in the *imagery debate* in cognitive science and will be briefly reviewed in the following section (for an overview, see Kosslyn, 1994a).
- Both *reconstruction* and *invention* is effective in mentally operating with images. This means that either pictorial knowledge stored in memory is recalled (or reconstructed) in a mental image, or that completely new images not corresponding to a real former perception are constructed in mind.[19] In fact, actually employing mental images in reasoning commonly makes use of a combination of both cases, as reconstruction of images from memory is complemented by fictitious details (cf. Logie, 2001).
- Mental imagery is directly interwoven with *visual perception*. That is why it seems sensible to assume that from the mental image processing point of view it cannot be distinguished (1) whether the source of a mental image is some information taken from memory, (2) whether it stems from an actual current visual perception of an object or an event, or (3) whether it is a combination of both conditions. Indeed, as will be shown on the basis of the Kosslyn (1994a) model in Section 2.3.3, visual perception always involves imagery processes (whereas the reverse is not necessarily the case).

2.3.1 The Imagery Debate

One of the influential foundations of the theory of mental imagery are the experiments performed by Shepard and Metzler (1971) and by Cooper and Shepard (1973). The authors showed participants different views of visual depictions of geometric objects. They asked the participants to decide whether two views that deviated from each other by different viewing angles showed the same object or two different objects. They found that reaction times varied proportionally with the rotation angle between the different views of the objects. This observation led to the conclusion that people use something like a spatio-analogical representation

---

[18] Note that, although mental imagery is characterized as an *experience* in the above definition, the subject matter of the theory of mental imagery are the underlying neuropsychological principles (i.e., the mental representations and processes), not the investigation of the experience of having mental images itself (cf. Kosslyn, 1994a: 3).

[19] This type of mental image can be related to *spatial mental models* in reasoning in the sense of Johnson-Laird (1983), cf. Section 2.1.3.

in mind and mentally rotate within this representation to decide about the equivalence of the depicted objects.

This conclusion of a particular spatio-analogical representation format in mind has been criticized in cognitive science (mainly by Pylyshyn 1973; 1981). The core of the discussion in this *imagery debate* dealt with the question of whether there is a specific (spatio-analogical) representation format for visuo-spatial types of information in the mind, or whether every information processed in the mind is represented in a unique (propositional) format.[20] On the one hand, a couple of empirical results suggested the existence of two different formats, for example experiments dealing with mental scanning by Kosslyn (1973), or the *dual coding theory* by Paivio (1971; 1986) who postulated two different representation formats that are used in the reproduction of words from memory. On the other hand, alternative explanation schemes have been presented for the phenomena mentioned above, which solely rely on a single, propositional representation format. The main issue of criticism was the question who might 'look' at the internal images, based on the assumption that truly pictorial representations require an observer (Pylyshyn, 1973).

This criticism, together with others, has been refuted by the protagonists of the theory of mental imagery (e.g., Kosslyn & Pomerantz, 1977; Kosslyn, 1980; 1994a). The evidences for the existence of quasi-pictorial memory structures seemed so plausible that it has received a wide acceptance in cognitive science. For example, Farah (1988) reports parallels between perceptual and imagery impairments in patients with brain damages. It seems plausible to assume that the same mental systems are used in vision and in imagery because of these parallels in malfunction.

However, Pylyshyn (in press) argues that neuroscientific evidences in favor of mental imagery do not yet prove that there are spatio-analogical representation structures involved in mental imagery. According to his position, reasoning with 'mental images' might be based on the same principles as mental reasoning in general, with the additional feature that some information is involved about how things would look like.

2.3.2 Psychological and Neuroscientific Foundations

Mental images are subjective phenomena. Therefore, to meet scientific criteria, their characteristics cannot be investigated in a direct, straightforward manner. Rather, the characteristics of mental images have to be inferred from indirect measurements performed in empirical experiments. Kosslyn (1980) gives an ample overview over the psychological investigations that have been conducted to

[20] To speak of just one imagery debate is a simplification: Kosslyn (1994a: 4) identifies a total of three imagery debates, which relate to different phases of the discussion. These are characterized (1) by the mere debate about the type of representation format (i.e., whether it is pictorial or propositional); (2) by the discussion about the empirical results gained during phase 1; and (3) by the inclusion of insights in the brain functions participating in imagery processes.

exhibit the underlying representation and processing principles of mental imagery. The four main characteristics that form the basis for the development of Kosslyn's (1980) computational model (cf. Section 2.3.3.1) are the following:

1. Mental images are *not epiphenomenal*. Rather, mental images are based on quasi-pictorial representation structures that account for the phenomena that can be experienced and investigated. The refutation of the arguments of the proponents of purely propositional representation structures has already been mentioned above.
2. Mental images are *not retrieved in a read-made form* from long-term memory, but are *constructed* when needed. It can be derived from experimental findings that mental images are not simply retrieved from long-term memory when they are needed in active working memory; neither are they stored in the same way as they occur in working memory. Mental images are constructed in working memory from pieces of information that are contained in memory in a fragmented and distributed manner.
3. Mental images are *neither retrieved as a whole nor in a piecemeal manner*. Rather, they are composed from organized units that are retrieved from hierarchically structured representations in long-term memory. This characteristic is derived from experiments that exhibit that the organizational form of a pattern and not just the quantity of material needed to evoke a mental image is critical for the image formation time (cf. Kosslyn, 1980: 99).
4. Mental images are not only generated from *pictorial* information, but they are also constructed using *descriptive (propositional)* information from long-term memory. This result, although thoroughly investigated in research on mental imagery, does not seem very surprising. For example, when considering that mental imagery is used in mental visualizations of states of affairs in text understanding, it becomes plausible that images can even be purely based on descriptive information (e.g. Intraub & Hoffman, 1992).

A conceptually different characterization of the properties of mental imagery is given by Finke (1989). Finke identifies five *unifying principles* underlying the phenomena that are empirically investigated in mental imagery:

1. The *implicit encoding* principle: one of the most important functions of mental imagery is to exhibit properties about entities that, despite the fact that they can be mentally visualized, are not otherwise encoded in memory. So mental imagery "is instrumental in retrieving information about the physical properties of objects, or about physical relationships among objects, that was not explicitly encoded at any previous time" (Finke, 1989: 7). This property of image-like representations is one of the main motivations of *diagrammatic reasoning* (DR) in AI (Glasgow et al., 1995; see Section 2.5). The main advantage of the implicit encoding principle lies in the efficiency regarding space requirements: a vast amount of spatial information needs not be computed and represented in advance but can be exploited on demand when needed.
2. The *perceptual equivalence* principle: "Imagery is functionally equivalent to perception to the extent that similar mechanisms in the visual system are activated when objects or events are imagined as when the same objects or events

are actually perceived" (Finke, 1989: 41). This principle is crucial for the development of the second functional mental imagery model by Kosslyn (1994a): mental imagery is found to be imperative in visual perception and object recognition (cf. Intraub et al., 1996; see Section 2.3.3.2). So the distinction whether imagery occurs together with or completely without external visual stimuli does not affect the operation of the neural structures interacting in imagery processes.

3. The *spatial equivalence* principle: the assumption that there is a quasi-pictorial representation medium effective in mental imagery processes has already been mentioned above. "The spatial arrangement of the elements of a mental image corresponds to the way objects or their parts are arranged on actual physical surfaces or in an actual physical space" (Finke, 1989: 61). The *visual buffer* is the central quasi-pictorial representation structure and therefore an essential part in mental imagery models (Kosslyn 1980; 1994a). Similar data structures have also been used in AI systems operating with pictorial knowledge (see Section 2.5).
4. The *transformational equivalence* principle: the empirical findings about mental rotation have already been mentioned above (Section 2.3.1). "Imagined transformations and physical transformations exhibit corresponding dynamic characteristics and are governed by the same laws of motion" (Finke, 1989: 93). This principle is essential for mentally operating with knowledge about dynamic systems, like mechanical devices or navigational issues. It has also been used in artificial systems dealing with reasoning processes about dynamic domains (see Section 2.5).
5. The *structural equivalence* principle: the fact that mental imagery is tightly interwoven with visual perception entails the possibility of dealing with structural properties of mental images instead of real world objects. "The structure of mental images corresponds to that of actual perceived objects, in the sense that the structure is coherent, well organized, and can be reorganized and reinterpreted" (Finke, 1989: 120). This property is essential, for example, for anticipating change, for planning motion processes, and for mentally performing spatial configuration tasks. Nevertheless this principle refers to capacity questions, i.e., the problem of how many items can be properly tackled simultaneously in mind (cf. Hegarty, 2000).

2.3.3 The Kosslyn Models

In the following, I will present two models that account for the findings and principles reported above. The first model (Kosslyn, 1980) is an implemented computational model that specifies data structures for image construction, inspection, and modification; the quasi-pictorial structure used is constructed according to a *cathode-ray tube* (CRT) *metaphor*. It allows for displaying images on a positional (i.e., raster-based) representation structure. The second model (Kosslyn, 1994a) is a functional model that does not come in an implemented form. It describes the

interaction of distinct functional subsystems that are identified according to recent results in neuroscience.

I selected these two models for two reasons. First, the 1980 model is the first model that accounts for the characteristics of mental imagery in a comprehensive and psychologically valid way. Previous models only simulated selected aspects of imagery (for a review see Kosslyn, 1980). Second, the 1994 model seems to be the most elaborated imagery conception that is in line both with psychological findings and neuroscientific results. As such, it is a significant step towards the attempt to bridge the gap between indirect measurement methods performed by empirical psychologists and the bottom-up approach taken by neuroscientists investigating brain structures.

2.3.3.1 The 1980 Model

Kosslyn's (1980) first mental imagery model is based on psychological findings about human reasoning processes in visuo-spatial domains. He presents an implemented computational model for image generation, image modification, and image inspection by building up the corresponding representation structures, as well as by defining the tentative processes operating on these structures. Regarding the representation structures, Kosslyn (1980) distinguishes between the *surface representation* and the *deep representation*.[21] The surface representation is intended to model the content of active working memory, which corresponds to the experience of having an image in the mind. The deep representation, on the other hand, contains the information provided by long-term memory structures to build up actual mental images.

The surface representation structure in which actual image generation and modification takes place corresponds to the *visual buffer* in the mind. It is realized by a raster matrix (or *surface matrix*) of cells that can be filled with specific image information (see Fig. 2.3).

The array data structure of the surface matrix has the following characteristics:

- It has a *limited spatial extent* and a *roughly circular shape*. Corresponding to the maximum angle of the visual field that provides visual information under real perception conditions, the visual buffer only allows for mental images with a limited size. Although realized in a rectangular array structure, the surface matrix – at least in the highly resolved center of the array – exhibits a round shape.
- It has a *specific grain* and a *limited resolution*. As an array structure with a limited number of cells, the surface matrix has a specific grain. Its resolution is limited, which may cause image content to be obscured when resolution is exceeded. To denote that resolution has been exceeded in an image, capital letters are used in the surface matrix to encode that more than one point has

[21] These terms derive from Kosslyn's cathode-ray tube metaphor. The surface representation corresponds to the image visible on the screen, whereas the deep representation holds the information necessary to produce this image.

been mapped to a single matrix cell. Otherwise, image points are denoted by lower case letters in the surface matrix.

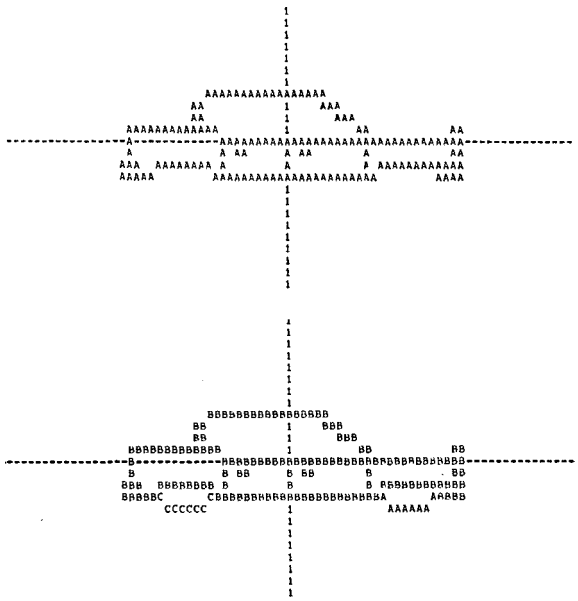

Fig. 2.3. Two examples of images in the visual buffer (Kosslyn, 1980: 154): a skeletal image of a car (above) and the elaborated image completed with details (below)

- The *resolution* of the surface matrix *decreases* towards the periphery. Since mental images are known to have the highest resolution in the center of the image (corresponding to the foveal area of the retina under visual perception conditions) the resolution of the surface matrix is lower near the boundary. This is modeled by using only one cell out of a square of nine cells for image representation.
- The image representation *fades out* over time. The visual buffer has a limited temporal storage capability, i.e., image components fade out and disappear after some time if they are not periodically refreshed. This property is modeled using different letters of the alphabet: letters nearer the beginning of the alphabet indicate that the corresponding image part has been generated or refreshed more recently (cf. Fig. 2.3).

The *deep representation* provides the information necessary to build up an actual mental image in the surface representation. Two sorts of deep representation content are distinguished in the 1980 model: a *perceptual* or *literal* image representation that provides information about what an object really looks like, i.e., its shape; and a *discursive* description that is a list of *propositional encodings* necessary for image compositions:

- The *literal* image representation gives the actual appearance of an object or an object part. For reasons of parsimony and ease of operation (e.g., varying size and location) the representation format chosen in the model consists of lists of polar coordinates. However, other forms like Fourier transforms and shape primitives are also conceivable. These shape definitions are each stored in a separate data file, and they are accessed by their names. Two general types of literal image representations are contained in the model: (1) *skeletal* encodings give an object's central or global shape which is used to provide the overall structure of an image; (2) *individual* encodings of literal information are used to further specify an image of the first type (cf. Fig. 2.3). The individual encodings are linked to the skeletal encoding they belong to by propositional definitions. These propositional definitions are hierarchically organized.
- The *propositional* encodings are lists of facts that are given in a propositional format. Like the shape definitions above, these lists are also stored in separate data files. They are sequentially interpreted during image construction. The information contained in each data file comprises (among others) data about the parts that constitute an entity, an entity's superordinate category, information about the location of an entity or a part thereof (given in qualitative descriptions like 'left of') together with routines for interpreting these descriptions, as well as rough size information.

The described data structures are manipulated by three sorts of processes: (1) processes for mapping information from the deep representation into the surface matrix (*image generation*), (2) processes for making use of the information contained in the image (*image inspection*), and (3) processes for transforming image contents (*image modification*).

1. The available *image generation* processes are PICTURE, FIND, PUT, and IMAGE. PICTURE performs the conversion of a deep representation into a surface representation regarding actual size and location. FIND is used to identify the location where an image component is to be placed. An image component is placed in the image by the PUT procedure with respect to the image parts already contained. The coordination of these three processes is done by the IMAGE procedure, which serves as the interface to the overall system.
2. *Image inspection* is done by the processes LOOKFOR, RESOLUTION, and REGENERATE, together with the image generation processes described above and the image modification processes ZOOM, PAN, ROTATE, and SCAN. The LOOKFOR procedure controls the complete inspection process by employing the other processes. RESOLUTION determines the resolution of the image, whereas REGENERATE is used to refresh the most-faded parts of the image until all parts are refreshed.
3. Processes for *image modification* are ZOOM, PAN, ROTATE, and SCAN. ZOOM moves all points in the surface matrix out from the center, filling in new information into the center area to depict more detail. PAN is the inverse operation to ZOOM. ROTATE moves all points of an image part by a specified angle and direction around a pivot. Finally, SCAN is used to move image contents

along a vector. New image components are filled into the image areas that become vacant during the SCAN process.

Together with the data structures, these processes exhibit a structure – process pair that produces an image processing device by using inherent spatial properties of pictorial representations. This characteristic has also been adopted in the AI subdiscipline of diagrammatic reasoning (e.g., Funt, 1980; Khenkhar, 1991; Glasgow & Papadias, 1992; see Section 2.5).

2.3.3.2 The 1994 Model

Kosslyn's (1994a) second major approach to explore the nature of visual mental images is based on neuropsychological findings about visual perception, i.e., high-level vision processes. As mentioned above, visual perception and visual mental imagery share certain neurological subsystems in the brain. Moreover, the same neural processes operate on real visual input from the visual apparatus and on image representations that are constructed on the basis of mentally stored information. As neural systems in the brain are formed by intertwined clusters of neurons that integrate the structural and the processing aspect of cognition, there is no use in distinguishing (static) representation structures from (dynamic) processes. For this reason, the Kosslyn (1994a) model is described in terms of interacting *functional systems* rather than in terms of distinct structures and processes. These systems and their principal interactions are shown in Fig. 2.4. In the following, the operations of these subsystems for image generation, image inspection, and image modification will be described (Kosslyn, 1994a: 383-8).

Image generation proceeds as follows: when a mental image is to be generated in the *visual buffer*, the appropriate pattern code is accessed in *associative memory*. This is done by one of the two property lookup subsystems, i.e., the *coordinate property lookup* system or the *categorical property lookup* system. These property lookup systems access a pattern code that is linked to the object to be imaged in associative memory. The distinction between *coordinate property lookup* and *categorical property lookup* is made to differentiate between specific shapes stored in memory (i.e., the shapes of specific objects) and general shapes (i.e., prototypical object shapes that are used when there is no specific shape available or required).

The pattern code accessed by the property lookup subsystems is sent to one of the *pattern activation* subsystems, i.e., the *exemplar pattern activation* is triggered by the coordinate property lookup system, whereas the *category pattern activation* is triggered by the categorical property lookup system. The pattern activation subsystems store shapes of objects: the exemplar pattern activation stores the shapes of particular exemplars, whereas the category pattern activation stores members of shape categories that are used as prototypical shape information. The representations in the pattern activation subsystems are activated for generating entire shapes during image generation: the pattern activation that fits best (i.e., exemplar or categorical) produces an activation (i.e., an actual image) in the visual buffer.

The described image generation procedure is used when an entire image is to be mapped into the visual buffer. When more than one image item is needed to

generate an image (*multiple part images*), the property lookup systems access representations that are connected with the object in question in associative memory. Among others, these representations contain information about size and position of the respective part with respect to the item it belongs to (i.e., its 'foundation part'). The order in which additional parts are added to a visualized object is given by their importance or significance for the image to be constructed: "For example, in visualizing a duffel bag, the representation of the handle may be the strongest representation, so it 'wins', and that part or characteristic will be added next to the image" (Kosslyn, 1994a: 384).

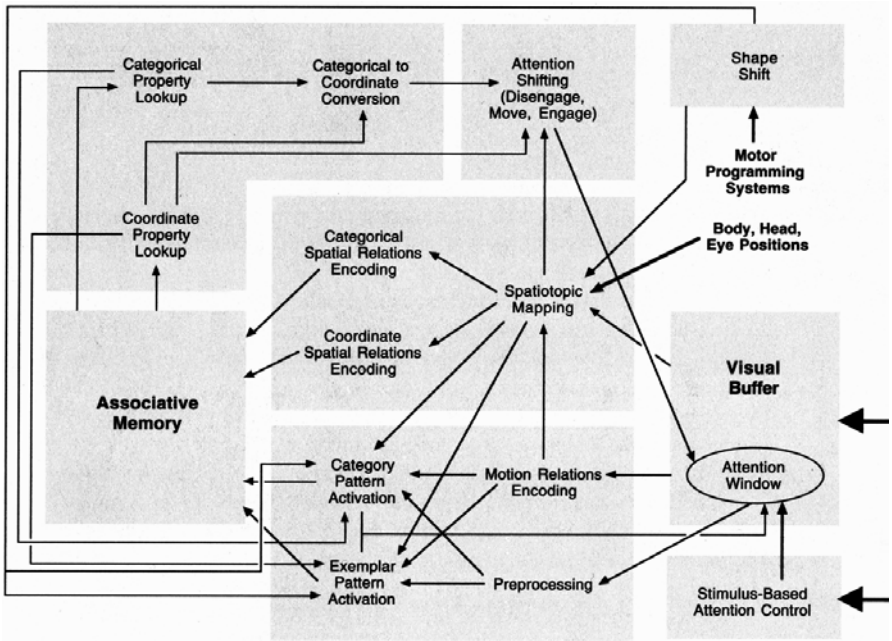

Fig. 2.4. Functional subsystems interacting in mental imagery (Kosslyn, 1994a: 383)

But how are the proper positions of the additional parts determined? In the case of a coordinate image representation (i.e., a specific object is imaged) the position is represented explicitly, and the *attention window* can be positioned accordingly by the *attention shifting* subsystem. The attention window determines the part of the image in the visual buffer actually under operation.[22] In the case of a categorical representation of the position of some part (i.e., when it is given in a categorical form like 'on the top') the actual position is computed by the

[22] Another way of generating mental images is by continuously moving the attention window on the visual buffer and by evoking arbitrary pictorial forms induced by this movement ('mental drawing', see Kosslyn, 1994a: 385).

categorical to coordinate conversion subsystem. In this case, the information passes through this subsystem, which in turn triggers the attention shifting subsystem. In both cases, the pattern code is sent to the pattern activation subsystems as soon as the attention window is properly positioned. The pattern activation subsystems activate the visual buffer accordingly, so that the additional part can be visualized in the proper position and size in the visual buffer.

The described mechanisms for combining objects and parts in a common visual mental image can be used to represent an object up to a specific level of detail, as well as to combine a number of objects – on a coarse level of detail – in a common arrangement (cf. the trunk packing example by Kosslyn, 1994a: 385).

Both ways to construct mental images, i.e., based on categorical and on coordinate information, can be combined in a single mental image. This may be the case when the overall shape of some object is not known, but some detail is represented explicitly. Or the other way around, when an overall shape is exactly represented, but to add a detail that is not explicitly represented, it may have to be referred to a categorical representation. By the image generation processes described above, it is also possible to combine objects and parts with each other that have not been combined before (cf. *creative imagery*; Finke, 1990; 1992).

As claimed above, mental images have to be maintained periodically to prevent them from fading out over time. This *image maintenance* is done by repeatedly activating the pattern activation subsystems. The interplay between the image's fading and the image maintenance controls the storing capacity of the visual buffer. The amount of information that can be maintained in the mental image depends on the degree of *chunking* within the image components to be contained in the mental image:

> "Visual-memory-based images are maintained by repeatedly activating a compressed image representation or set of such representations in a pattern activation subsystem. The amount of material that can be retained in an image depends on how effectively it can be organized into chunks, which is accomplished by the preprocessing subsystem" (Kosslyn, 1994a: 325).

Now, how does *image inspection* proceed? Generally, it has to be distinguished whether object properties (like an object's color or texture) or spatial properties (like relative sizes or locations) are to be inspected in the mental image. Moreover, in both cases properties can be of the exemplar type or the categorical type.

Inspecting object properties activates the exemplar pattern activation subsystem or the category pattern activation subsystem by passing the preprocessing subsystem or the motion relations encoding subsystem, respectively.[23] More precisely, the portion of the visual buffer that is in the attention window causes a specific pattern activation (which results in an object recognition by the respective pattern

[23] As motion is an important factor in object recognition, the purpose of the motion relations encoding subsystem is to detect motion patterns in the visual buffer. Moreover, motion is also being inferred from static visual input to reason about an arrangement's dynamic properties (e.g., Freyd & Finke, 1984; Hegarty, 1992).

activation subsystem), which is matched against some object representation stored in associative memory. A specific situation occurs when one is looking for a particular object property, i.e. a property that is expected in an image due to some kind of expectation. In this case, like in the image construction process (see above), the pattern activation subsystems are primed by some information taken from associative memory by the property lookup systems.

When spatial properties (like the relative size or orientation of two objects or object parts) are to be inspected in a mental image, there are again two subsystems involved: the *categorical spatial relations encoding* subsystem and the *coordinate spatial relations encoding* subsystem. To inspect a rough (qualitative) spatial relation (for instance a coarse spatial orientation), the former becomes effective, whereas the latter is used to decide about precise metric spatial relationships (like a specific distance between two objects or object parts).

In any case, the respective subsystems just process what is contained in the specific area of the visual buffer that is focused on by the attention window. So it may be the case that the attention window is in the wrong position with respect to the visual buffer to inspect some property, or the area covered by the attention window may provide the wrong resolution, i.e., it may be either too small to get an overall impression or too large to inspect a specific detail. In these cases, image modification processes, like scanning or zooming operations, become necessary.

Scanning the visual buffer to focus on some other portion of the image is done by moving the attention window when distances are small. In other cases, *image modification* is necessary. When a part of the image in a farther distance has to be inspected, the image as a whole must be transformed, i.e., the mapping properties from the pattern activation subsystems to the visual buffer have to be changed. This operation is controlled by the *shape shift* subsystem, which is triggered by the *spatiotopic mapping* subsystem. The same may happen when it becomes necessary to perform a zooming operation on the image. Again, the shape shift subsystem causes the pattern activation subsystems to provide an altered image in the visual buffer that fits the needs of the image inspection to be performed.

Having reported on the processes that are performed by the subsystems operating in mental imagery, it is interesting to compare the imagery subsystems to the memory structures identified in Section 2.2.[24] I have described above that working memory is constituted by parts of activated long-term memory and by short-term memory structures.[25] The short-term memory structure that is effective in visual mental imagery is the visual buffer together with the attention window. This memory structure needs periodical maintenance to prevent it from fading, and it has only a limited capacity for representing image contents. Thus, the amount of information activated in long-term memory with respect to a mental imagery task

[24] Both the identified memory structures and the characteristics of mental imagery are essential for the model to be developed.

[25] When considering the notion of working memory not only with respect to representation structures but also to processes, then the control processes effective in mental imagery (e.g., the property lookup subsystems) also need to be mentioned (cf. Kosslyn, 1994a: 324).

typically exceeds the information that can be visualized in the visual buffer: "Given the severe capacity limits of short-term memory ... more information often may be activated in long-term memory than can be represented in short-term memory" (Kosslyn, 1994a: 324). Thus, from the perspective of memory structures, activated parts in long-term memory participate in mental imagery tasks, although they may not be visualized in the visual buffer all the time.

Finally, when comparing the two modeling approaches by Kosslyn (1980; 1994a), it is seen that they are perfectly compatible. Although they seem to differ significantly at first sight, the operational properties of the two models can be mapped onto each other (see Kosslyn, 1994a: 388-95). Whereas the earlier model is derived from empirical findings of psychological experiments, the later model is based on details revealed by neuroscientific investigations. These results underlying the later imagery model verify the assumptions the functional relationships in the earlier model are based upon.

2.4 Spatial Reasoning

AI is concerned with spatial knowledge in several of its subareas. The areas most relevant to this work are qualitative spatial reasoning (QSR) and diagrammatic reasoning (DR). In the present section I will give an overview of spatial and qualitative spatial reasoning techniques. Section 2.5 will report on diagrammatic reasoning.

In computer science, spatial information can be dealt with in a highly precise way. Geometric principles and techniques can be used to operate on metric representations of two- or three-dimensional spatial entities. This is done in the sub-field of computational geometry.[26] However, there are reasons to abandon this precision in information processing tasks that require spatial capabilities:

> "Classical mathematical global spaces (coordinate spaces) have been successfully used in computational geometry. In the knowledge representation area of AI, the necessity to cope with imprecision, incompleteness, and uncertainty of knowledge, both in physical space and cognitive space, led to drop this approach for a more qualitative one." (Vieu, 1997: 10).

The idea of QSR derives from qualitative reasoning techniques in AI (e.g., Allen, 1983; Weld & De Kleer, 1990; Kuipers, 1994). Instead of aiming at homogeneously precise representations, qualitative techniques concentrate on representing *relevant* aspects of the state of affairs under consideration. Relevant in this context may refer to what is required to decide upon a given question, what can be detected by sensory capabilities, or what is necessary to perform a specific behavior (Freksa, 1991; Freksa & Röhrig, 1993; Hernández, 1994).

To put the focus on representing relevant aspects of spatial information is often related to dealing with uncertain and vague knowledge: although it is not possible

[26] In Section 2.4.5 I will give a short overview on computational geometry.

to specify exact values or parameters in a given situation, a cognitive agent may have a sufficiently distinctive representation to cope with spatial problems. From this perspective, qualitative spatial reasoning is interesting for modeling or imitating cognitive skills, for example in the context of geographic information systems, in robot navigation, or in natural language processing.

A key question in qualitative spatial reasoning is what types of basic entities are used to describe a spatial situation. Approaches to qualitative spatial reasoning therefore can be classified according to their ontologies. Important issues are the mathematical and physical characteristics the entities have, together with the basic operations that can be performed with them. Theories of QSR use point objects, array cells, tuples of intervals, region-based entities, or combinations of the former as primitives.

Another classification of QSR approaches can be made according to a hierarchical ordering of expressiveness of the diverse theories. One "... can think of theories of space as forming a hierarchy ordered by expressiveness (in terms of the spatial distinctions made possible) with topology at the top and a fully metric/geometric theory at the bottom" (Cohn, 1997: 15; cf. Kuipers, 2000). In the remainder of this section I will report on representing topological knowledge, orientation knowledge, distances, and shapes. Finally, I will give a short sketch of computational geometry issues.

2.4.1 Topology

Topology describes properties of mathematical spaces independent of angles and distances (e.g., Edgar, 1990; Kong & Rosenfeld, 1996). Topology can be regarded as the most fundamental form of qualitative spatial descriptions. Informally, topological relations can be related to the metaphor of a rubber sheet (Lynch, 1960, cf. Section 2.1.2). Spatial entities are related with respect to each other in terms of whether they are connected with each other, whether one entity contains the other one, or whether two entities overlap.

An important aspect in qualitative spatial reasoning is that the spatial relations provided to describe a spatial phenomenon (e.g., topology) both cover all possible cases, and that they are mutually exclusive.[27] The first systematically developed set of topological relations with this property is given by Egenhofer and Franzosa (1991) by evaluating the possible relationships between the boundaries and the interiors of two spatial regions ('4-intersection' model). Two entities can be completely disjoint, they can be connected with each other either from outside or one being inside the other one, they can overlap, one can be inside the other one (without touching each other), or they can be simply equal. Taking into account that in the case of non-symmetrical relations between two entities it has to be distin-

[27] This characteristic that each spatial situation corresponds to one and only one relation is usually referred to as the JEPD (jointly exhaustive – pairwise disjoint) property in QSR (cf. Cohn, 1997). However, there are also qualitative representation formalisms that are not based on the JEPD property (e.g., Freksa 1992a, b).

guished whether the first entity is related to the second one or the other way around (e.g., one object being inside the other one), eight binary topological relations are obtained (see Fig. 2.5). In their *region connection calculus* (RCC) Randell and co-authors (1992) derived the same set of topological relations based on a logical conception of *connection* between regions.

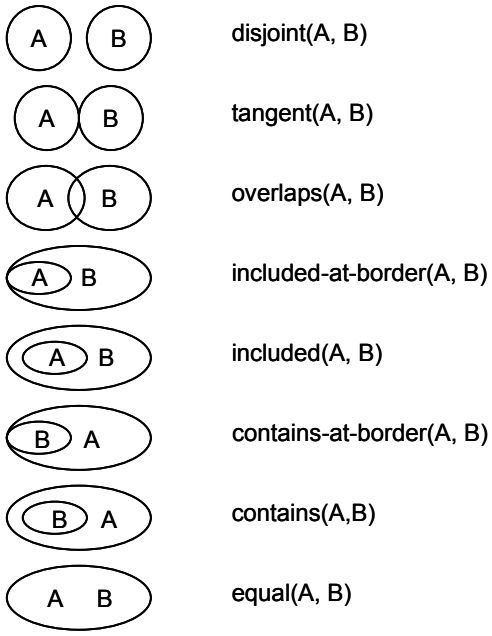

Fig. 2.5. The eight possible topological relations between two regions

In his '9-intersection' model (in contrast to the '4-intersection' model, see above) Egenhofer (1991) also takes into account the relation between the two entities and the embedding space. Grigni and co-authors (1995) use Egenhofer's model at two different levels of resolution. In the high resolution model they distinguish the eight topological relations shown above, whereas in the medium resolution model there are only five different relations (the *tangent* and the *overlap* relation, the *included* and the *included-at-border* relations, and the *contains* and *contains-at-border* relations, respectively, are merged).[28]

The theories reported so far use general regions as basic spatial entities. When considering topological relationships between entities of different dimensionality

[28] There is also a low resolution version of the RCC relations. RCC-5 (in contrast to RCC-8) also contains only five topological relations. In contrast to the medium resolution model by Grigni and co-authors (1995), in the RCC-5 model the *tangent* relation coincides with the *disjoint* relation, not with the *overlaps* relation.

(i.e., points, linear, and areal entities) more specific distinctions have to be made (e.g., Clementini et al., 1993; Clementini & Di Felice, 1995; Egenhofer & Mark, 1995a). These distinctions are especially needed in application contexts, for instance in spatial databases or in reasoning with cartographic entities (Isli et al., 2000).

With respect to the model to be developed, point entities and extended regions will be most interesting. In the possible pairwise combinations between point entities and extended entities, the possible topological relationships are more or less restricted with respect to the general case of two extended regions. Table 2.1 shows which topological relations may hold between two point entities, between a point entity and an extended region (and vice versa), and between two extended entities.

Table 2.1. Topological relations that can hold between two point entities, between a point entity and an extended region and vice versa, and between two extended regions

	point – point	**point – extended**	**extended – point**	**extended – extended**
disjoint	✓	✓	✓	✓
tangent				✓
overlaps				✓
included-at-border		✓		✓
included		✓		✓
contains-at-border			✓	✓
contains			✓	✓
equal	✓			✓

Although in topological relations no orientation information can be encoded (e.g., cardinal directions between geographic entities) the relations *disjoint*, *tangent*, *overlaps*, *contains-at-border*, and *included-at-border* imply the existence of an orientation between the two entities involved (cf. Hernández, 1994). In the following, I will review approaches that explicitly deal with orientation knowledge.

2.4.2 Orientation

To determine an orientation relation, three elements are required: a *primary object* (the object under consideration), a *reference object* (with respect to which the position of the primary object is described), and a *frame of reference* (the system of orientation relations used) (cf. Hernández, 1994). A reference frame is given by "the orientation that determines the direction in which the primary object is located in relation to the reference object" (Retz-Schmidt, 1988: 95). Three types of reference frames can be distinguished: *intrinsic*, *deictic*, and *extrinsic* reference frames:

- Intrinsic reference frames are established by inherent properties of the reference object used (e.g., the front side of an object: "the bike is in front of the house").

- Deictic reference systems are given by an observer's perspective on the reference object (e.g., "the bike is behind the house" – regarded from my point of view).
- Extrinsic reference frames are imposed on the reference object by external factors (e.g., the earth's gravitation, the reference object's accessibility, or the georeference system).

Since I deal with geographic knowledge in this work, I am primarily interested in the extrinsic georeference system, i.e., in *cardinal directions*.

A systematic derivation of orientation relations can be given by considering three points in the plane (cf. Hernández, 1994). With respect to a straight line through two points (*reference axis*) and one of these two points selected as reference point, the third point can be either to the right, to the left, or directly on this line (i.e., collinear). So three orientation relations can be distinguished (see Fig. 2.6a).

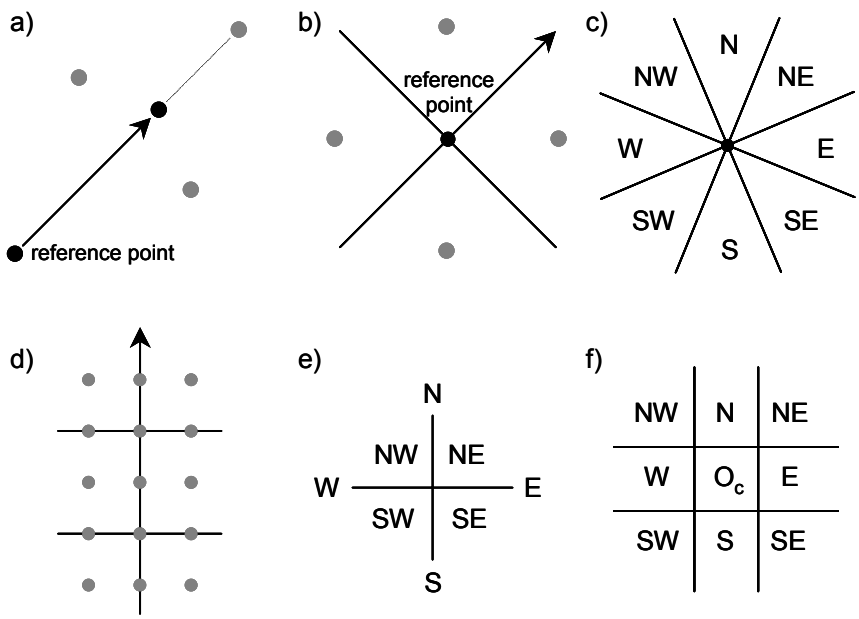

Fig. 2.6. Different models for orientation relationships (see text)

A second system arranged in the reference point, perpendicular to the first one, allows for distinguishing four sectors (e.g., front-left, front-right, right-back, left-back, Fig. 2.6b).[29] By superposing further systems of this kind and by rotating the whole system (for adjusting the sectors with the intended directions), arbitrary

[29] Besides the four sectors there are also four different qualitative positions on the lines.

orientation relations can be constructed (for instance the common cardinal direction systems distinguishing between 4, 8, or 16 cardinal directions, Fig. 2.6c).

The question of whether the directions that correspond to positions directly on the axes (i.e., the collinear cases) should be considered as orientation relations in their own right or not is especially interesting with respect to their cognitive plausibility. Freksa (1992b) argues for the plausibility of mental reasoning with 'exact' directions (and also qualitatively exact locations). In his approach, he uses two axes perpendicular to the reference axis instead of just one, which yields 15 qualitative locations of a point with respect to two given locations: six distinguishable sectors, 6 positions on the axes, 2 positions coinciding with one of the two given positions, plus the position on the straight connection between the two given points (Freksa, 1992b, Fig. 2.6d).

For similar considerations, Frank (1991, 1992, 1996) discusses two approaches for dealing with cardinal directions: (1) the *cone-based* approach, which describes cardinal directions as sectors (cf. Fig. 2.6c), and (2) the *projection-based* approach, which uses axes for the four main directions (north, south, east, west), whereas the intermediate directions are quadrants (cf. Fig. 2.6e). Moreover, he also proposes a *neutral zone* for the projection-based approach. The neutral zone characterizes the area in which someone does not want to assign a cardinal direction to an object, because it is too close to the reference object. So by this model, the plane is divided into nine orientation sectors (Fig. 2.6f).

Isli and co-authors provide a calculus that combines the approaches of Frank (1991, 1992) and Freksa (1992b) in a common reasoning system (Isli et al., 2001). An application of (human) qualitative reasoning with cardinal directions is discussed by Kulik and Klippel (1999). They present a formal description of reasoning about cardinal directions (qualitative geographic coordinates) in grid-based reference systems in maps solely based on ordering information.

2.4.3 Distance

Closely related to orientation knowledge in space is the notion of distance (cf. Cohn, 1997; Vieu, 1997). In the plane (i.e., in Euclidean space) the triangle inequality directly relates orientation and distance.[30] The addition of linear distances (i.e., with respect to a given direction) directly entails orientation. Moreover, for spatially extended entities distance may vary depending on the orientation of two entities with respect to each other. So most approaches dealing with spatial distance information combine orientation and distance knowledge.

Frank (1992) proposes a combined approach to reasoning with both distances and cardinal directions to answer questions like: "given the distances and directions from A to B and from B to C, what is the direction and distance from A to C?" For the distance-related part of this task, Frank uses the two qualitative distances *close* and *far*, whereas the orientation task is based on direction concepts as

[30] For three points x, y, and z the (metric) distance d is constrained by $d(x, z) \leq d(x, y) + d(y, z)$.

reported above. After discussing the distance and orientation parts of the task separately, he indicates how they can be integrated. It shows, however, that distances only combine well when directions are similar, whereas reasoning with directions requires more or less equal distances.[31]

Also the work on *qualitative distances* in geographic space by Hernández and co-authors (1995) is inspired by results from reasoning with qualitative orientation. Instead of just two qualitative distance values, the authors induce a scale of variable qualitative *distance relations* to describe distance intervals which partition the plane in circular regions around a reference object (Fig. 2.7). *Distance systems*, besides the distance relations also contain *structure relations* that describe how the individual distance values relate to each other. These structure relations impose restrictions on the distance relations (e.g., monotonically increasing interval length) and are required for combining distance relations with each other (see also Clementini et al., 1997).

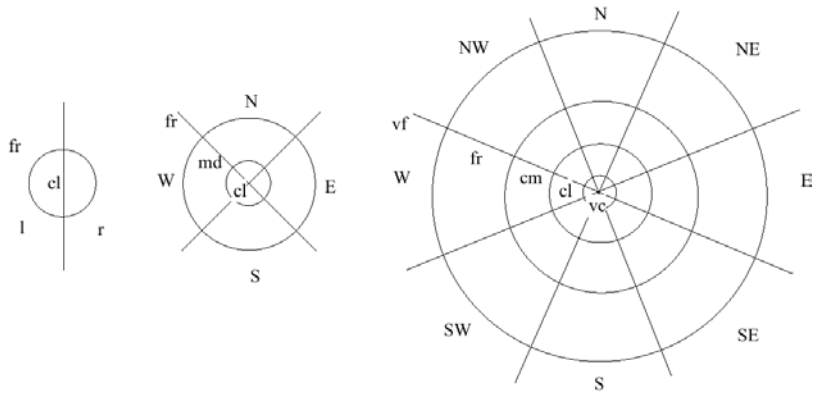

Fig. 2.7. Various granularity levels of distance (and orientation) relations. Orientation and distance granularities may vary independently of each other (Clementini et al., 1997)

Zimmermann (1993, 1995) extends Freksa's (1992b) orientation representation system by a linear order on the edges between the three points that are related to each other (Δ-Calculus). Through this implicit comparison of the edges' lengths, the plane is partitioned in qualitatively distinct areas. This partitioning allows for a mapping between the position knowledge described in Freksa's approach and qualitative (comparative) distance information (see also Zimmermann & Freksa, 1996).

[31] As a consequence of the qualitative approach, complete integration of distance and orientation information is generally impossible (Vieu, 1997).

2.4.4 Shape

Some shape properties of spatial objects are implicitly represented in topological relations (e.g., whether there are holes or whether the object is in one piece). Also, the possible shapes of an object are restricted by its topological relationships to other objects (e.g., the 'included' relations). However, when a more detailed description of shape properties is required, it is necessary to go a step further toward a complete geometric description (Cohn, 1997).

An early approach to describe shape properties of spatial entities is given by Jungert (1993). Jungert uses *symbolic slope projections* (i.e., projections of the contours of an object to two orthogonal coordinate axes) to describe convexities and concavities (i.e., acute and obtuse angles) of spatial entities (see Fig. 2.8a). Schlieder (1996) uses *ordering information* to describe shape characteristics. His approach distinguishes between convex and concave vertices of a polygon by specifying the respective triangle orientations for every triple of consecutive vertices.

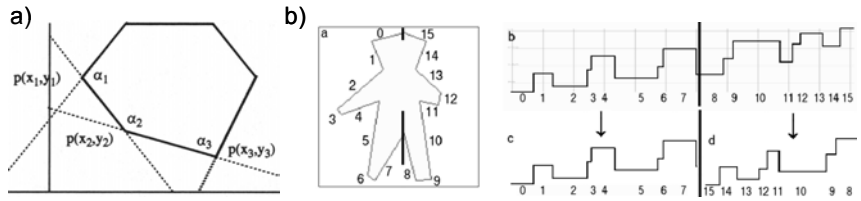

Fig. 2.8. a) The symbolic slope projection method (Jungert, 1993: 447). b) A polygonal curve together with its tangent function (right, above). The similarity in the parts c and d of the tangent function (right, below) exhibit the symmetry of the polygonal curve (Latecki & Lakämper, 2000)

Cohn (1995) proposes a logical approach to describing shape properties of convex entities. Using two primitives, the connectedness between entities and the convex hull operator, he shows how a great variety of shapes can be distinguished. Applying the method recursively to the insides of the primary region under consideration allows for constructing a hierarchical description of an object's shape properties, thus providing shape descriptions on different levels of granularity.

A framework comprising several types of shape characteristics is proposed by Clementini and Di Felice (1997). This framework is based on few primitives that are arranged along the three conceptual dimensions of topological (connectedness, compactness), projective (convexity), and metric (qualitative symmetry, elongation) shape properties, thus forming a *configuration space* for qualitative shape descriptions. However, the issue of the mutual implications between the three dimensions is not dealt with.

An important application for shape representation is in matching shapes with each other (e.g., in comparing objects in image databases). Latecki and Lakämper

(1999a, b) developed the *discrete curve evolution* method for systematically reducing the number of vertices in polygonal curves while preserving the shape's perceptual appearance to the maximum extent. The transformation of the simplified shape to its *tangent function* exhibits shape characteristics like symmetry. Moreover, it allows for retrieving similar shapes from large sets of image data (see Fig. 2.8b).

Another important application domain for representing and reasoning with shape information is in geographic information processing. Discrete curve evolution (Latecki & Lakämper, 1999a, b) has been applied to simplifying geographic shape information. In dealing with geographic shape, it is important that shape modification of one object does not affect essential spatial relationships (e.g., topological and ordering information) between neighboring entities (Barkowsky et al., 2000). The construction of areal objects from entities depicted in cartographic data as part of human map interpretation has been modeled by Steinhauer and co-authors (2001). Also in this task, the context of the shaped entity is crucial. The authors use conceptual representations of *abstract regions* in maps to aggregate appropriate shapes from given map objects.

2.4.5 Computational Geometry

The approaches reported in the last subsection employed more and more quantitative techniques for dealing with spatial information. I will end this section by briefly sketching the purely quantitative field of computational geometry (Preparata & Shamos, 1985; O'Rourke, 1998). Although computational geometry is neither considered as cognitively motivated nor part of AI, it nevertheless serves for differentiating the qualitative AI reasoning techniques against algorithmic methods outside AI.

Actually, many techniques are used inside and outside AI. In the non-AI disciplines, problems are usually properly defined and require optimal solutions, whereas in AI, the available information is often incomplete. Sub-optimal but efficiently computed solutions are sufficient (Schlieder, 1996).

The notion of computational geometry has been coined in 1975 by M. I. Shamos (cf. Preparata & Shamos, 1985: 6; Shamos, 1978). It starts from the insight that from a computational point of view, most classical characterizations of geometric objects are not appropriate. Therefore, to allow for the development of efficient algorithms, it is necessary to work out new concepts for describing geometric problems.

Among others, the field of computational geometry comprises the development of efficient representation structures and computation strategies for geometric search problems (i.e., identifying objects in a data set that fulfill certain geometric criteria), for convex hull computation, for assessing proximity between geometric entities (e.g., by triangulation and Voronoi diagrams), as well as for computing intersections between the diverse types of geometric entities.

2.5 Diagrammatic Reasoning

Human thinking and problem solving can often be enhanced through the use of visual media like pictures, sketches, and diagrams (e.g., Sloman, 1971; 1975). The AI sub-discipline of *diagrammatic reasoning* (DR) aims at adopting the features of diagrammatic representation structures for knowledge representation and reasoning tasks (Narayanan, 1992; Glasgow et al., 1995; Anderson, 1998; Olivier, 2001). The term *diagrammatic* in this context refers to the visual properties of a representation structure that can be utilized by appropriate processing facilities.[32] The information itself that is represented in such a diagrammatic representation structure, however, does not necessarily have to be visual (cf. Kulpa, 1994).

The operations humans carry out with diagrams can be classified in *generation*, *perception* and in *reasoning*. Accordingly, diagrammatic representations are employed in AI and computer science for information *presentation*, *representation*, and *reasoning*.

- The aim of information *presentation* is in encoding (typically non-visual) information in a visual format for ease of communication. The related sub-fields are (computer-aided) graphic design, data or knowledge visualization, or graphical user interfaces (Kosslyn, 1994b; Tufte, 1998; Card et al., 1999).
- The purpose of diagrammatic *representation* is to encode data in a diagrammatic form in a data or knowledge base, for storage and retrieval, either by humans or by computers (Kulpa, 1994).
- The most interesting purpose of using diagrams for the concerns of this work is in diagrammatic *reasoning*. Diagrammatic reasoning focuses on the information and knowledge processing advantages of diagrammatic representations for deriving new pieces of information from given facts. The reasoning process is performed either by humans (be it on external diagrams like sketches on paper or on internal, i.e. mental, images), solely by computers (in DR systems), or in interaction between humans and computers in interactive systems.

As will be motivated below, the distinction between diagrammatic representation and diagrammatic reasoning is not sharp, since the efficiency of the reasoning process substantially depends on the form of representation chosen. In fact, reasoning can even be completely substituted by suitable representations.

Applications of diagrammatic reasoning are among others in spatial reasoning (e.g., Funt, 1980; Khenkhar, 1991), in configuration, layout, and design (e.g., Marks, 1991; Yi-Luen Do & Gross, 2001), in qualitative physics (e.g., Forbus et al., 1991), in problem solving in expert systems (e.g., Glasgow & Papadias, 1992), in didactics and instruction (e.g., Barwise & Etchemendy, 1994; Narayanan & Hegarty, 1998), in theorem proving (e.g., Barwise & Etchemendy, 1995), in visual programming (e.g., Glinert, 1990; Myers, 1990), in automatic diagram under-

[32] Kulpa (1994) defines *diagrammatic* representations as a subspecies of *visual* representations (i.e., related to the visual sensual modality). Other types of visual representations are graphical, pictorial, or geometrical representations.

standing (e.g., Carriero et al., 1992), or in information visualization (e.g., Card, 1999).

2.5.1 Propositional vs. Analogical Knowledge Representation

Diagrammatic representations are *analogical* representations (Sloman, 1971; 1975; Rehkämper, 1995). As such, they are contrasted with *propositional* representations.[33] Analogical representations structurally correspond to the semantics of the problem domain they are about:

> "If R is an analogical representation of T, then there must be parts of R representing parts of T, … and it must be possible to specify some sort of correspondence … between properties or relations of parts of R and properties or relations of parts of T" (Sloman, 1975: 161; cf. Palmer, 1978).

Due to this structural correspondence between the representation and the domain it is about, analogical representations are *specialized* representations (cf. Kulpa, 1994).

Propositional representations, on the other hand, do not have these structural correspondences with the semantics of what they stand for. So in contrast with analogical representations, propositional representations are universal, since a single propositional formalism (say, predicate logic) can be used – at least in principle – to cover all representational requirements in every domain conceivable. As a consequence, a given analogical representation will never be the unique means for representing a state of affairs. Two forms of representation can be *informationally equivalent* (i.e., the same information can be derived from both of them), but they can be *computationally inequivalent* (i.e., an inference may be easier and faster in one of the representations; Simon, 1978). In this sense, a diagrammatic representation may be computationally inequivalent to another, non-analogical representation: it may provide a specific organizational form that is advantageous for the task to be performed with the representation (cf., Freksa, 1988). The main advantages of analogical (or more specific: diagrammatic) representations are the following:

- In diagrams, pieces of information that are used together or that semantically belong together are typically located close to each other. This *locality advantage* (Koedinger, 1992) allows for effective control of the reasoning process, because it significantly can reduce search effort both in data and in problem space (Larkin & Simon, 1987). Moreover, since objects in the representation are spatially related to each other, it is often not necessary to assign individual symbolic labels to them.
- Diagrams possess *emergent properties* (Koedinger, 1992), which means that often inferences are already carried out by just representing the state of affairs

[33] Other notions that have been used instead of '*propositional*' are '*Fregean*' (Sloman, 1971), '*sentential*' (Larkin & Simon, 1987), and '*logical*' (Nilsson, 1980). For the term '*analogical*', also '*direct*' has been used (Barr & Feigenbaum, 1981).

in the diagram. Many solutions can be simply read off after the construction of the diagram, since they are perceptually obvious. This effect makes those consequences explicit that are already implicitly encoded in the given premises (for examples see Kulpa, 1994).

- Another implication of emergent properties in diagrams is that geometrically impossible situations are not considered in the reasoning process, simply because they cannot occur in the diagram. This effect again reduces the problem space (see above). Moreover, these *structural constraints* (Koedinger, 1992) allow for efficiently controlling the inference process when part-whole relations are used as a means for efficient knowledge organization.
- Finally, reasoning processes in diagrams rely on principles that are close to human perception facilities. This implies that reasoning processes and results are usually natural and can be easily understood by humans. So conducting and presenting inferences using diagrams can be advantageous when the inference process or the outcome has to be communicated to humans.

To make use of the advantages of diagrammatic representations outlined above, it must be ensured that the formalism chosen is both *expressive* and *effective* with respect to the requirements at hand (Kulpa, 1994). Whereas effectiveness refers to the ease of representing facts and performing inferences, expressiveness deals with the question whether *all facts* and *only the required facts* are contained in the diagrammatic representation. Particularly the latter requirement, which points to the possible inferences that can be drawn from a diagrammatic representation, is important to consider. All pictorial forms of representation allow for being *over-interpreted*: relations may be evaluated that are contained in the diagram just due to spurious coincidences (Kulpa, 1994), and that do not correspond with what they represent in the presumed way (cf. over-interpretation in map-like representations, e.g. Berendt et al., 1998b).

It is important to note that most existing diagrammatic reasoning systems are *hybrid* systems, i.e., they use both analogical (diagrammatic) and propositional representation structures. The analogical part can be made highly task- or domain-dependent to utilize the advantages of analogical representations summarized above, whereas the overall general control of the system is provided by propositional structures. Especially in text understanding, propositional representation formats are required due to their close relationship to natural language (e.g., Habel, 1990; Habel et al., 1995). Although it is possible to develop sufficiently general purely analogical reasoning systems, it may not be practical (Kulpa, 1994).

2.5.2 Types of Diagrammatic Reasoning Systems

What does '*diagrammatic*' mean when a diagrammatic reasoning system is realized in a digital computer, and how can the above advantages be reached? It is important to note that diagrammatic properties are not provided by a representation structure per se, but that they always result from the representation structure *together with the processes* that operate upon them and that utilize the analogical

properties of the structure (Freksa & Barkowsky, 1999).[34] Regarding the representation structures used in a computational diagrammatic reasoning architecture, we can distinguish between *positional* and *relational* forms of representation, as well as between *integrated* forms that use both forms of representation (Schlieder, 1998).[35]

- In *positional* diagrammatic representation structures, the two-dimensional plane (analogously the three-dimensional space) is conceptualized as a set of positions. So positional representation structures directly represent the (empty) space, i.e., the possible places in space where material objects can be. As representational medium, positional representations employ raster structures. These raster structures may either be regular rectangular structures (e.g., Khenkhar, 1991; Furnas, 1992; raster-based geographic information systems), or they may employ any other kind of positional structure (e.g., Kosslyn, 1980; Funt, 1980; Barkowsky et al., 1996).
- *Relational* diagrammatic representation structures represent objects together with the spatial relations that hold between them. So relational diagrammatic representation structures are descriptions of the structure of spatial relationships between entities in space. Relational representation structures are more suitable for representing incomplete spatial relationships than positional structures. In relational representations there is no need for specifying all spatial relations as required by a positional realization of a state of affairs. A diagrammatic reasoning approach that solely relies on relational representations is discussed by Larkin and Simon (1987). Also, the representation formats used in vector-based GISs are an example for relational structures.
- Both positional and relational representations have specific advantages. Systems that *integrate* positional and relational diagrammatic representation structures are interesting both for computational and for cognitive reasons (cf. Couclelis, 1992). Especially, the memory model for mental image processing proposed by Kosslyn (1980) has been adopted for diagrammatic reasoning (cf. Schlieder, 1998). A technical diagrammatic reasoning system that directly refers to Kosslyn's psychological model has been developed by Glasgow and Papadias (1992, see below).

In the next section I will briefly review three diagrammatic reasoning systems that exhibit positional and relational representation structures, and that are related to mental reasoning with diagrammatic representations: the DEPIC-2D system (Khenkhar, 1991) that operates on regular rectangular cell matrices, the WHISPER system (Funt, 1980) that employs a retina-like representation structure, and

---

[34] This characteristic of diagrammatic reasoning systems corresponds to the representation – process relationship discussed in cognitive modeling contexts: whether an observable phenomenon is due to a specific representation structure or whether it results from the process that operates on the structure can not be decided by analyzing an agent's behavior (cf. Section 1.4.3).

[35] *Integrated* systems using both positional and relational diagrammatic representation structures must not be confused with *hybrid* DR architectures that integrate analogical and propositional reasoning techniques (see above).

the computational imagery approach (Glasgow & Papadias, 1992) that is directly based on the mental imagery model by Kosslyn (1980).

2.5.3 Examples for Diagrammatic Reasoning Architectures

2.5.3.1 DEPIC-2D

The DEPIC-2D system (Khenkhar, 1991) is used in a hybrid spatial reasoning architecture (Habel, 1990) to visualize and inspect propositionally stored spatial knowledge. The visualization is performed on a *cell matrix*, a representation structure built up from regular *cells* each standing for a discrete section of the plane. The concept of the cell matrix is inspired by the visual buffer structure in the mental imagery model by Kosslyn (1980). In DEPIC-2D, however, the cell matrix exhibits a uniform resolution over the entire plane (cf. Section 2.3.3.1).

Several connected cells are used to represent a spatial entity, i.e., they form a *depiction* of an object. The topological relations that hold between the represented entities (i.e., punctual, linear, and areal depictions of spatial objects) are realized by the neighborhood structure of the cells. Figure 2.9 shows a cell matrix depicting two streets meeting at a corner together with two regions: one representing the immediate surrounding of the corner, the other one representing the area alongside the street.

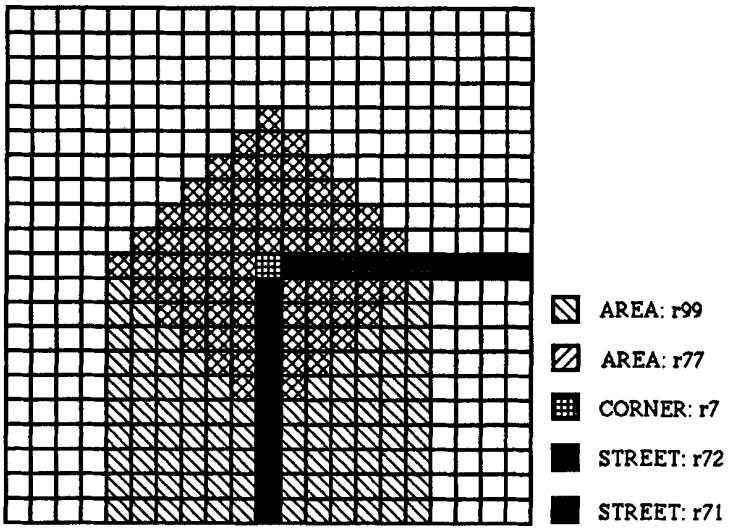

Fig. 2.9. Example of a cell matrix as used in the DEPIC-2D system (Khenkhar, 1991)

The cell matrix is operated on by neighborhood-based processes (e.g., spreading activation), which construct new entities on the cell matrix (e.g., a

region around a given position, see above), or which spatially explore the entities depicted in the cell matrix to yield a spatial relation. The purpose of this quasi-pictorial representation structure is to perform inferences with the depicted entities by constructing the scene under consideration on the cell matrix. As the inference processes performed in DEPIC-2D strongly rely on the spatial structure of the cell matrix, all main advantages of diagrammatic representations outlined above are exploited (cf. Section 2.5.1). Moreover, the reasoning processes are immediately visualized and can be directly observed.

2.5.3.2 WHISPER

A positional representation structure is also used in the diagrammatic reasoning system WHISPER (Funt, 1980). The purpose of the system is reasoning about the mechanics of unstable constellations of solid objects in a blocks world. The initial constellation to be considered is presented to the system as a raster image of a two-dimensional projection (the *diagram*). The individual entities are uniquely identified in the representation, and it is differentiated between the objects' boundaries and their interiors (see Fig. 2.10a).

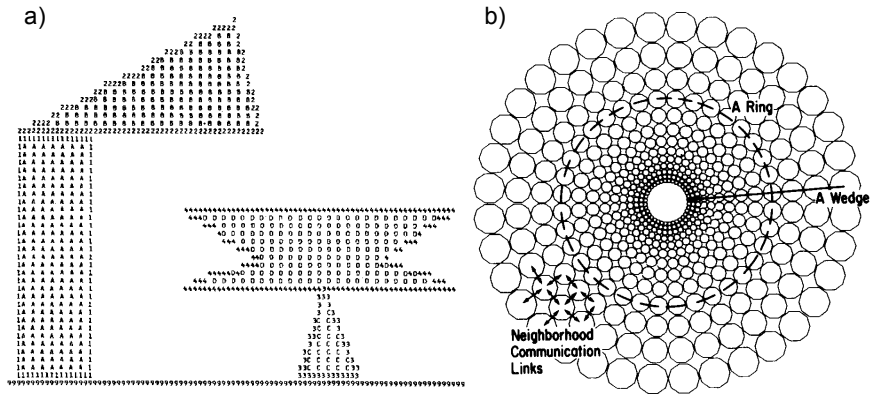

Fig. 2.10. a) Diagram representation of an initial state of a set of unstable solid objects in WHISPER (Funt, 1980); b) WHISPER's *retina*: neighboring processors are arranged on rings and wedges

WHISPER identifies unstable situations in the given scenario and predicts how and in which order the represented entities will move. The intermediate results of the reasoning process are immediately depicted in the diagram. WHISPER is a hybrid diagrammatic reasoning architecture. Besides the diagram holding the blocks world scenario, there is a *high level reasoner* and the so-called *retina*. The high level reasoner is a procedural propositional problem-solving system that contains the necessary qualitative physical knowledge about stability and motion of rigid bodies under gravity conditions.

WHISPER's retina is a positional diagrammatic representation structure that is used to 'perceive' (details of) the scene depicted in the raster image. However, the retina's structure is not rectangular. It is a circular structure built up by a number of processors which are arranged in concentric rings and on regular wedges (Fig. 2.10b). The processors operate in parallel and exchange information with their neighbors: they immediately interact with the two neighboring processors on the same ring and with the two processors of the neighboring rings in the same wedge. A superordinate processor, the *retinal supervisor* coordinates all processors in the retina. Determined by the arrangement of the processors, the retina's resolution decreases towards the periphery like in natural retinas and like in the visual buffer in the mind (cf. Section 2.3.3.1).

The position of the retina is variable with respect to the diagram that holds the scene. For reasoning purposes, information is transferred from the diagram into the retina where it is processed by *perceptual primitives*. The perceptual primitives comprise routines for focussing on the center of an entity, for checking for symmetry within one entity and for congruence of different entities, for scaling or rotating an object in the retina, for detecting connections and collisions between different entities, for assessing boundary features of objects (e.g., convexity or slope), and for finding neighboring entities with respect to an entity that is focused on.

In the retina, the tentative movement of the objects with respect to each other is simulated. After this retinal simulation, the resulting conditions are used for updating the representation in the diagram. So in the diagram the overall reasoning process can be observed step by step, and each updated representation in the diagram serves as precondition for the next retinal inference process.

2.5.3.3 Computational Imagery

Also the *computational imagery* approach presented by Glasgow and Papadias (1992) is based on a hybrid architecture. Moreover, it is an integrated architecture: it employs a positional and a relational diagrammatic structure. Computational imagery is motivated by Kosslyn's (1980) mental imagery model.[36] The intention of Glasgow and Papadias (1992) is to develop a diagrammatic reasoning architecture based on mental imagery principles that can be applied to reasoning tasks in technical contexts, for instance in chemistry or in processing geographic information.

The overall architecture of computational imagery is shown in Fig. 2.11. It distinguishes between three subsystems: the *visual representation*, the *spatial representation*, and the *deep representation*.[37] With respect to the mental imagery

[36] In essential aspects, however, the computational imagery model by Glasgow and Papadias (1992) is based on more recent neuropsychological results described in (Kosslyn, 1987).

[37] This architecture can be conceived as a *standard architecture for integrated diagrammatic reasoning systems*. It has been further differentiated for purposes of artificial intelligence and of cognitive psychology (cf. Schlieder, 1998).

conception by Kosslyn (1980) it is important to note that Kosslyn's *surface representation* (cf. Section 2.3.3.1) has been split into two subsystems: one for the *visual* aspects (*what* an object looks like, e.g. its shape, color, or texture), and one for the *spatial* aspects of mental imagery (*where* an object is located with respect to other objects; cf. Mishkin et al., 1983; Kosslyn, 1987; 1994a).

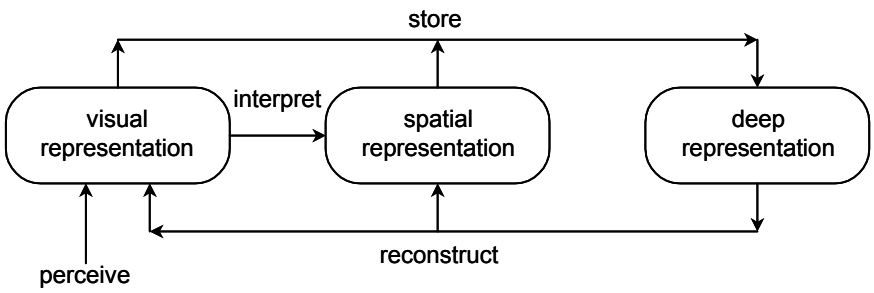

Fig. 2.11. The overall architecture of the computational imagery approach (Glasgow & Papadias, 1992)

The *deep representation* (corresponding to long-term memory in mental imagery) contains the underlying knowledge that is operated on in the visual and in the spatial representation (which together correspond to human working memory). After the reasoning process in one or both of these systems, information is restored in the deep representation. The deep representation contains propositional knowledge in a hierarchical frame structure.

The *visual representation* is the positional component of the architecture. It is realized by *occupancy arrays* that are built up from individual cells. Each cell stands for a definite spatial region.[38] The visual representation is used to encode an object's shape and size, but also the relative distances between (parts of) objects. Moreover, the cells that represent an object encode texture and color information, as well as the object's surface orientation. In contrast to mental imagery assumptions, occupancy arrays are three-dimensional, and they represent spatial entities in a viewer-independent form. Standard processes as common in computer graphics or computer vision operate on the visual representation for rotating, moving, or scaling entities, as well as for identifying volume and shape properties. The visual representation can be interpreted to extract spatial facts which are used by the spatial representation. Figure 2.12a depicts a visual representation of a molecule structure at three different levels of resolution.

The *spatial representation* in computational imagery represents qualitative spatial relationships between objects. It abstracts from exact shapes and sizes, but preserves topological information and coarse orientation information. The spatial

[38] Although this form of representation is inherently *spatial*, Glasgow and Papadias (1992) use it to encode properties that they assign to the visual part of their architecture.

representation is realized by a relational structure that is built up using multidimensional symbolic arrays. The symbolic arrays are nested to allow for hierarchical spatial representations: single array elements themselves can contain complete symbolic arrays, that describe how this array element further subdivides at a finer resolution. Figure 2.12b shows an example of a symbolic array that represents the spatial relationships between the states of a part of Europe. Observe that to capture all spatial relationships it may be necessary that an object occupies more than one array element (e.g., in the case of France or Spain). In Fig. 2.12b the cells representing Great Britain have been replaced by another symbolic array that exhibits the internal structure of Great Britain.

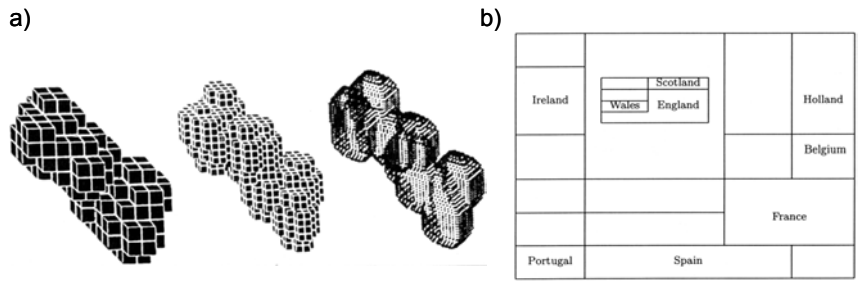

Fig. 2.12. a) Example of a visual representation. A molecule structure is shown at three different levels of resolution (Glasgow & Papadias, 1992: 370; figure inverted). b) Example of a spatial representation. Embedded symbolic arrays represent qualitative spatial relationships (updated figure from reprint in Glasgow et al., 1995: 455)

Besides processes for transferring information between the spatial representation and the deep representation, there are basic functions for placing, detecting, deleting, and moving objects in the spatial representation, as well as for testing for adjacency. Moreover, for pattern recognition tasks, there are processes that realize a concept of attention control (like in mental imagery). From these basic processes more complex functions can be generated, especially domain-specific routines (e.g., for small-scale space or geographic scenarios).

2.6 Summary

In this chapter I have first reported on metaphors for spatial knowledge representation and processing in the human mind. It has shown that the cognitive map metaphor is misleading, since empirical investigations reveal that spatial knowledge does not resemble a uniform, coherent representation in the mind. However, when assuming that spatial representations are constructed in mind when needed, we can still think of a spatio-analogical structure used as a representation medium for spatial knowledge. Under this perspective, the construction processes that are used for

envisioning a spatial state of affairs can be identified as the sources of the distortions in mental spatial knowledge documented by the psychological investigations (e.g., rotation or alignment). The mental apparatus used for mental spatial reasoning resembles a *cognitive atlas* (partial representations usable for the mental construction task are stored in an organized manner) or rather a (human) GIS (which accounts for the operational aspect of spatial mental knowledge processing). The rubber sheet analogy of spatial mental representation characterizes the distorted reconstruction on the spatio-analogical representation medium in the mind.

The constructive characteristic of mental spatial knowledge processing is also pointed out by the conception of (spatial) mental models. However, spatial mental models can be visual (i.e., realized in mental images) or they can be spatial (in the sense of an analogical mental representation) without being visual. As I follow the conception developed by Kosslyn (1994a) in this work, I will assume that the mental construction of geographic knowledge is done using visual mental images. So spatial mental models in the context of this thesis are conceived as being visual.

Visual mental images are constructed in working memory. I have shown how working memory differs from long-term memory. Working memory processes information that either stems from the senses (i.e., the visual sense in our case) or that is retrieved from long-term memory. The working memory subsystems that are active in visual mental imagery are the visuo-spatial scratchpad (which serves as the spatio-analogical representation medium for the mental construction) and the central executive (which controls the image generation and exploration processes and which maintains the image).

For long-term memory, it has been distinguished between declarative (explicit) and non-declarative (implicit) memory. The geographic knowledge considered in this thesis is of the explicit type: it is knowledge consciously available to the individual person who owns it (semantic memory). Knowledge in long-term memory is assumed to be organized hierarchically. Retrieval from long-term memory requires search processes conceived as spreading activation processes in the brain. With respect to mental imagery, long-term memory must be differentiated in activated and non-activated parts. Activated parts contain knowledge that has been retrieved and that currently participates in the imagery process. As such, it is part of working memory. Non-activated long-term memory has not been retrieved but it may contain further important information that is used later on in the imagery process.

Mental images are reconstructions from facts from long-term memory augmented by relations required for the image generation process. Mental images are built up using pictorial and propositional information. With the structures employed in diagrammatic reasoning they share the advantages of spatio-analogical representations. The implicit encoding principle allows for inferring pieces of information from the mental image. These pieces of information are not explicitly represented in memory, but they emerge when available relations are visualized.

I reported on AI techniques for representing qualitative spatial knowledge resembling the lean knowledge available from long-term memory in the mental imagery process. However, the emphasis in reasoning with qualitative spatial knowledge is more on logic reasoning techniques than on reasoning with diagrams. I have motivated that spatial knowledge can be ordered according to the expressiveness of the types of knowledge with topological relations as the most basic type up to fully metrically specified representations as dealt with in computational geometry. Within this thesis I will first of all consider topological and orientation information. Rough shape information is also used. However, I will not use qualitative representation techniques for shape knowledge, but I will use explicit (pictorial) shape descriptions instead. In mental imagery processes both propositional and pictorial knowledge is involved. I will not deal with distance knowledge in this work.

From an AI perspective, the model developed here is a diagrammatic reasoning system. A diagrammatic representation medium is used to describe the content of the visual buffer, and images are built up based both on propositional and on pictorial representations retrieved from long-term memory (so it is a hybrid system). Besides the positional description of the visual buffer content, there are relational structures employed to hold the working memory representation needed to generate and maintain the image proper. Thus, the model to be developed is an integrated representation system.

Besides the diagrammatic reasoning systems by Funt (1980) and Khenkhar (1991), which use ideas of mental imagery, I presented three modeling conceptions directly related to mental image processing: the two systems by Kosslyn (1980; 1994a) and the architecture by Glasgow and Papadias (1992). Whereas the 1980 Kosslyn model comes as an implemented system that uses a positional representation structure as visual buffer, his 1994 model describes a conceptional approach that accounts for the complex interaction processes between the several functional components operating in the imagery process. The implemented computational imagery system presented by Glasgow and Papadias (1992) treats imagery as a technical problem solving paradigm. As it distinguishes between spatial and visual aspects of mental imagery, it operates with two distinct types of representations.

The architecture presented in this thesis is a computational model that uses a relational (vector graphic like) description of the positional content of the visual buffer. The long-term memory components involved in the imagery process (i.e., the components which form the non-pictorial working memory content) are in an analogical (i.e., graph-like) representation of propositional and pictorial facts used for the image generation proper. Thus, unlike the architecture by Glasgow and Papadias (1992), the approach taken here is intended as a cognitive science contribution (rather than a technical solution), and it relies on the functional conception of Kosslyn's 1994 model. In the following chapter, the MIRAGE model will be developed.

3 MIRAGE – Developing the Model

In this chapter I will develop the MIRAGE[39] model for representing and processing lean geographic knowledge in the human mind. The principal architecture described in Section 1.3 is elaborated, and the components and their operations are described in detail.

First, I will describe the characteristics of the model. The description is structured according to the distinction of human memory in *long-term memory* and *working memory*. Second, I will discuss two contrasting ways of constructing and evaluating the working memory representation. I will elaborate a synthesis of these two contrasting possibilities and present an overview of the model's architecture. Third, I will define the types of entities and spatial relations used in MIRAGE. Fourth, I will characterize the individual components of the model and explain the processes operating on them.

3.1 Characteristics of the Model

In the following I will define the characteristics of the long-term memory and working memory components of MIRAGE. Geographic knowledge available in long-term memory is used to construct a working memory representation of a spatial configuration. The working memory representation is evaluated to obtain the required spatial relation (cf. Fig. 3.1). I will distinguish between psychologically grounded (or at least plausible) characteristics and additional characteristics introduced for modeling purposes (cf. Section 1.4.2). The former are due to empirical results gained from psychological investigations, whereas the latter are needed for the technical realization in form of an implemented computational model.

[39] MIRAGE stands for Mental Images in Reasoning About Geographic Entities.

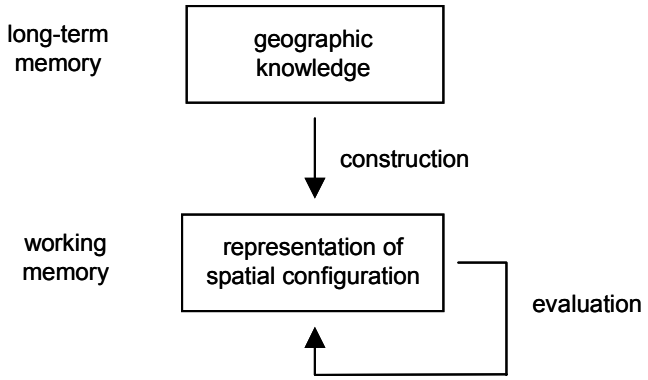

Fig. 3.1. Construction of working memory representation based on geographic knowledge from long-term memory (cf. Section 1.2)

3.1.1 Long-Term Memory

Geographic knowledge stored in long-term memory is often *underdetermined* or *lean*. On the one hand pieces of geographic information that are needed in a specific context are often not explicitly stored (*scarce knowledge*). On the other hand knowledge available from long-term memory that can be utilized for inferring missing information is often of the qualitative type and not represented in a precise, metrically specified form (*coarse knowledge*).

MIRAGE deals with scarce knowledge and also mainly with qualitative knowledge.[40] Mental reasoning processes about geographic configurations can also be based on precise information (for instance knowledge about metrical distances), or answering a geographic question can consist of just retrieving the information from memory (when the required information has been explicitly acquired before). Since MIRAGE's concern is the *construction* of spatial knowledge representations in mind, the discussion is restricted to lean geographic knowledge.

Geographic knowledge stored in long-term memory is *fragmented*; it is not represented in a single, coherent representation structure like a map. Spatial knowledge representations in long-term memory may be spread over several partial representation structures, which may or may not be connected with each other in a coherent way. Especially when taking into account that the parts of spatial knowledge may stem from different sensory modalities (e.g., visual, haptic, auditory, kinesthetic, etc.), it is plausible that spatial knowledge is spread over different representation structures (cf. Tversky, 1993; Hirtle, 1998).

For modeling purposes I assume that geographic knowledge in long-term memory is highly fragmented. I make this assumption to avoid having to deal with

[40] An exception to using only qualitative knowledge will be when explicit representations of shape information are involved in the mental construction task (for example the shape of an extended geographic entity like a state or a continent).

partial spatial knowledge structures of different complexity and of different degree of aggregation. Although also partially coherent representation structures of different size and complexity may be stored in long-term memory, this case is not considered in MIRAGE. MIRAGE's main intention is to model the construction of mental geographic knowledge representations from available pieces of information from long-term memory. The basic entities of this fragmentary long-term memory representation are the *spatial knowledge fragments*: geographic knowledge represented in long-term memory is modeled using spatial knowledge fragments.

Geographic knowledge is *hierarchically organized* in long-term memory. It is known from empirical findings that human memory is organized in hierarchical structures (see Section 2.1.2). These hierarchies become effective in accessing the knowledge represented in long-term memory: for instance, when the organizational structure in long-term memory is conceived as tree-like[41], accessing knowledge about a state of affairs means traversing the tree down to the level of detail needed for the problem to be solved.

For the purposes of MIRAGE, I assume that the long-term memory representation formed by spatial knowledge fragments exhibits a single representation structure. This assumption is made for not having to deal with retrieval-related problems, for example, that some information cannot be accessed due to structural gaps in the long-term memory representation. I assume that the hierarchical structure in long-term memory is encoded in a *directed graph*.

Hierarchy applies both to the *type* of spatial knowledge represented and to the *degree of coarseness*: the spatial knowledge types are assumed to be ordered hierarchically (for instance, topology ranking higher than orientation information), and coarser spatial relations are further up in the hierarchy than finer ones (for example A is north of B ranking higher than A is north-northwest of B). Thus, the organization is in the form of a two-dimensional hierarchical structure. This twofold hierarchy is effective in retrieving information from long-term memory. For modeling purposes, I assume that *spreading activation* processes operate on the graph structure, and that access of pieces of information is done in a sequential manner. Spatial knowledge fragments are extracted from long-term memory according to their hierarchical organization in the graph structure.

---

[41] In general, the conception of human memory as being tree-like is oversimplifying. It seems to be more sensible to conceive of memory structures as being organized in the form of overlapping trees or *semi-lattices* (cf. Alexander, 1982). A specific hierarchy can be represented as a tree-like structure, but in general a *graph structure* has to be assumed. Graph structures allow for spreading activation processes suggested by psychological experiments (e.g. in associative priming, see Meyer & Schvanefeldt, 1971; Ratcliff & McKoon, 1981). These processes can be simulated in artificial neural networks (Rumelhart et al., 1986).

3.1.2 Working Memory

Since knowledge about geographic spaces frequently is not stored in a ready-made form, geographic knowledge is *constructed* in working memory from available pieces of information when needed. The purpose of the geographic knowledge representation constructed in working memory is to provide a piece of spatial knowledge that is not explicitly represented in long-term memory.

In the context of this work, the goal of the knowledge representation built up in working memory is to decide about a specific geographic question. Although spatial knowledge in long-term memory is lean, the working memory representation will provide determinate results with respect to the question to be answered (cf. Section 1.1.2). To compensate for scarce knowledge in long-term memory, the working memory representation is complemented by *default components*.

The working memory representation is built up using *mental imagery*. I argued in favor of the assumption that mental images are used for constructing the working memory representation (cf. Section 1.2.4). The mental image representation is *visualized* in a *quasi-pictorial medium* in the mind. Available pieces of knowledge are *successively integrated* in the image leading to a *stepwise refined* result. The mental representation constructed in this way is inspected by *image inspection* processes.

For the purposes of MIRAGE, I assume a pictorial representation structure (*visual buffer*) in which the image-like representation is built up, maintained and modified. However, MIRAGE is not committed to a specific physical realization of this representation structure (i.e., whether it is more suitably modeled in a positional or in a relational structure, or what the physical properties regarding size, varying resolution, etc. are). The mental image is constructed by appropriate processes in the visual buffer; these visualization processes operate in a successive manner. As the contents of mental images fade out over time, mental images have to be periodically refreshed by image maintenance processes. Although the dynamic characteristics of these maintenance processes are beyond the scope of MIRAGE, still some kind of (non-pictorial) backup representation has to be considered in working memory. This representation is used to refresh the image contents without completely rebuilding the image from long-term memory.

Visuo-spatial knowledge is *successively integrated* into the mental image representation. Mental images are built up in working memory in a step-by-step manner. With respect to the construction of geographic knowledge, the spatial knowledge fragments taken from long-term memory have to be arranged to form a coherent representation in working memory. Spatial knowledge fragments are integrated in the working memory representation according to the order in which they are retrieved from long-term memory. The retrieval is sequenced by the hierarchical order in long-term memory. Subsequent mental systems involved in the image generation process operate in parallel using partial result yielded by preceding systems.

For the purposes of MIRAGE, I assume that the mental image representation is stepwise refined. Although in principle image construction processes can proceed

in a highly parallel way, the conceptualization in terms of discrete (sub-)systems pursued here requires that image construction proceeds in a step-by-step manner. The image construction starts at a coarse level (i.e., it contains only few pieces of information from long-term memory) and is successively enriched by further facts. Missing information is complemented by default knowledge. Default knowledge can be replaced later in the image construction process, when appropriate pieces of knowledge are retrieved from long-term memory. So default components are used to compensate for both missing information from long-term memory (be it that knowledge is coarse, be it that it is completely missing) and missing information due to an early stage of the construction process.

The advantage of this stepwise refinement characteristic is that it enables an *anytime behavior* of the system. Under resource restrictions (for instance under time pressure) it is possible to produce fast albeit rough results when required. By refinement, more and more pieces of knowledge are integrated during the image construction process, and knowledge already contained in the image is further specified according to more detailed information retrieved from long-term memory.

Before developing the architecture of MIRAGE, it has to be considered how the mental knowledge representation constructed in working memory is used for answering the geographic question at hand.

3.1.3 Evaluating the Working Memory Representation

The purpose of the construction of the working memory representation is to obtain spatial information not explicitly contained in long-term memory. So far it is not clear how specific the constructed working memory representation will be in comparison with the lean long-term memory representation used, and how the representation in working memory is evaluated to gain the intended spatial result. How is the leanness of the long-term memory representations compensated in working memory, and how is the working memory representation used?

In Section 1.4.3 I motivated that in computational models, cognitive phenomena can be rebuilt both by representation structures and by processes. Thus, it may be indistinguishable, whether a cognitive characteristic relies more on a process or more on a representation structure. With respect to the question how the required geographic relation is obtained from a constructed working memory representation, two contrary possibilities are conceivable:

1. Either the representation built up in working memory is *completely specified* with respect to the required spatial state of affairs. In this case the required solution is given by the *representation structure*, and the result only has to be inspected by an appropriate process.
2. Or the working memory representation is still *underspecified*, i.e., it still does not contain the required relation. In this case the solution is provided by the *process* operating on the constructed working memory representation.

I will now first discuss the implications, advantages, and drawbacks of the two possibilities. Based on this discussion, the overall architecture of MIRAGE will be developed below.

1. When the working memory representation is *completely specified*, default knowledge is used in the *construction* of the working memory representation. Lean knowledge from long-term memory is complemented by default components to build up a working memory representation that explicitly contains the required spatial relation. The result is a determinate working memory representation that directly can be utilized by a simple inspection process that reads the required result off the representation (see Fig. 3.2a). From an AI perspective such a representation structure can be realized through a positional diagrammatic representation structure (e.g., Funt, 1980; Khenkhar, 1991; Glasgow & Papadias, 1992).

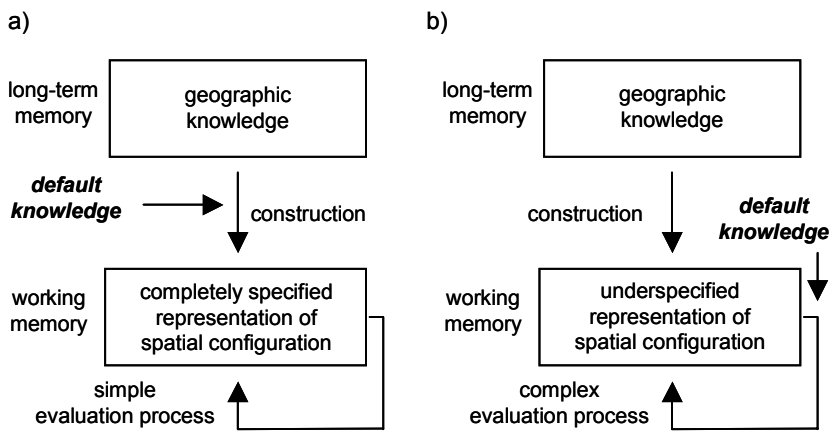

Fig. 3.2. Two possibilities of constructing and evaluating a working memory representation

2. The working memory representation on the other hand may be *underspecified* when only the available information from long-term memory is integrated into the working memory representation. In this case the working memory representation is still lean, like the corresponding long-term memory representation the used pieces of knowledge stem from. To evaluate this representation structure, default knowledge has to be provided by the process that makes use of the constructed working memory representation. In this variant, the inspection process has to complement the working memory representation by default knowledge. Therefore, the inspection process has to be much more complex than an inspection process that simply reads off a readily represented relation (see Fig. 3.2b). A representational structure of this type can be modeled using relational representation structures (e.g., Larkin & Simon, 1987; Habel et al., 1995).

Both variants have advantages and drawbacks. In the completely specified working memory representation, the explicitly represented knowledge is only partially based on knowledge taken from long-term memory. As a consequence, this work-

ing memory representation has to be partially corrected when further and more specific pieces of information are integrated into the representation.[42] To distinguish between aspects that are due to determinate information from long-term memory and information that is only contained as a result of the construction process, additional knowledge about the pieces of information built in the representation so far would be necessary, which would need administrational effort. The main advantage, however, is the simplicity of the inspection process that can be used to explore the result of the visualization.

In the underspecified variant of the working memory representation only determinate information is explicitly represented. Refinement of the working memory representation can easily be done, as already built-in information easily can be exchanged by more specific pieces of knowledge, or it simply can be supplemented by additional information. However, the main work in achieving the required result in working memory is left to the inspection process. Determinacy has to be provided while inspecting the underdetermined working memory representation.

The architecture presented in the following section makes use of the advantages of both discussed variants by combining them in a hybrid model. By this synthesis, some of the drawbacks can be avoided.

3.2 MIRAGE – Outline of the Model

MIRAGE's components will be described according to the detailed distinction of human memory in three memory components: *non-activated long-term memory*, *activated long-term memory*, and *short-term memory* (cf. Section 2.2.3). Long-term memory is differentiated in whether its content has been activated during processing or not. Working memory is partially constituted by activated long-term memory structures and partially by short-term memory. Figure 3.3 gives an overview of the architecture of MIRAGE.

Lean geographic knowledge used for the construction of customized working memory representations of a geographic state of affairs stems from non-activated long-term memory. This knowledge is organized hierarchically to allow for access processes adapted to the problem to be solved. When pieces of information are requested in the context of a working memory representation to be constructed, this non-activated long-term memory structure is accessed (i.e., parts of it are activated). For reasons of representational and processing clarity, spatial knowledge fragments are extracted from non-activated long-term memory by a specific *access process* in MIRAGE. The fragments accessed this way are used to construct an activated long-term memory representation.[43]

---

[42] Another possibility would be to completely rebuild the representation in every refinement step. This variant, however, does not seem viable for time-economical reasons, and it is not in line with what is known about mental operations on images.

[43] In terms of distributed neural structures, these access and construction processes have to be conceived as a step-by-step activation of hitherto not activated memory areas by

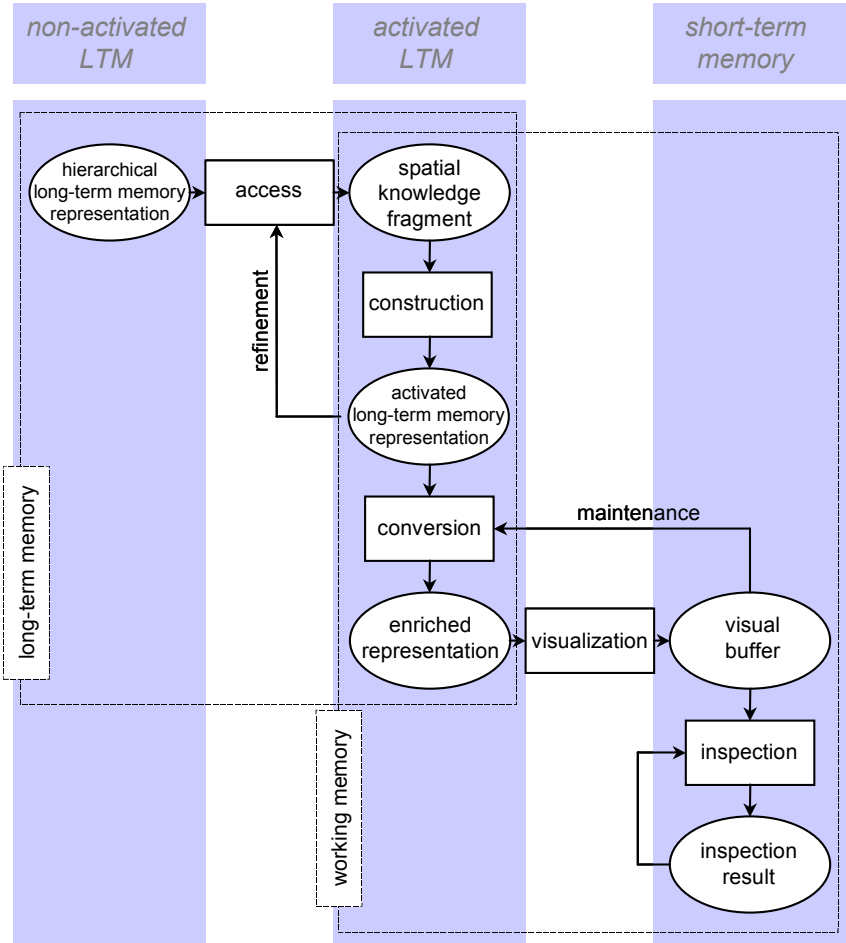

Fig. 3.3. Outline of the architecture of MIRAGE

The result of the access process and the construction process is an activated representation in long-term memory (cf. Fig. 3.3). As it is just a selection of information stored in non-activated long-term memory, this activated long-term memory representation is still as underdetermined as the original contents in non-activated long-term memory. For visualization in a quasi-pictorial medium, however, it has to be converted into a more specific form of representation.

As the reconstruction of geographic knowledge in working memory is done using visual mental imagery in MIRAGE, a visual representation medium or *visual buffer* is used. For the purpose of the visualization, the still lean working mem-

neural spreading activation processes. So these processes may exhibit no explicit extraction and no physical transfer between brain structures.

ory representation has to be further specified by additional spatial information. This is done by a *conversion process* (comparable to the *categorical to coordinate conversion* subsystem in Kosslyn's 1994 model, cf. Section 2.3.3.2). The result of the conversion is an *enriched working memory representation* in MIRAGE. This enriched working memory representation is still part of activated long-term memory as it combines two types of knowledge taken from long-term memory.[44]

The enriched working memory representation is now used for the visualization in the visual buffer. The representation in the visual buffer is inspected by an *inspection process*. Since it is a visualization of a representation that previously has been made determinate (i.e., by converting the underdetermined working memory representation into the enriched representation) visualization and inspection of the resulting representation are quite simple.

Image representations held in the visual buffer fade out over time and therefore need to be periodically refreshed. As such, visual mental images are part of short-term memory, which periodically is *maintained* by the underlying representation constructed in activated long-term memory (cf. Fig. 3.3).

Two main characteristics can be exhibited when regarding this architecture:

1. the *anytime characteristic* of processing geographic knowledge and
2. the treatment of lean geographic knowledge while at the same time producing determinate results.

Both characteristics will be explained in the following.

1. It is seen in the depiction of the architecture that three subsystems[45] are identified as working widely independently of each other, each one possessing its own feedback structures (see Fig. 3.4): first, the *long-term memory activation* system, which accesses hierarchical long-term memory knowledge to construct the (still lean) working memory representation; second, the *visual mental image construction* system comprising the conversion process for obtaining the enriched representation and the visualization process to evoke the image in the visual buffer; and third, the repeated *inspection* of this visualization.

   As can be observed in the figure, each of these three subsystems has its own feedback loop: the *refinement* loop to activate further pieces of knowledge from long-term memory, the *maintenance* loop to rehearse the visualization, and the *inspection* loop, which enables repeated inspection of the visualization. Due to these three subsystems, the processes subsequent to the activation of the long-term memory representation can already start operating on quite rough a representation built up so far. The conversion and visualization processes can proceed while further pieces of knowledge are accessed in long-term memory and

[44] Where does the employed default knowledge stem from? It is assumed in MIRAGE that it is provided by the conversion process. I.e., there is no separate knowledge representation structure in long-term memory that holds default knowledge; rather, it is the task of the conversion process to transform lean knowledge into the enriched representation.

[45] These subsystems will be referred to later on again when describing the representation structures and the processes operating between them in more detail. At present it is important not to confuse them with the three memory structures displayed in Fig. 3.3.

are integrated into the activated long-term memory representation. In the same way, the visual buffer can be inspected as soon as the first rough construction has been visualized. So the working memory representation is stepwise refined, while preliminary results can already be used by subsequent processes.

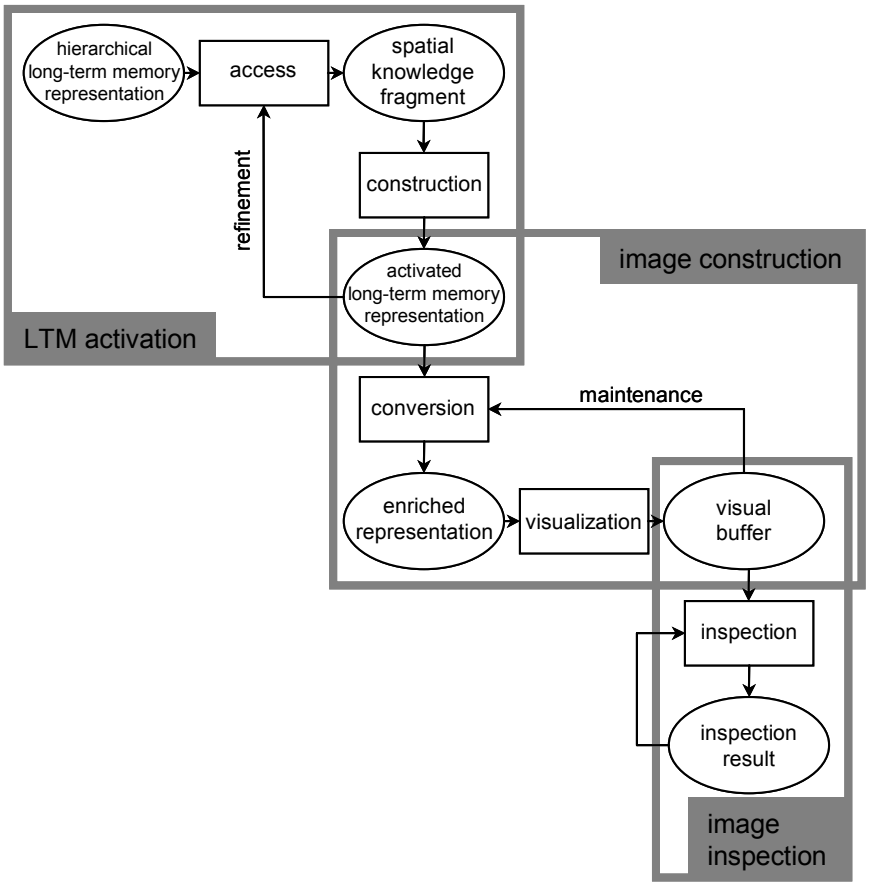

Fig. 3.4. The three subsystems *long-term memory activation*, *visual mental image construction*, and *image inspection*

2. The problem of whether the working memory representation is lean or completely specified (cf. Section 3.1.3), i.e., whether default knowledge is used in the construction of the working memory representation or only induced by a complex inspection process, is solved by the architecture shown in Fig. 3.3 as follows: the working memory representation is constructed in a two-step manner. In a first step, a lean representation is constructed based on spatial knowledge fragments taken from non-activated long-term memory. In a second step,

this underdetermined representation is complemented by default knowledge. This complemented representation is visualized in the visual buffer and inspected by a simple inspection process.

Due to this two-step processing, the main disadvantages of the two principal possibilities discussed above (cf. Section 3.1.3) are avoided. There is no administrative effort needed in the first step, since the representation in working memory for the present remains lean. During the subsequent visualization, there is no additional information regarding the status of the image's components needed, since the enrichment is only done for just one visualization turn. So after the enriched representation has been visualized in the visual buffer, it is dropped and rebuilt for the next turn of visualization, possibly after including further spatial knowledge fragments and thus making the construction more specific.

3.3 Types of Entities and Spatial Relations in MIRAGE

Before describing MIRAGE's subsystems in detail in Section 3.4, I will define the types of geographic entities and spatial relations that can be represented and processed in the model.

3.3.1 Entities

The ontological type of a geographic entity represented in MIRAGE is either punctual or areal, i.e., linear objects are not considered. Extended entities are expected to be simply connected (i.e., not spread over several distinct entities), and they contain no holes, which both are common assumptions in qualitative spatial reasoning (cf. Section 2.4). Generally, the ontological type of entities is not explicitly represented in long-term memory. Rather, the ontological type of objects is assigned during the image construction, which allows for varying the ontological type of geographic entities according to the problem to be solved. For example, a lake may be either conceived as punctual or as extended, depending on whether its location with respect to other distant points is considered, or whether it is regarded as an areal object containing other objects, for example islands. To uniquely refer to the entities, every entity has its own label.

3.3.2 Relations

In MIRAGE, topological and orientation relations (cardinal directions) are represented as binary relations between pairs of entities. For extended spatial entities, shape information is represented in form of polygonal curves. All spatial relations may be represented at different levels of resolution. The resolution a spatial relation is represented at refers to suitable interpretations of the relation. For example, an orientation relation like 'A is west of B' has to be interpreted differently, depending on whether it refers to a distinction of 16 or of just 4 cardinal directions

(cf. Section 2.4.2). The different resolution levels of the spatial relations are denoted by numerical values starting from 1 for the lowest resolution up to some n for the highest resolution.

Topological relations between two spatial entities are represented according to the 8 relations described in Section 2.4.1 at the highest resolution level possible. A lower resolution level only allows for 5 topological relations. These 5 relations are roughly according to the 5 relations in Egenhofer's medium resolution topology representation (Grigni et al., 1995). Table 3.1 shows the possible topological relationships at both resolution levels. In Section 2.4.1 I showed which of the topological relations may hold between either combination of punctual and extended geographic entities (cf. Table 2.1).

Table 3.1. Topological relations in MIRAGE

resolution	number of relations	topological relations
1 (low)	5	disjoint, tangent-or-overlaps, equal, in, contains
2 (high)	8	disjoint, tangent, overlaps, equal, in-at-border, in, contains-at-border, contains

As orientation relations, cardinal directions can be represented in MIRAGE at 4 different resolution levels. All levels comprise a relation for a neutral direction relation accounting for two objects being either (roughly) at the same place or for one object being centered within another (extended) object. The lowest resolution level just allows for a dichotomic distinction between north – south or east – west, which together with the neutral direction relation results in three possible relationships between two objects. The highest resolution level describes sixteen cardinal directions plus the neutral direction relation. Table 3.2 gives an overview of the cardinal direction relations used in MIRAGE. A cardinal direction can be represented between two point entities, between a point entity and an extended entity, and between two extended entities. For extended entities, the respective centers of gravity are used to determine a cardinal direction.

Table 3.2. Cardinal directions in MIRAGE

resolution	number of relations	cardinal directions
1 (lowest)	3	N / S / neutral ; E / W / neutral
2	5	N / E / S / W / neutral
3	9	N / NE / E / SE / S / SW / W / NW / neutral
4 (highest)	17	N / NNE / NE / ENE / E / ESE / SE / SSE / S / SSW / SW / ... / neutral

Shapes of spatially extended entities are represented through polygonal curves. Like the other relations described above, also the shapes of entities may be represented at different levels of granularity (a person may have both a very rough and a finer representation of an extended entity in her mind). In MIRAGE, shapes can be represented at two different levels of granularity. Similar to resolution encoding in topological relations, the two different granularities of shape descriptions are denoted by 1 for the lower resolution and 2 for the higher resolution. Figure 3.5 shows the shape representation of California at two different levels of resolution as an example.

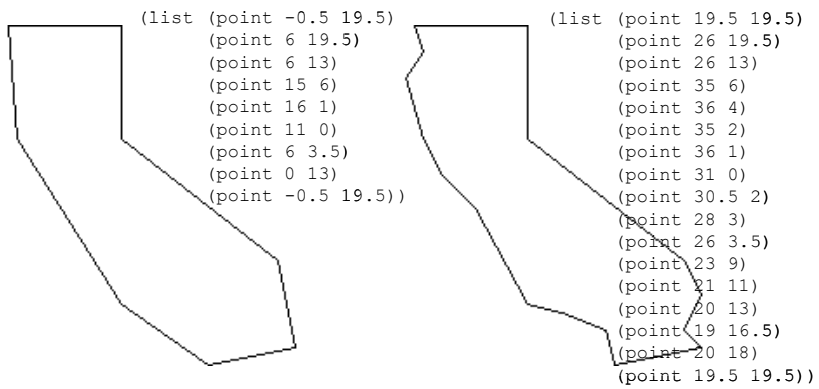

Fig. 3.5. Shape representation of California at two different resolutions

3.4 Subsystems, Structures, and Processes

In this section I will describe MIRAGE's three subsystems *long-term memory activation*, *visual mental image construction*, and *image inspection* in more detail (cf. Fig. 3.4). I will define the representation structures, explain the processes operating on them, and give examples to illustrate MIRAGE's dynamic behavior. Also, the relation to Kosslyn's (1980, 1994a) models is exhibited. The visual mental image construction subsystem is more thoroughly focused on in Chapter 4.

3.4.1 Long-Term Memory Activation

3.4.1.1 Spatial Knowledge Fragments

First of all, *spatial knowledge fragments* are defined. Spatial knowledge fragments serve as the elementary units for representing pieces of geographic knowledge in MIRAGE.

Definition: A *spatial knowledge fragment* is an n-ary spatial relation between geographic entities. The geographic entities involved are uniquely referred to by identifiers. The relation is annotated (a) by information about the type of spatial knowledge and (b) by the degree of resolution that is suitable for its interpretation.

The definition of spatial knowledge fragments as n-ary spatial relations comprises unary relations, which are used to express *properties* of spatial entities. In MIRAGE, the shape descriptions of extended spatial entities are unary relations. This relation defines a pictorial property of the object it represents; it is given by its polygonal description. Spatial relations between two or more geographic entities are captured by binary, ternary, etc. relations. In MIRAGE, binary relations are topological and orientation relations. An example for a ternary spatial relation (not realized in MIRAGE) is the spatial *betweenness* relation, representing that an entity A is located between two other entities B and C.

In principle, arbitrary spatial relations can be expressed in spatial knowledge fragments. Especially, they can cover every type of spatial knowledge worked with in (qualitative) spatial reasoning (cf. Section 2.4). So this basic form of representation can be used to cover the whole range of types of spatial knowledge from topological relations to completely metrically specified geometric knowledge. However, for the concerns of this work, coarse (qualitative) spatial relations are most interesting.

The annotations claimed in the definition specify the spatial knowledge fragment with respect to its relational type (i.e., whether it is topological knowledge, orientation knowledge, or a shape representation) and with respect to its degree of resolution. The annotation of the degree of resolution is sensible for representational reasons, since it seems plausible that a piece of spatial knowledge in the mind comes together with information about its accuracy. For example, if someone has stored in her mind that a location A is north of another location B, it is likely that the person also knows whether this north-of relation may be interpreted as the direction between east and west, between north-east and north-west, or maybe whether it is just meant as opposition to a southern direction (i.e., referring to a sector of 180 degrees).

The use of unique identifiers as references to the geographic entities involved guarantees that different knowledge fragments describing properties of and relations between identical spatial entities can be referred to. Thus, this forms the basis for the representation of complex structures of underdetermined spatial knowledge in long-term memory, which can be conceived as a graph structure. This property is needed in this work for the symbolic cognitive modeling paradigm used. In neural structures, however, the correspondence between various pieces of spatial knowledge is provided by other mechanisms, for example excitatory connections between neural sub-structures that encode knowledge about common entities.

By means of spatial knowledge fragments, the hierarchically organized long-term memory representation for lean geographic knowledge can now be defined.

3.4.1.2 The Hierarchical Long-Term Memory Representation

Definition: The *hierarchical long-term memory representation* is a directed graph structure built up by spatial knowledge fragments. Its nodes are formed by the identifiers of the geographic entities involved, and its edges denote the spatial relations that hold between these entities. The edges are annotated by the type of relation that is encoded, and they carry information about the granularity of the spatial relation.

An example of a hierarchical long-term memory representation is depicted in Fig. 3.6. This example has been constructed according to the 'Reno – San Diego' scenario introduced in Section 1.1.2. The nodes are formed by the geographic entities 'U.S.', 'Nevada', 'California', 'Reno', and 'San Diego'. Binary spatial relations that hold between two entities (i.e., topological or orientation relations in MIRAGE) are given by the edges that connect the respective entities. Unary relations describing properties of objects (i.e., shapes of extended entities in MIRAGE) are attached to the nodes they refer to.[46]

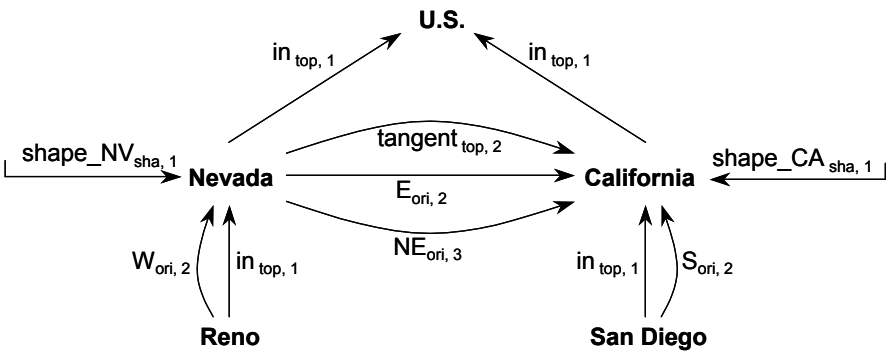

Fig. 3.6. Hierarchical organization of spatial knowledge fragments in long-term memory

---

[46] Note that the graph depicting the exemplary hierarchical long-term memory representation (Fig. 3.6) is just a visualization for illustration purposes. The hierarchical long-term memory representation also could have been given in the following way:

in $_{top,\ 1}$ (Reno, Nevada);	E $_{ori,\ 2}$ (Nevada, California);
in $_{top,\ 1}$ (San Diego, California);	W$_{ori,\ 2}$ (Reno, Nevada);
in $_{top,\ 1}$ (Nevada, U.S.);	NE $_{ori,\ 3}$ (Nevada, California);
in $_{top,\ 1}$ (California, U.S.);	shape_NV$_{sha,\ 1}$ (Nevada);
tangent $_{top,\ 2}$ (Nevada, California);	shape_CA $_{sha,\ 1}$ (California);
S $_{ori,\ 2}$ (San Diego, California).	

Especially the spatial arrangement of the entities and relations in the graph depicted in Fig. 3.6 must not be interpreted (e.g., 'U.S.' being located at the top of the figure has no meaning with respect to the contents represented in this representation structure).

The spatial relations contained in the hierarchical long-term memory representation are further specified by two annotated indices. The first index refers to the type of spatial relationship: 'top' denotes topological relations, 'ori' orientation relations, and 'sha' shape descriptions. The second index denotes the granularity at which the spatial relation is represented (cf. Section 3.3.2). Both kinds of annotations will be further explained in the following.

As discussed above (cf. Section 3.1.1), a twofold hierarchical order is encoded in this representation: one hierarchy is defined by the types of spatial relations, the other one is given by the levels of granularity of the relations. The order in which spatial knowledge fragments can be accessed in long-term memory depends on these hierarchical structures:

- With respect to the hierarchy of the types of spatial relations, the order of mathematical expressiveness known from qualitative spatial reasoning is used (cf. Section 2.4). Thus, with respect to the three types of spatial relations employed in MIRAGE (i.e., topology, cardinal directions, and shapes), topology is the most basic type of spatial knowledge, followed by orientation knowledge. Shape information, which is given as pictorial descriptions rather than in a qualitative form, is considered most complex. The idea is that types of knowledge that are less complex are accessed prior to those which are more expressive. However, from a psychological point of view, it is not clear beforehand, whether this hierarchical order may be suitable or not. Moreover, it may differ among individuals, among specific situations, or it may depend on how some piece of knowledge has been acquired. In MIRAGE, varying types of hierarchical structures can be described and tested.
- The second hierarchy concerns the granularity aspect of the represented knowledge fragments as defined in Section 3.3.2. Lower indices denote coarser granularities, whereas higher indices stand for finer resolutions. For example, 'ori, 2' denotes an orientation relation that distinguishes between five orientation relations (i.e., north, south, east, west, and neutral), whereas 'ori, 3' denotes the next finer resolution distinguishing between eight cardinal directions plus the neutral relation. In MIRAGE, it is assumed that the order of access of spatial knowledge fragments in long-term memory depends on the resolution level a piece of information is encoded at. Thus, coarser relations will be accessed prior to finer ones. More detailed pieces of information are retrieved later from long-term memory and are used for refining the working memory representation.

3.4.1.3 The Access Process

The two hierarchies realized in the long-term memory representation are used in accessing and extracting the spatial knowledge fragments required to build up the activated long-term memory representation. So the problem is, what are the relevant fragments for a given spatial question? Besides by the hierarchical structure, the access process is controlled (1) by the entities (i.e., the nodes of the graph) that shall be spatially set in relation to each other, and (2) by the type of spatial relation required.

To further illustrate this issue, the example of the relative orientation between Reno and San Diego, together with the hierarchical long-term memory representation shown in Fig. 3.6 will be considered again. The spatial question to be decided is the relative orientation between the two locations, i.e., all information about the two entities 'Reno' and 'San Diego' that are appropriate for reconstructing the relative orientation between them have to be exploited by the access process. The access will be performed sequentially in accordance with the hierarchical structure encoded in the long-term memory representation.

The access process is modeled as a bidirectional graph search that starts from either of the two nodes that are to be spatially related with each other. This bidirectional search process is performed until an orientation relation is found that connects two paths starting from the two nodes under consideration. When two nodes are related by more than one edge, the order of access is controlled by a cost criterion, i.e., cheapest paths are considered first. To meet the stepwise refinement requirement (see Section 3.1.2) this graph search is done for all possible paths between the two nodes that involve an orientation relation. The spatial knowledge fragments provided by each path are used to construct the activated long-term memory representation (see Section 3.4.1.4).

Now, how are the costs of a path assessed? It has been assumed that there is one hierarchical structure given by the type of spatial knowledge and another one by the degree of resolution. Both factors are assessed by a numerical value and added up to obtain a cost assessment of a spatial knowledge fragment at a specific resolution. For reasoning about orientation relations, the cost assessment is done as follows:

- It has been motivated above that topological relations are considered most basic, whereas shape information are most elaborate. As we are interested in an orientation relation between two entities, however, it has to be further differentiated between the class of topological relations. Topological relations that spatially subordinate one entity to another one (i.e., the *inside*, *contains*, and *equal* relationships) are considered most essential for inferring orientation information, as they allow for reasoning in superordinate spatial structures. The remaining topological relations on the other hand (i.e., disjoint, tangent, and overlaps) only provide useful information after an orientation relation has been considered. Therefore, four classes of relations are identified, that are assigned cost indices of 1 (for the essential topological relations) through 4 (for shape information).
- The cost index corresponding to the degree of resolution is computed by the resolution index annotated to the relation in the hierarchical long-term memory representation divided by the number of possible classes of resolutions (i.e., 2 for topological and shape relations, and 4 for orientation relations). This division is done to standardize the index with respect to the possible stages of resolution. So the resolution index for topological relations and shape information is .5 for the low resolution and 1 for highly resolved topological relations. The possible values for orientation relations are .25, .5, .75, and 1.

Table 3.3 shows the cost assessment for all types of spatial relations at all stages of resolutions. The values are obtained by adding the cost index for the class of spatial relation and the resolution cost index.

Table 3.3. Cost assessment for all types of spatial relations and all resolutions

resolution	topological relations *in-at-border*, *in*, *contains-at-border*, *contains*, *equal*	all orientation relations	topological relations *disjoint*, *tangent*, *overlaps*	all shape representations
1	1.5	2.25	3.5	4.5
2	2	2.5	4	5
3	-	2.75	-	-
4	-	3	-	-

Now the dynamic behavior of the access process can be illustrated using the 'Reno – San Diego' example. Figure 3.7 shows the activated long-term memory representation with costs annotated to the paths. Observe that the costs of shape representations are assigned to the nodes, as they represent properties of single entities. The cheapest path starting from 'Reno' and 'San Diego' that meets in an orientation relation is:

in (Reno, Nevada) – E (Nevada, California) – in (San Diego, California)

with total costs of 5.5. The spatial knowledge fragments contained in this path are returned by the access process and passed to the construction process to build up the activated long-term memory representation.

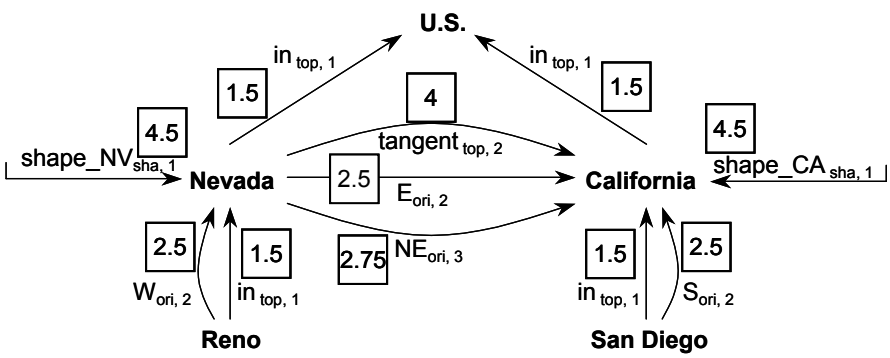

Fig. 3.7. Hierarchical long-term memory representation with costs annotated

In subsequent access steps, knowledge fragments refining this path are accessed and passed one by one to the construction process. The refining spatial knowledge fragments are ordered by their costs values. The order in which they are returned in the example is shown in Fig. 3.8.

NE (Nevada, California)
W (Reno, Nevada); S (San Diego, California)
tangent (Nevada, California)
shape_NV (Nevada); shape_CA (California)

Fig. 3.8. The order in which refining spatial knowledge fragments are returned in the example (see text). The vertical ordering of the spatial knowledge fragments indicates the order in which they are returned. Horizontal arrangement indicates that the order among them is not determined by the algorithm

3.4.1.4 The Activated Long-Term Memory Representation

Before I can show how the construction process makes use of the spatial knowledge fragments provided by the access process, the structure of the activated long-term memory representation has to be defined.

Definition: The *activated long-term memory representation* is defined like the hierarchical long-term memory representation (cf. Section 3.4.1.2) with the following restriction: for each tuple of geographic entities the directed graph does not contain spatial relations of the same type at different levels of granularity.

The activated long-term memory representation is built up using the spatial knowledge fragments sequentially provided by the access process, which explores the hierarchical long-term memory representation. The hierarchical long-term memory representation, however, may contain knowledge fragments involving spatial entities that are of the same type, but provide spatial information at different levels of granularity. An example for this case taken from the structure depicted in Fig. 3.6 is given by the fragments 'E (Nevada, California)' and 'NE (Nevada, California)'. Since the activated long-term memory representation contains the fragments that are relevant for constructing the visual mental image in working memory, only the most refined fragment of the same type is kept. Thus, during the construction of the activated long-term memory representation the former, less detailed fragment 'E (Nevada, California)' will be replaced by the more specific fragment 'NE (Nevada, California)' when it is provided by the access process (see next section).

Since the activated long-term memory representation is constructed from spatial knowledge fragments provided by the access process, it follows that the activated long-term memory representation contains only fragments that belong to a relevant path between the nodes that are to be spatially related to each other.

3.4.1.5 The Construction Process

The construction process for the activated long-term memory representation follows in a straightforward manner from the requirements of this representation structure and the form in which spatial knowledge fragments are provided by the access process. The purpose of the construction process is to build up a directed graph from spatial knowledge fragments extracted from hierarchical long-term memory, while checking for granularity conflicts between fragments of the same type involving the same set of geographic entities.

I will illustrate the construction using the spatial knowledge fragments taken from the example (see Fig. 3.8). The first three spatial knowledge fragments that are provided by the access process are 'in (Reno, Nevada)', 'in (San Diego, California)', and 'E (Nevada, California)'. A depiction of the resulting activated long-term memory is shown in Fig. 3.9a.

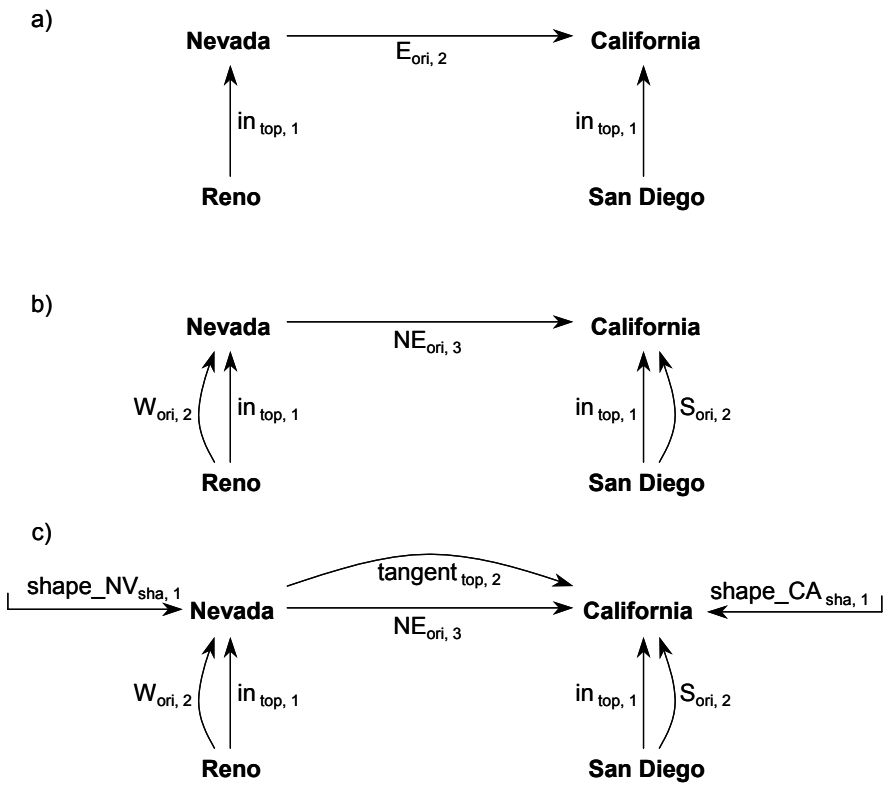

Fig. 3.9. Depictions of some stages of the activated long-term memory representation according to the spatial knowledge fragments returned by the access process (cf. Fig. 3.8)

The next spatial knowledge fragment delivered by the access process, i.e., 'NE (Nevada, California)' (cf. Fig. 3.8), causes the construction process to replace the formerly represented orientation relation between 'Nevada' and 'California' ('E (Nevada, California)') by the new piece of information. The orientation relations 'W (Reno, Nevada)' and 'S (San Diego, California)' that are provided next are simply added to the activated long-term memory representation, because there is yet no orientation relation between these pairs of entities. The activated long-term memory representation resulting from the last construction steps is depicted in Fig. 3.9 b.

Figure 3.9c shows the resulting activated long-term memory representation after all six spatial knowledge fragments listed in Fig. 3.8 have been successively integrated by the construction process.

The resulting representation of the long-term memory activation subsystem is the activated long-term memory representation, which at any stage of its construction may be used by the visual mental image construction subsystem. The visual mental image construction is described in the following subsection.

3.4.2 Visual Mental Image Construction

The visual mental image construction subsystem comprises the *conversion* process, which uses the *activated long-term memory representation* to convert it into the *enriched representation*; the enriched representation in turn is used by the *visualization* process to evoke the visual mental image in the *visual buffer*. The purpose of the *maintenance* loop is to periodically update the image in the visual buffer. Within the maintenance loop, further pieces of knowledge are integrated in the mental image as soon as they are available in the activated long-term memory.

3.4.2.1 The Enriched Representation

The enriched representation forms the basis for constructing the mental image in the visual buffer. Every piece of geographic information in the activated long-term memory that is not specific enough for immediate visualization is complemented by further details and subsequently represented in the enriched representation.

Definition: The *enriched representation* is defined like the activated long-term memory representation with the following two modifications: (1) shape information is assigned to every geographic entity that is required to be visualized as spatially extended; and (2) for every pair of geographic entities for which a spatial knowledge fragment is represented, both a topological and an orientation relation is specified.

Every geographic entity contained in the enriched representation is either defined as punctual or extended. For extended objects, standard shapes are assigned when no specific shape information is represented in the enriched representation.

The spatial relations represented in the enriched representation are specified up to an extent that allows for a straightforward parameterization that is done by the visualization process. This means that every represented spatial relation between

two geographic entities is determined with respect to its orientation and with respect to its topological relationship.

3.4.2.2 The Conversion Process

The purpose of the conversion process is to transform the activated long-term memory representation into the enriched representation. The conversion process is related to the *categorical to coordinate* conversion subsystem in Kosslyn's (1994a) mental imagery model (cf. Section 2.3.3.2). However, the conversion process does not assign specific values to the representations of the entities and the relations that hold between them. Rather, it qualitatively specifies the relations up to a degree which allows for an easy transformation into the visual buffer by the subsequent visualization process.

When looking at the activated long-term memory representation in the above example (cf. Fig. 3.9), it becomes evident that not every entity contained in the representation is specific regarding the ontological type it belongs to. Obviously, the nodes 'Nevada' and 'California' in Fig. 3.9.a must be extended objects (i.e., regions) to fit the 'in (Reno, Nevada)' and 'in (San Diego, California)' relationships (cf. Section 2.4.1). However, this property is not explicitly represented in the exemplary activated long-term memory representation, although it is essential for visualizing this spatial state of affairs.

In a first step, every entity is assigned an ontological type (*ontological typing*). As far as possible, entities are assumed to be points. Although, all entities can be conceived as extended, for reasons of simplicity, entities shall only be represented as extended when this is necessary. For example, the ontological type of the nodes representing the cities Reno and San Diego may be either punctual or extended. As it is simpler to visualize a point entity (compared to an extended region) these entities consequently are assumed to be points. The nodes representing Nevada and California, on the other hand, need to be represented as extended entities.

Now having decided on the ontological types of the objects, their properties and the spatial relations that hold between them can be further specified (*relational completion*). First of all, a standard shape will be assigned to those extended entities whose shape is not specified in the activated long-term memory representation. The question what kind of standard shape this may be (i.e., round, rectangular, hexagonal, or other shape) is an empirical question which is beyond the scope pursued here. In MIRAGE, a square shape is assigned as standard form for extended entities.

After having decided on the ontological type of the entities, the spatial relations between the represented entities can be further specified. For every pair of entities for which a topological relation is already represented in the activated long-term memory representation, an orientation relation is assigned (if not yet represented), and vice versa. So for every pair of entities for which a spatial relation has been represented in the activated long-term memory representation before, both a topological and an orientation relation is now represented. Note that the spatial relationships represented in the enriched representation are still qualitative (i.e., no

explicit metric values are assigned). So the actual task of the visualization is left to the interpretation of the enriched representation by the visualization process.

Table 3.4 shows which orientation relations can be assigned to hold between two entities when a topological relation is given and vice versa. The table is arranged according to the possible combinations of ontological types of two entities. The relations that are used as default in MIRAGE are underlined.

Table 3.4. Orientation relations to be added for two entities when a topological relations is given and vice versa. The table is arranged according to the possible combinations of ontological types of two entities. MIRAGE's default relations are underlined

relation given	**point – point**	**point – extended**	**extended – point**	**extended – extended**
topological relations:				
disjoint	all orientation relations except neutral; N	all orientation relations; N	all orientation relations; N	all orientation relations; N
tangent, overlaps				
in-at-border		all orientation relations; N		
in		all orientation relations; neutral		all orientation relations; neutral
contains-at-border			all orientation relations; N	all orientation relations; N
contains			all orientation relations; neutral	all orientation relations; neutral
equal	neutral			neutral
orientation relations:				
all orientation relations except neutral	disjoint	disjoint, in-at-border, in	disjoint, contains-at-border, contains	all topological relations except equal; disjoint
neutral	equal	disjoint, in-at-border, in	disjoint, contains-at-border, contains	all topological relations; equal

In the example (cf. Fig. 3.9), the 'in' relationships are complemented by a neutral orientation relation, and the specific orientation relations (e.g. 'E', 'NE') are complemented by the topological relation 'disjoint'. The resulting enriched representations that correspond to the activated long-term memory representations shown in Fig. 3.9a and c are depicted in Fig. 3.10a and b, respectively.

The enriched representation is now used by the visualization process to produce a mental image of the represented states of affairs in the visual buffer. So the visual buffer will be defined next.

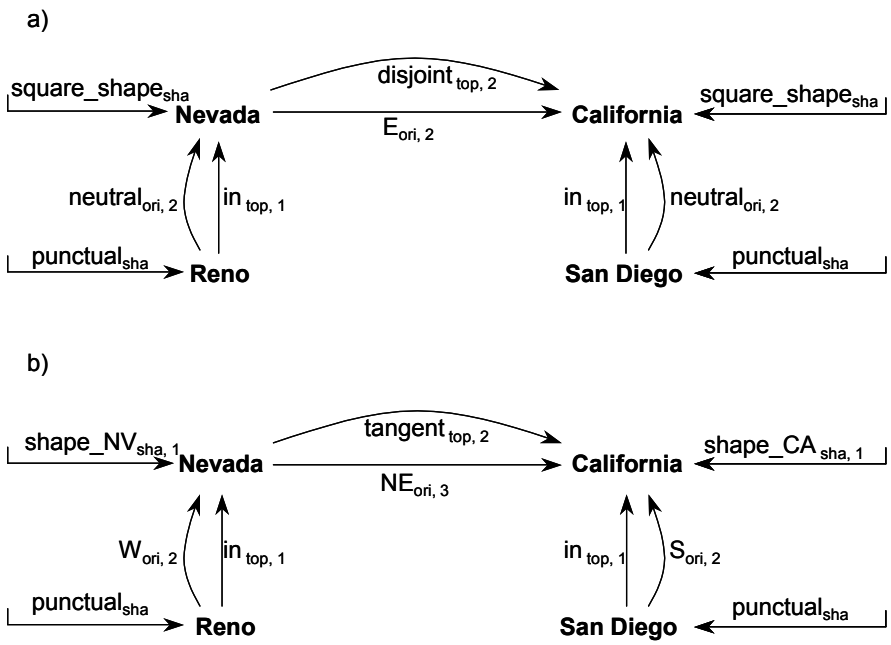

Fig. 3.10. The resulting enriched representation from the activated long-term memory representation depicted in Fig. 3.9a and c

3.4.2.3 The Visual Buffer

The visual buffer is the quasi-pictorial medium that is used for pictorially representing spatial information contained in the enriched representation. Its purpose is to exhibit geographic knowledge that is not explicitly represented in long-term memory. According to the ideas of diagrammatic reasoning (cf. Section 2.5) it is used to make implicit spatial knowledge explicit. For this purpose, qualitative spatial relations are depicted in the spatio-analogical representation medium.

However, essential for this task is not the representation structure of the visual buffer per se, but rather the functional characteristics provided by the processes that operate on this representation structure. Therefore, from a representation-theoretic point of view it is less important how the pictorial representation structure is physically realized (for instance, whether the image is represented in a positional or in a relational diagrammatic representation structure), but rather how the pictorial information is explored by the inspection process that utilizes the visual buffer.

In MIRAGE, the content of the visual buffer is described as an image representation in terms of elementary pictorial components like points, line segments, and polygons. The representation structure used for modeling the visual buffer

may be regarded as an intermediate representation format that can be used to evoke a positionally represented image, for example in a raster matrix.[47]

3.4.2.4 The Visualization Process

The task of the visualization process is twofold: first, the entities and their relations represented in the enriched representation are assigned specific (metric) values (*image specification*). This first step is needed for subsequently evoking specific images in the visual buffer. Second, the specified geographic entities are mapped into the visual buffer to evoke the image proper (*image mapping*).

The image specification step is a necessary precondition to performing the visualization proper: entities that will be used to produce a specific image need to be defined in terms of geometric descriptions. The second step (image mapping) restricts the specification generated by the image specification step by determining the part of the image to be visualized in the visual buffer. For this purpose, an appropriate clipping and scaling is determined. The restrictions imposed by image mapping are necessary as geographic configurations represented in the activated long-term memory representation may be too extensive to sensibly fit into the visual buffer. The area in the visual buffer that can be inspected in a detailed way (i.e., the highly resolved center area or the attention window, cf. Kosslyn, 1980; 1994a) is restricted both in size and resolution. Thus, an image resulting from visualizing everything contained in the activated long-term memory representation may have too low a resolution to be able to inspect the information the image originally has been generated for. The activated memory representation may comprise more information than can sensibly be mapped into the visual buffer:

> "Given the severe capacity limits of short-term memory ..., more information often may be activated in long-term memory than can be represented in short-term memory. Thus, there often will be a complex 'swapping' process between the two types of memory, which shuffles information in and out of short-term memory" (Kosslyn, 1994a: 324).

Under these conditions it may be the case that the visualization (more specifically, the image mapping step) has to be done in an iterated manner: the result of the inspection of a first image mapping may require another mapping with clipping and scaling changed. This iterated image mapping is based on the same information represented in the activated long-term memory representation. It may become effective in the maintenance loop necessary to keep the mental image vivid in the visual buffer (see next section).

Therefore, the spatial knowledge represented in the enriched working memory representation potentially may be more complex than what (due to the scaling and clipping performed in the image mapping part of the visualization process) actually can be mapped into the visual buffer.

---

[47] For observing the operation of the implemented model the representation in the visual buffer is (positionally) visualized on the computer screen (see Section 5.3).

When considering the 'Reno – San Diego' example again (cf. Fig. 3.10), visualization may proceed as follows. In Fig. 3.10a, the shapes of the entities 'Nevada' and 'California' need to be metrically specified before the positions of all entities can be determined. The square shapes of extended entities are given by the point coordinates of two diagonal corners (e.g., [0, 10; 10, 0]). This shape is assigned to both of the extended entities. After that, the position of the entity 'Nevada' can be modified (by horizontally shifting) to fit the relations 'E (Nevada, California)' and 'disjoint (Nevada, California)'. So the position of Nevada may be given by [12, 10; 22; 0]. Finally, the positions of the entities 'Reno' and 'San Diego' can be determined to fit the 'neutral' and 'in' relationships with respect to their reference objects 'Nevada' and 'California'. Their positions can be determined by the point positions [17, 5] and [5, 5], respectively. The pictorial representation resulting from these specifications is depicted in Fig. 3.11.

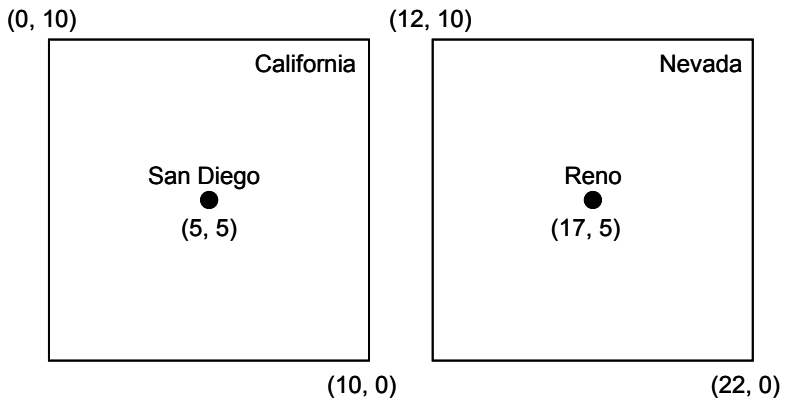

Fig. 3.11. Depiction of the resulting representation of the image specification step of the visualization process. Specific sizes and locations have been assigned to the square standard shapes of the extended entities (cf. Fig. 3.10a)

As explained above, in most cases it will not be the case that the whole image is mapped into the visual buffer. Rather, a specific part is selected in the image mapping step. According to the part that is focused on, the image is scaled appropriately and properly located in the visual buffer. In the example, the relative orientation between the two entities 'Reno' and 'San Diego' is required. Thus, when mapping the image into the visual buffer, it can be focused on these entities. A resulting image representation in the visual buffer that is usable for inspecting the required relation is depicted in Fig. 3.12.

In the example in Fig. 3.10b the entities 'California' and 'Nevada' are already represented together with their specific shapes; thus, in this case no standard shape needs to be assigned by the image specification process. Again, the positions of the regions have to be determined such that they fit the 'NE (Nevada, California)'

and 'tangent (Nevada, California)' relationships represented in the enriched representation. Finally, the positions of the cities can be defined within the represented regions. The representation resulting from image specification is depicted in Fig. 3.13a. Figure 3.13b shows the resulting visual buffer representation after focusing on the relevant entities.

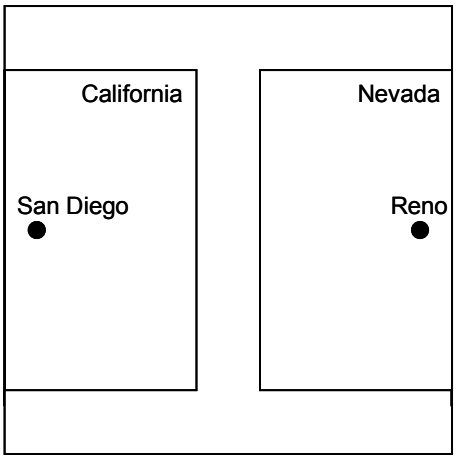

Fig. 3.12. Depiction of a possible result of the image mapping step performed on the representation shown in Fig. 3.11

3.4.3 Image Inspection

The image inspection subsystem utilizes the image constructed in the visual buffer to evaluate the required spatial relation.

3.4.3.1 The Inspection Result

The result of the image inspection is a qualitative spatial relation. Thus, it is represented as a spatial knowledge fragment in MIRAGE (cf. Section 3.4.1.1). More specifically, the result of an image inspection is a topological or an orientation relation between a pair of geographic entities represented in the system. Like in any spatial knowledge fragment, the type and the resolution of the spatial relation is annotated in the inspection result. So, for example an orientation relation can be inspected at 4 different resolution levels.

As being a spatial knowledge fragment, the structure of the inspection result is compatible with the other subsystems of MIRAGE. Therefore, a spatial relation inspected in the visual buffer can be used for subsequent spatial reasoning processes by storing it in the hierarchical long-term memory representation or by integrating it directly in the activated long-term memory representation for further use. However, these options are beyond the scope of this work.

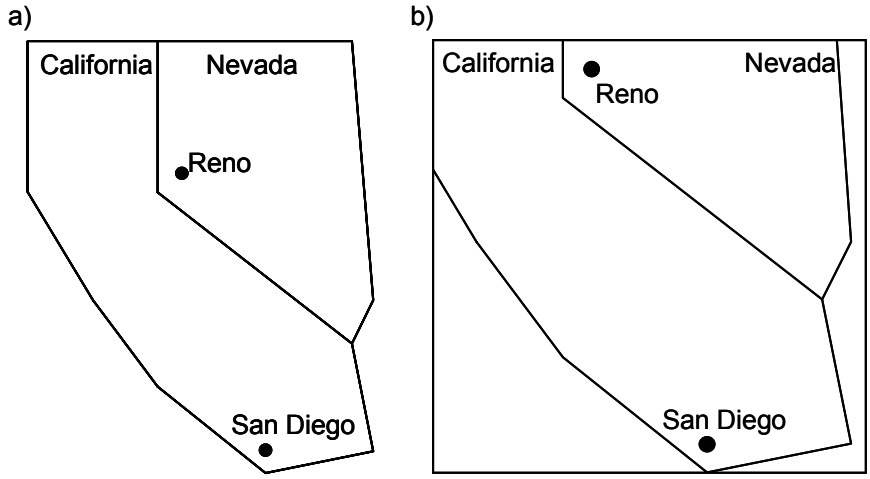

Fig. 3.13. a) Depiction of the resulting representation of the image specification step of the visualization process. Shapes explicitly represented in the enriched representation have been used (cf. Fig. 3.10b). b) Depiction of the result of the image mapping step performed on the representation shown in a)

3.4.3.2 The Inspection Process

The image content in the visual buffer is represented in MIRAGE in the form of a description of its image components (cf. Section 3.4.2.3). These components have to be evaluated with respect to the question to be elaborated. For this purpose, the picture elements contained in the image representation are interpreted in terms of qualitative spatial relations, i.e., their relative positions are translated into a spatial knowledge fragment.

The task of the image inspection process is specified by the entities under consideration, by the type of relation required (i.e., topology or orientation), and by the intended resolution of the inspection result.

To illustrate the operation of the inspection process in MIRAGE let us consider the 'Reno – San Diego' example again (Fig. 3.12 and Fig. 3.13b). We are interested in the cardinal direction of San Diego with respect to Reno. As the positions of the entities 'Reno' and 'San Diego' are determined in the visual buffer representation, the orientation of the two point entities can be evaluated with respect to each other. The resulting spatial knowledge fragment is of the form

$$\langle\text{orientation\_relation}\rangle_{\text{ori, }\langle\text{resolution}\rangle}(\text{San Diego, Reno}),$$

where <resolution> specifies the level of granularity at which the orientation of San Diego with respect to Reno is evaluated, and <orientation_relation> is the cardinal direction that is read off the image. The cardinal direction is determined on the basis of the angular orientation between the two entities in the visual buffer. The angle of San Diego with respect to Reno is measured counterclockwise,

where 0° is horizontally right of Reno. The resulting angles are 180° in the image representation depicted in Fig. 3.12 (San Diego is straight left of Reno) and 287° in the image according to Fig. 3.13b. The possible image inspection results for both images at every resolution possible are listed in Table 3.5.

Table 3.5. Inspection results based on the mental image representations depicted in Fig. 3.12 and 3.13b for the orientation relation of San Diego with respect to Reno at all possible resolutions

resolu-tion	**cardinal direction**	**angle interval**	image according to Fig. 3.12 (angle San Diego – Reno: 180°) **inspection result**	image according to Fig. 3.13b (angle San Diego – Reno: 287°) **inspection result**
1	S	[0; 180)		$S_{ori,\ 1}$ (San Diego, Reno)
	N	[180; 0)	$N_{ori,\ 1}$ (San Diego, Reno)	
	W	[90; 270)	$W_{ori,\ 1}$ (San Diego, Reno)	
	E	[270; 90)		$E_{ori,\ 1}$ (San Diego, Reno)
2	E	[315; 45)		
	N	[45; 135)	$W_{ori,\ 2}$ (San Diego, Reno)	$S_{ori,\ 2}$ (San Diego, Reno)
	W	[135; 225)		
	S	[225; 315)		
3	…	[112.5; 157.5)		
	W	[157.5; 202.5)	$W_{ori,\ 3}$ (San Diego, Reno)	
	SW	[202.5; 247.5)		
	S	[247.5; 292.5)		$S_{ori,\ 3}$ (San Diego, Reno)
	SE	[292.5; …		
4	…	[146.25; 168.75)		
	W	[168.75; 191.25)	$W_{ori,\ 4}$ (San Diego, Reno)	
	WSW	[191.25; ...		
	...	[258.75; 281.25)		
	SSE	[281.25; 303.75)		$SSE_{ori,\ 4}$ (San Diego, Reno)
	SE	[303.75; …		

For every resolution applicable to orientation relations the table (in parts) lists the angle intervals that correspond to the respective cardinal directions. For instance, at resolution 2 the direction 'north' corresponds to an angle between 45° (inclusive) and 135° (exclusive). The resulting interpretations for both images are shown in the last two rows. As in the image representation shown in Fig. 3.12 San Diego is straight left of Reno (180°) the inspection at all resolutions will be that San Diego is west of Reno (at resolution 1, also the interpretation that San Diego is north of Reno is possible, depending on which dichotomic distinction is used). In the more elaborate version of the image representation (Fig. 3.13b), the result of the inspection varies with the resolution chosen. At resolutions 1 through 3, San Diego is inspected to be south of Reno (at resolution 1, also the interpretation that San Diego is east of Reno is possible). At the resolution 4, however, the interpretation yields the more specific spatial relation that San Diego is south-southwest of Reno.

4 Visual Mental Image Construction in Detail

In Section 3.4.2 I showed how mental images can be generated in the visual buffer based on an activated long-term memory representation. In the presented 'Reno – San Diego' scenario the visualization of the working memory content could be performed in a straightforward way. However, the generation of images from activated long-term memory representations is not always that easy.

In this chapter I will investigate the image construction process in more detail (cf. Fig. 4.1). I will discuss a more demanding scenario that points to problems related to the construction of image-like spatial knowledge representations. I will discuss strategies to cope with those problems both from an analytical point of view and under the computational modeling perspective. MIRAGE will be extended by image construction strategies that allow for managing the image generation process with reasonable cognitive effort.

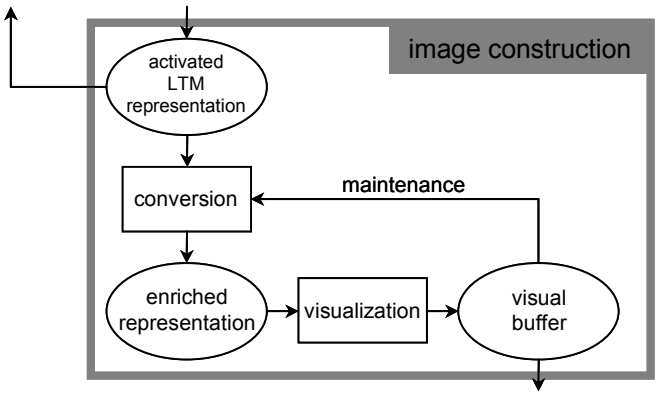

Fig. 4.1. The image construction subsystem of the MIRAGE architecture (cf. Fig. 3.4)

4.1 A More Demanding Scenario

As in the 'Reno – San Diego' scenario (Section 3.4), the scenario used here deals with deciding about the spatial orientation between two geographic entities. Suppose a person wants to decide about the relative orientation of the two cities of Nice (France) and Geneva (Switzerland). Like in the 'Reno – San Diego' example above, available pieces of information may be the two states involved and the relative orientation between those states. Let us assume that knowledge is available about the relative orientation between Lake Geneva (Switzerland) and the two cities of Geneva and Nice: Lake Geneva is known to be roughly east of Geneva and roughly north of Nice. Figure 4.2a shows these pieces of knowledge on a map, and Fig. 4.2b shows them in form of the tentatively corresponding activated long-term memory representation (cf. Section 3.4.1.4). This activated long-term memory representation will now serve as a basis for the construction of the visual mental image to exhibit the relative orientation between Nice (France) and Geneva (Switzerland).

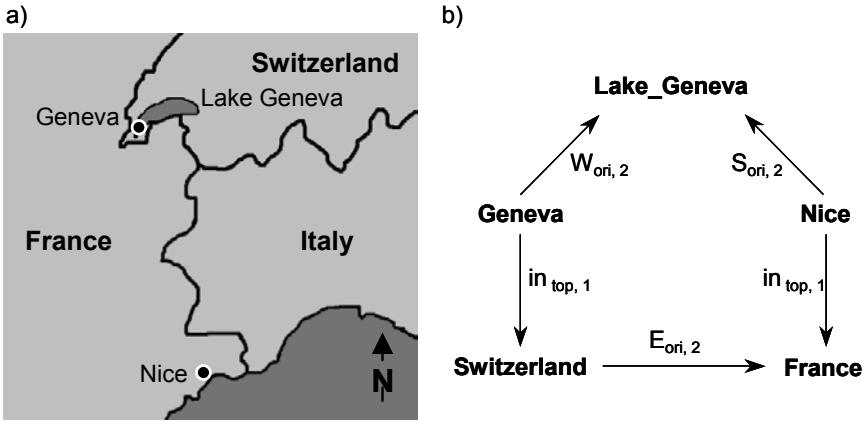

Fig. 4.2. Spatial relationships between Geneva, Lake Geneva, Nice, Switzerland, and France (a) on a map and (b) as a tentative activated long-term memory representation

When comparing the activated long-term memory representation depicted in Fig. 4.2b with the 'Reno – San Diego' example in the previous chapter (cf. Fig. 3.9, Section 3.4.1), it shows that this representation is a cyclic graph: each geographic entity represented in this graph is constrained by spatial relationships with respect to two other entities. Although each of the five relations independently appears reasonable and correct, when they are jointly used in a mental image construction in MIRAGE, they generate a conflict. Remember the default assumptions about shapes and relations introduced in Chapter 3: square shapes are employed as standard forms for the extended objects, point objects are located at their centers, and entities are oriented in their prototypical directions. If one tries to visualize

this representation in the same way as proposed in the previous chapter, it shows that no visualization can be constructed that is consistent with all five spatial constraints represented.

However, when one of the five represented relationships is ignored, a mental image construction is possible. Depending on which of the relations is ignored, five different partial solutions can be constructed. Figure 4.3 depicts all mental images that can be obtained when one spatial relationship is omitted from the activated long-term memory representation. Observe that the respective omitted spatial relation is in conflict with the image generated by the other constraints. Although the activated long-term memory representation shown in Fig. 4.2b corresponds to the map representation in Fig. 4.2a, each of the visualized variants depicted in Fig. 4.3 is in conflict with one spatial relation represented in the underlying long-term memory representation.

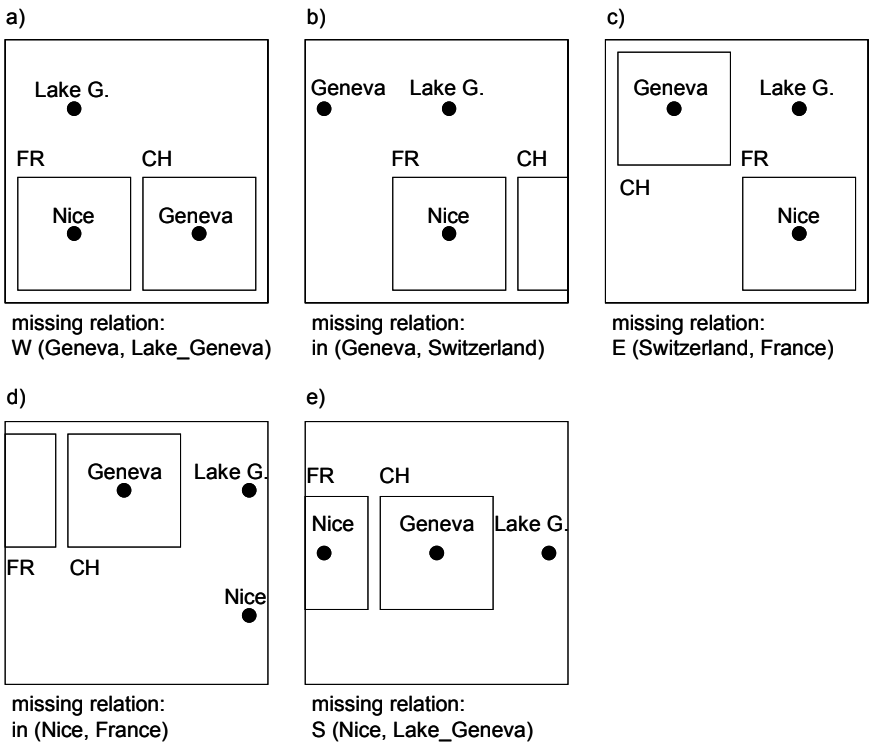

Fig. 4.3. Five possible image constructions resulting from the activated long-term memory representation shown in Fig. 4.2b. Each visualization violates one of the represented spatial relationships

So there is no easily achievable visual mental image construction for this activated long-term memory representation. How can such problems be tackled in

the mind such that a mental representation is obtained that can be used to decide upon a given spatial question?

4.2 Diagrammatic Representations of Lean Knowledge

When attempting to construct a visuo-spatial representation (like a mental image) from a set of underdetermined spatial representations (like the pieces of knowledge from long-term memory), several degrees of freedom have to be dealt with. In MIRAGE, coarse orientation information, topological relationships, or both are provided from long-term memory; also, rough shape information about extended spatial entities may be available. Missing orientation, topology, or shape information is added in the conversion process, whereas the actual specification of positions, sizes, and distances is left to the visualization process (more specifically, the image specification process).

From a computer science point of view, the problem of determining a spatio-analogical representation of a given set of spatial information can be described in the form of a constraint satisfaction task. The set of the given spatial binary relationships is expressed as a *binary constraint network* (BCN; cf. Ligozat, 1998). A BCN is a directed graph; its nodes are interpreted as point locations, and its edges denote the given spatial relations. A core problem of spatial reasoning in artificial intelligence (AI) is the question whether a spatial realization of the states of affairs represented in a given BCN exists, i.e., whether the BCN describes a possible spatial situation or not. This *consistency problem* is a precondition for constructing a spatio-analogical representation, e.g. in a cell matrix or in a symbolic array (cf. Section 2.5.3). However, even for subsets of spatial relationships the consistency problem is computationally intractable. For instance, the consistency problem of a BCN that represents orientation relations between pairs of point entities is known to be NP-complete.[48] As a consequence, the type of spatial reasoning task considered here cannot be solved on the basis of directly computing a spatial solution from the given constraints.

From a psychological point of view, restrictions in mental resources are the limiting factor of possible computational processes in the mind. Only a few items can be operated on at the same time, and solutions are often needed fast. The strategy in mental spatial knowledge processing is to eliminate degrees of freedom at early processing stages by committing to specific interpretations instead of general constellations with a variety of options. The problem space (i.e., the universe of possible constellations that potentially contains solution(s) to a given problem) is rapidly reduced in an attempt to find a solution in a straightforward way. If the

[48] The class of NP-complete problems refers to computations whose complexity grows exponentially with respect to the number of items involved. As a result, only very small instances of NP-complete problems can be computed on digital computers, whereas in the general case no solution can be obtained in reasonable time.

chosen strategy is successful, i.e. if a visuo-spatial representation can be successfully constructed, the problem solving process is straightforward and fast.

The cognitive characteristic of restricting the problem space and committing to specific cases is related to the theory of *preferred mental models* (Knauff et al., 1995; Rauh et al., 1997; Schlieder, 1999). When people construct a mental model to perform a reasoning task and several alternative models are valid, the selection of the candidate model is not random. Rather, there are cognitive preferences that lead people to construct certain models first, whereas others only are constructed later, when the question to be solved is reflected more thoroughly.

However, the preferred mental model strategy entails a number of drawbacks. First of all, it is not clear beforehand whether it will be possible to find a spatio-analogical representation at all. This is due to the fact that consistency of the underlying representation cannot be considered beforehand. So if one tries to find a solution and fails, this may be due to the wrong image construction strategy or due to inconsistencies in the underlying facts.

Moreover, the early restriction of the problem space may cause conflicts that result in an unsuccessful attempt to construct a spatio-analogical representation. The problem presented in the previous subsection occurred because default assumptions used in the construction of the visual mental image have been too restrictive. So compared to the AI approach to compute a solution while considering all possible options, the psychologically motivated modeling approach pursued her operates conversely:

1. The problem space is radically restricted towards finding a solution in a straightforward way.
2. Only when a solution is not found, the restrictions are relaxed to allow for more potential solutions.

In the following section I will discuss alternative options for the image construction in MIRAGE.

4.3 Consequences for Image Construction

In this section I will consider three options for constructing visual mental images that can be applied when the straightforward image generation strategy fails. These options are:

1. the reduction of spatial constraints,
2. variations in the completion of qualitative spatial relations, and
3. modifications in the interpretation of qualitative spatial relations during the image construction.

4.3.1 Relaxation of Spatial Constraints

When conflicting facts are represented in the given activated long-term memory representation, the simplest way to still construct a mental image is to omit one or more facts that conflict with other facts during the image generation process. This

strategy presupposes that no visualization of all facts can be found (with reasonable effort), and that a visualization that does not use all facts is sufficient to obtain a reasonable answer.

Two arguments favor the omission of facts. Either, one or more facts are assessed as being less reliable than others (in this case the image is generated by using only the most reliable facts); or it is assumed that employing a fast and straightforward visualization strategy that does not include all facts still yields a correct result. In the latter case, facts are excluded although they are assumed to be true, because they conflict with other facts in the chosen image construction strategy. Of course, whether or not the obtained image is correct immediately depends upon the decision which facts are used and which are excluded.

When considering again the 'Geneva – Nice' example depicted in Fig. 4.3 we observe that in two of the five images (i.e., in Fig. 4.3a and in Fig. 4.3e) Nice is found to be west of Geneva, which is not correct in comparison with the given map. In the other three cases (Figs. 4.3b-d) Nice is located south-east of Geneva, which corresponds to the relation found in the map.

4.3.2 Completion of Qualitative Spatial Relations

In the conversion process, topological relations are complemented by orientation relations and vice versa. Moreover, default shapes are assigned to extended entities. Table 3.4 (Section 3.4.2.2) showed all possible combinations of topological and orientation relationships between two entities together with the relations that are added as defaults. When the default relations do not allow for an image construction, however, other possible relations may be used.

All variations of the relations for the completion of qualitative spatial relations can be tested. In the 'Geneva – Nice' example (cf. Fig. 4.2) it shows that the topological "in" relationship can be complemented by any orientation relationship. The orientation relation between the two extended entities ('E (Switzerland, France)') can be complemented by all topological relations except the 'equal' relation. For the remaining two orientation relations 'W (Geneva, Lake_Geneva)' and 'S (Nice, Lake_Geneva)' only the 'disjoint' relationship may be added. For every binary relation in the example, Table 4.1 shows the default relations to be added, the other possible relations, and the number of possible relations. It shows that a total of 100 combinations (1*1*5*5*4) is possible in this example and can be used to check for suitable image constructions.

Regarding the assignment of default shapes it has been assumed so far that square shapes are employed. However, when a regular default shape is to be assigned, any other forms may be suitable as well (e.g. circles, hexagons, etc.). Other forms may be useful when specific neighborhood structures between extended entities are required (for example, rectangular forms differ from hexagonal or octagonal shapes with respect to their neighborhood structures, which may be important for adequately depicting orientation relations).

Besides varying the completion of qualitative spatial relations, it can also be sensible to leave spatial relations unspecified. For example, when the 'in (Geneva,

Switzerland)' relation is represented in the activated long-term memory representation, it can be sensible to leave the orientation relation between Geneva and Switzerland unspecified. The interpretation of this relation in the visualization subsystem is that Geneva can be at any place within the extended entity of Switzerland. I will further explain this option in the next section.

Also, the shape of extended entities can be left unspecified in the relational completion. This option allows for generating arbitrary shapes in the visualization process. This option will be elaborated in Section 4.4.4.2.

Table 4.1. Qualitative spatial relations that can be used to complement the relations represented in the activated long-term memory representation according to Fig. 4.2b. The total number of possible combinations in the example is 1*1*5*5*4 = 100

given relation	**default relation to be added**	**possible relations**	**number of possible relations**
W (Geneva, Lake_Geneva)	disjoint	disjoint	1
S (Nice, Lake_Geneva)	disjoint	disjoint	1
in (Nice, France)	neutral	all orientation relations	5
in (Geneva, Switzerland)	neutral	all orientation relations	5
E (Switzerland, France)	disjoint	all topological relations except equal	4

4.3.3 Interpretation of Qualitative Spatial Relations

The image specification process uses the qualitative spatial relations provided by the enriched representation to determine the parameters of the entities to be mapped into the visual buffer. In the examples presented so far, it has been assumed that orientation relations determine point locations that lie on a straight line through the reference object. For instance, the orientation relation 'W(Geneva, Lake_Geneva)' has been interpreted as Geneva being located left of 'Lake_Geneva' and on the horizontal line through 'Lake_Geneva' (Fig. 4.4a). However, besides this prototypical interpretation that interprets an orientation relation as a point location, also areal interpretations are possible. In this case, a cardinal direction relation refers to a sector that denotes the potential positions of the respective entity (Fig. 4.4b).

Relaxing the interpretation of a given qualitative spatial orientation increases the options for finding a visualization of a given set of spatial constraints. By considering one or more location areas instead of point positions, entities can be located appropriately. I will give an example of this case in the next section. Clearly, the visualization task becomes more complicated since more and more complex geometrical operations are needed during the image construction.

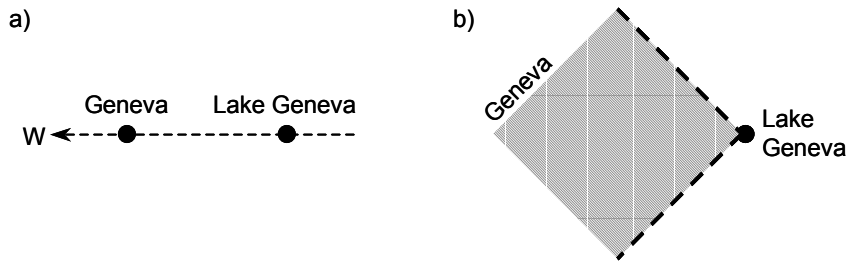

Fig. 4.4. a) Interpreting the orientation relation "$W_{ori,\,2}$ (Geneva, Lake_Geneva)" as a point location and b) as areal interpretation

Regarding the interpretation of shapes of extended entities, it has been assumed that in the relational completion process a regular standard square shape is assigned to extended entities whose shapes are not determined in the activated long-term memory representation. However, when standard shapes are too restrictive to cover a specific situation, it may be helpful to generate an arbitrary shape that allows for constructing an appropriate image. This generation of complex, irregular shapes corresponds to the *mental drawing* facility in human mental image generation: arbitrary shapes are generated in the visual buffer by continuously moving the attention window (Kosslyn, 1994a; cf. Section 2.3.3.2).

4.4 Image Revision Strategies in MIRAGE

In the following I will extend MIRAGE to allow for more flexible image construction strategies. First, I will demonstrate how inconsistencies in the long-term memory representation may result in *unstable images*. Second I will integrate

1. the option of omitting facts from activated long-term memory,
2. the revision of relational completion, and
3. the revision of image specification as image construction strategies.

I will give examples of how modifications in the relational completion and image specification processes can help to obtain a visualization when the default strategies fail due to conflicts during image specification. As will be seen, all alternative options for image construction are not independent of each other, and sensible strategies often require a combination of strategies.

Before going on I will refine the image construction subsystem. This refinement is needed since the partial processes participating in image construction will be extended and modified in the following.

As explained in the previous chapter, both the conversion and the visualization processes subdivide into two steps (cf. Section 3.4.2.2 and Section 3.4.2.4, respectively). The conversion process first determines the ontological types of the represented entities (*ontological typing*), and then it completes the representation with the missing shape properties and qualitative spatial relations (*relational comple-*

tion). The visualization process first determines the metric specification of the qualitative relations provided by the enriched representation (*image specification*), and then it maps the section of the image that contains the aspects of interest into the visual buffer (*image mapping*). Figure 4.5 shows MIRAGE's image construction subsystem with refined conversion and visualization processes. The refinement implies two new intermediate representations, the *typed representation* (provided by ontological typing) and the *specified image* (as the result of the image specification).

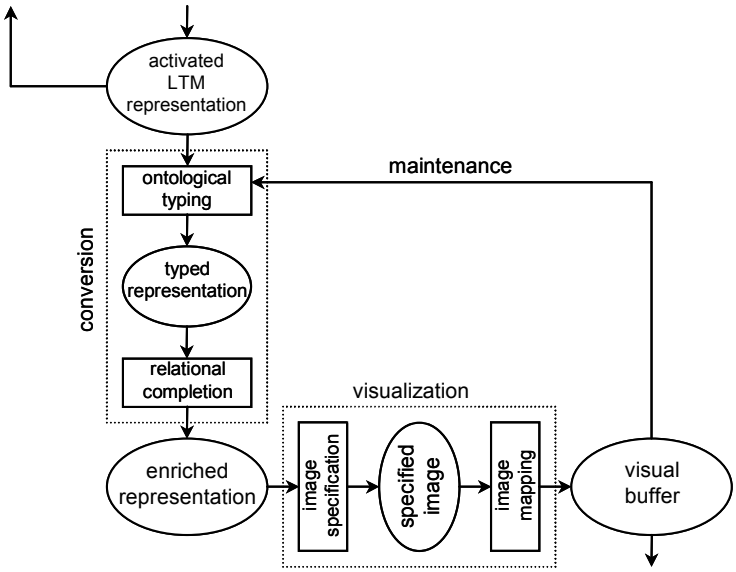

Fig. 4.5. The image construction subsystem with refined conversion and visualization subsystems

4.4.1 Unstable Images

Mental images in the visual buffer need to be regularly refreshed to prevent them from fading out (*image maintenance*; cf. Section 2.3.3). In the MIRAGE model, image maintenance is done by repeatedly performing the conversion and visualization processes. An unstable image may be induced by the image specification process. It results when entities cannot be uniquely localized due to contradictory spatial constraints. As a consequence, the system tries to reposition these entities again and again.

To illustrate this case let us consider again the activated long-term memory representation depicted in Fig. 4.2b. The conversion process can treat this representation in the usual manner as explained in Chapter 3. So the ontological typing process will model Geneva, Nice, and Lake Geneva as point entities, and

France and Switzerland as extended entities. In the relational completion process, the topological and orientation relations are completed as explained above. Figure 4.6 shows the resulting enriched representation.

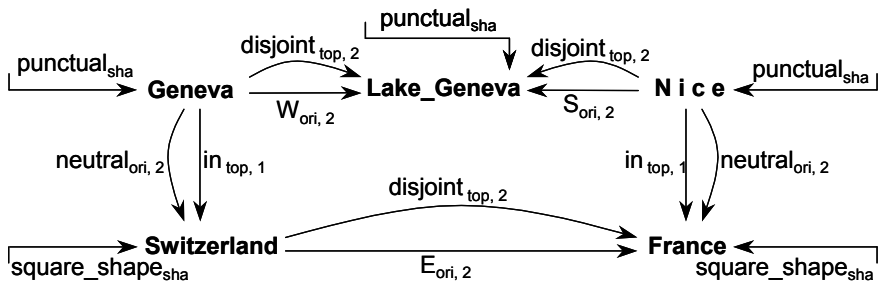

Fig. 4.6. The enriched representation that results from the conversion process performed on the activated long-term memory representation that is depicted in Fig. 4.2b

Now let us assume that the image specification process starts with 'Lake_Geneva' and processes in a clockwise manner (this case corresponds to the situation depicted in Fig. 4.3a). After defining the position of 'Lake_Geneva', 'Nice' is positioned straight below 'Lake_Geneva'. 'France' and 'Switzerland' are placed, and finally 'Geneva' is located centered in 'Switzerland'. The last spatial relation in the enriched representation, i.e. 'W (Geneva, Lake Geneva)', is in conflict with the image constructed so far.

If in this case the conflicting situation was not detected, the image specification process would not terminate. To fit the 'W (Geneva, Lake_Geneva)' relationship, 'Lake_Geneva' would be located once again straight to the right of 'Geneva'. Now the subsequent entities can be visualized again, too, with respect to the new location of 'Lake_Geneva'. The attempt to continuously visualize this representation in the visual buffer would result in an unstable, continuously moving image (in the example the image would move down and right with respect to the visual buffer), while older parts of the image fade out. This situation is illustrated in Fig. 4.7.

The image specification is repeated over and over again without reaching a stable representation for each image component. The described situation will not only occur when an image generation cannot be performed in an intended way. It also would occur when the representation in the activated long-term memory is contradictory and no visualization is possible at all. Remember that in the 'Geneva – Nice' example an image construction analogous to the map depicted in Fig. 4.2a would be possible. To decide upon the question whether an image construction is possible or not, further (logical) reasoning steps would be required.

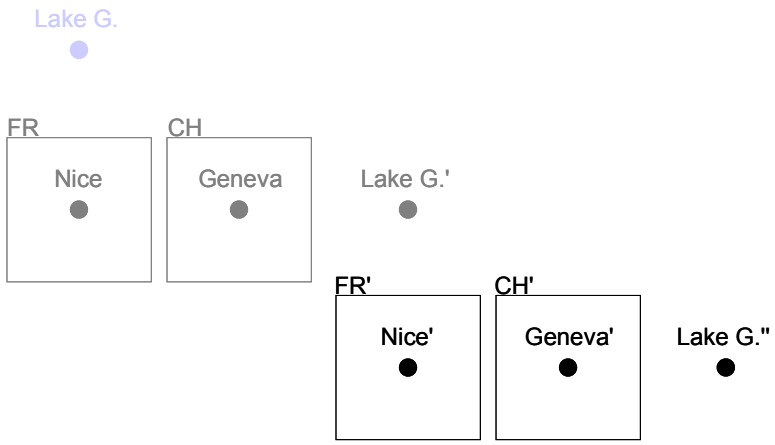

Fig. 4.7. Illustration of the attempt to generate an unstable image. While former instances of visualizations of entities are fading out, new instances are placed in new locations. The image moves down and to the right

Besides in moving images, unstable representations in the visualization process also can result in 'flickering' images. When two conflicting relations force a spatial entity to be positioned in two different locations, this entity may consistently 'jump' from one location to the other. This situation of an unstable image corresponds to the attempt to interpret 'impossible' figures and optical illusions in visual perception, as they are used in psychological investigations (e.g., Schacter et al., 1991).

In MIRAGE, the conflicting situation is detected during image specification as a precondition for a terminating image specification process, as well as for enabling alternative image generation strategies. During image specification it is checked whether the currently considered spatial entity is already specified, and if so, whether there is a conflicting situation.

The simplest conceivable solution to overcome an unstable image is to reject the image as a whole. The image specification breaks, and the image construction process stops. In this case no visualization of the corresponding activated long-term memory representation can be obtained, either because the representation in the activated long-term memory representation is over-constrained (in this case no complete image can be constructed), or because time constraints force to abandon the image construction attempt. This option may be sensible when an arbitrary timely decision is better than a delayed decision. However, when an image is needed because a proper solution to a given problem is required, other reasoning strategies have to be employed to deal with the conflicting situation.

4.4.2 Omission of Facts

Two reasons have been stated for reducing the number of facts in the activated long-term memory representation:

1. an omission of facts is considered sensible to enable a fast decision, or
2. one or more facts are considered less reliable than the others.

Corresponding to these two reasons, two strategies for reducing the number of facts are possible: either the omission is performed randomly, or specific pieces of knowledge are excluded from the given activated long-term memory representation. In the case of random omissions, the image is specified by the image specification process up to the maximum stage that still allows for determining positions and extensions of entities before a conflicting constraint forces to redefine an already specified entity. Visualization proceeds with the specified image developed so far and the specified image is mapped into the visual buffer for interpretation (see Fig. 4.8, case b).

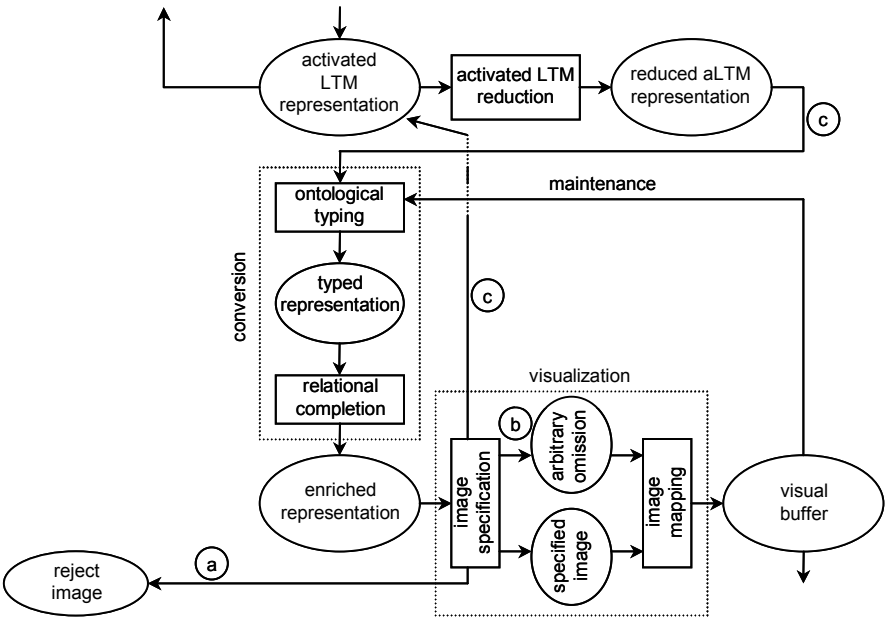

Fig. 4.8. Strategies for dealing with image construction conflicts: (a) the image is rejected and the visualization stops, (b) facts are omitted arbitrarily and visualization continues with image mapping, and (c) specific relations are eliminated from the activated long-term memory representation and the image construction continues with the conversion process

When specific pieces of knowledge are to be excluded from image generation due to conflicts the in image specification process, we have to go back to the activated long-term memory representation. A *reduction process* evaluates the

pieces of knowledge contained in the activated long-term memory representation and assesses the knowledge fragments with respect to their reliability. The least reliable fact is abandoned from the activated long-term memory representation, thus yielding a *reduced activated long-term memory representation*. With this reduced activated long-term memory representation the image construction can start all over again (i.e., beginning with the conversion process), attempting to produce a visualization in the usual way (see Figure 4.8, case c).

I assume that the activated long-term memory reduction process checks the spatial knowledge fragments contained in the activated long-term memory representation with respect to their reliability by referring to their long-term memory representation. Induced by the complications in the image construction the spatial knowledge fragments retrieved in the long-term memory activation subsystem are re-assessed to obtain a criterion for the exclusion of the least reliable spatial knowledge fragment.

Both variants may be sensible under certain circumstances. When there is a criterion for the reliability of every fact, those facts that are least reliable can be excluded from the visualization. The case of the arbitrary exclusion (which occurs when the visualization stops as a consequence of the chosen order of visualization) must be used when there is no a priori possibility to exclude a certain fact, or when time restrictions force to come out with a fast decision.

4.4.3 Revision of Relational Completion

When referring to the refined image construction subsystem as depicted in Fig. 4.5 once again two processes can be identified that can be modified to check for alternative visualization facilities: the relational completion process and the image specification process. The other two processes, i.e. ontological typing and image mapping do not provide sensible options: testing for alternative ontological types severely augments the complexity of the image construction task without providing advantages for alternative image constructions; and image mapping merely evokes the image proper in the visual buffer after the image has been specified by the image specification process.

The relational completion process and the image specification process, on the other hand, provide valuable options to check for alternative image constructions. In the relational completion process, alternative default knowledge components can be used. For example, alternative default shapes for extended entities can be employed, and other topological and/or orientation relations can be assigned for relations not determined by the long-term memory representation. In the image specification process the interpretation of the qualitative relations added in the relational completion process can be varied, i.e., the degrees of freedom with respect to how a qualitative spatial relation is depicted in terms of a diagrammatic representation can be utilized to generate alternative images.

In the next two sections I will treat the options outlined here in more detail, and I will extend MIRAGE's image construction subsystem accordingly.

4.4.3.1 Variation of Relational Completion

A revision in the relational completion subsystem can be performed in two different ways: on the one hand, the activated long-term memory representation can be complemented by other qualitative spatial relations than in the standard case (cf. Table 3.4); on the other hand, the relational completion can be partially reduced.

The relational completion process decides on default shapes for extended objects whose shape is not determined in the activated long-term memory representations. Moreover, orientation relations are added for pairs of entities that are topologically related to each other, and topological relations are added for pairs of entities whose orientation with respect to each other is given in the activated long-term memory representation.

The possible options for completing the qualitative spatial relations offers a number of variations in the relational completion process that can be used to check for an image that is consistent with all relations represented in the activated long-term memory representation. In the 'Geneva – Nice' example, such a completion can be found. Figure 4.9a depicts an enriched representation that allows for a straightforward image specification and visualization. The image constructed in the visual buffer is depicted in Fig. 4.9b. The effect that the two countries of France and Switzerland overlap and that the point entities belong to either of them, though not consistent with reality, is not a fault of this representation. The image has been constructed on the basis of the available spatial knowledge fragments; the realization chosen is not in conflict with any of the pieces of information used.

To enable the modifications in the relational completion process described in this section, MIRAGE's image construction subsystem needs to be extended (cf. Fig. 4.5). The inconsistency that prevents the image representation from being immediately visualized is detected by the image specification process. So the image specification process must interact with the relational completion process to trigger a varied completion by other qualitative relations possible. In MIRAGE, the relational completion process makes use of a set of parameters that determines the relations that are added depending on the relations given in the activated long-term memory representation. In Section 3.4.2.2, I presented a catalog of spatial relations that are used as defaults in the relational completion (cf. Table 3.4). This catalog determines that, for example, a 'neutral' orientation relation is added when a topological 'inside' relation holds between two entities, or that a 'disjoint' topological relation is added when a specific orientation relation is given between two entities. This catalog is represented in the *completion parameters* representation in MIRAGE.

The completion parameters' representation is modeled by an association list that assigns a default relation to every given relation. For modifying the spatial relations that are added by the relational completion process, the image specification process modifies the representation of the completion parameters to cause employing another than the usual set of completion parameters (see Fig. 4.10). The relational completion process will then generate another enriched representation based on the modified set of completion parameters. With this newly generated enriched representation, the image specification process retries to

produce a consistent specified image. This modification of completion parameters may be repeated until a consistent image specification succeeds (as in the example above, Fig. 4.9) or until an alternative image construction strategy is chosen.

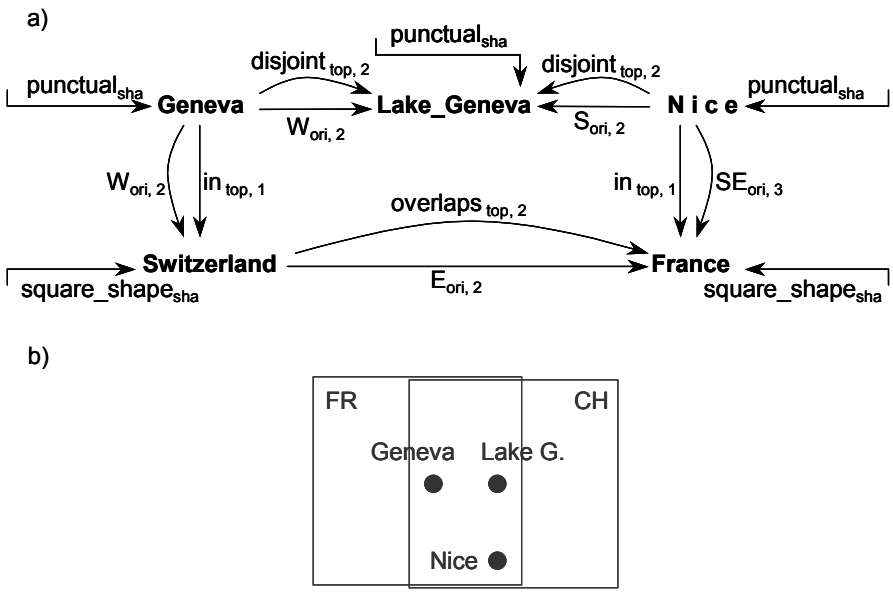

Fig. 4.9. a) An enriched representation that allows for a straightforward image specification and visualization; b) the corresponding image evoked in the visual buffer

Although the amount of possible combinations in completion parameters may be considerable (cf. Section 4.3.2), this strategy of producing a consistent image may be sensible. If a set of suitable completion parameters is found, the image construction can proceed in the same direct way as in the standard case.

4.4.3.2 Relaxation of Relational Completion

The other variant of revising the relational completion process is *reducing* the spatial relations that are added by the relational completion process. Although this option is incompatible with the overall idea of the image construction subsystem in MIRAGE (i.e., that working memory representations are made more and more specific for finally being visualized in the visual buffer) it nevertheless offers sensible options for constructing a mental image representation. The core idea of reducing the relational completion is that one or more of the spatial relations that are added in the relational completion process are abandoned to leave room for variations during the interpretation of the image specification process: the task of determining the spatial relations missing in the activated long-term memory representation is left to the image specification process.

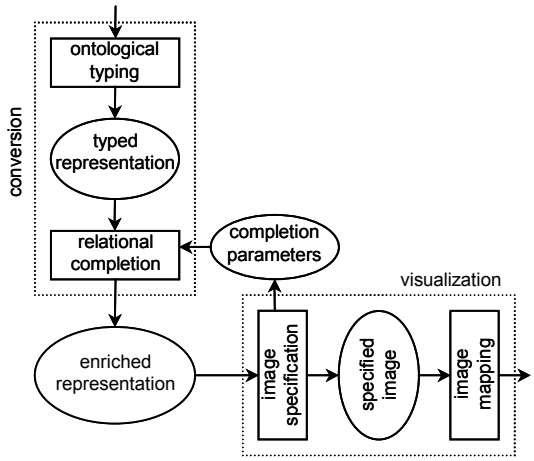

Fig. 4.10. Extension of MIRAGE to enable variations in the relational completion process. The image specification modifies the completion parameters that are used by the relational completion process

For example, a topological relation (say, an 'inside' relationship) is usually complemented by an orientation relation (for example 'neutral'). When omitting this orientation relation as a relaxation of the relational completion, and just leaving the topological 'inside' relation specified in the activated long-term memory representation, the entity can be placed at *any* location within the reference entity. The entire area of the reference object is treated as a potential location, thus allowing for a flexible interpretation performed by the image specification process (see Section 4.4.4.1).

The mechanism for relaxing the relational completion in MIRAGE is the same as above (cf. Section 4.4.3.1): the image specification process causes a modification in the completion parameters, which then force the relational completion process to leave certain relations unspecified.

Reducing the relational completion is especially interesting in the case of determining the shapes of extended entities. Normally, a standard shape is assigned to extended entities that do not come with a specific shape of their own. When the specifications of shapes are relaxed in the relational completion process, arbitrary shapes can be employed in the image specification process. This option allows for more complex visualizations and extends the possibilities for finding an adequate image specification. I will explain this option in the following section.

4.4.4 Revision of Image Specification

The image specification process uses the enriched representation produced by the conversion process (i.e., the ontological typing process and the relational comple-

tion process) to assign specific values to the represented entities. Its purpose is to interpret the provided qualitative spatial relations and to convert them into quantitatively specified representations. This process allows for variations in the interpretation of the qualitative relations. These variations will be explained in Section 4.4.4.1.

Besides interpreting given qualitative spatial relations, the image specification process also has to define spatial properties and relations that have not been specified by the relational completion process: as motivated in Section 4.4.3.2, a sensible strategy for constructing a mental image may be to leave the specification of spatial properties or relations to the image specification process. This may especially be sensible with respect to default shapes assigned to spatially extended entities. I will explain this case in Section 4.4.4.2.

4.4.4.1 Depicting Qualitative Spatial Relations

In Section 4.3.3 it has been shown that qualitative spatial relations may be interpreted as point locations or as sectors. In the standard case, relations are interpreted as point locations. When the image specification process indicates a conflicting situation, it is sensible to modify the way in which the spatial relations are interpreted. To illustrate this option let us consider the situation shown in Fig. 4.11. Figure 4.11a depicts the enriched representation of the 'Geneva – Nice' scenario. This representation resembles the situation depicted in Fig. 4.6 with some differences in the orientation relations that complement the topological 'inside' relations. When trying to visualize this enriched representation in the usual way, we will encounter quite the same situations as depicted in Fig. 4.3: there is no image construction found that is consistent with all relations contained in the enriched representation.

An alternative way of interpreting orientation relations (besides the prototypical interpretation as point locations used so far) is to assign *sectors* to each direction. The sector that is assigned to a direction is understood as the potential area in which an entity may be located. In Fig. 4.11b a visualization is seen that uses a sector interpretation for the spatial knowledge fragments 'S (Nice, Lake_Geneva)' and 'W (Geneva, Lake_Geneva)'. The position of 'Lake_Geneva' can be determined such that both relations hold.

The interpretation of the qualitative relations provided by the enriched representation is done by the image specification process. Also, the variations in these interpretations are performed by this process. Thus, a representation is required that controls the interpretation of the qualitative spatial relations during image specification. For this purpose, a representation of *specification parameters* is used by the image specification process. When the way qualitative relations are interpreted by the image specification process is changed, image specification may interact with its specification parameters. Parts of this set of specification parameters may be changed to allow for other interpretations that may enable a visualization that cannot be found using the standard parameters (see Fig. 4.12).

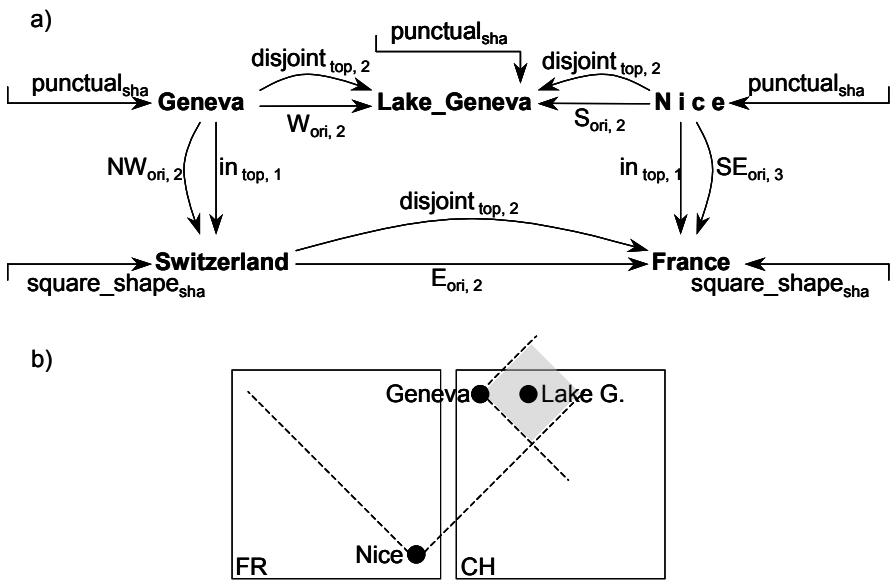

Fig. 4.11. Example of an enriched representation (a) and a visualization that (partially) uses sectors as interpretations of orientation relations (b)

4.4.4.2 Depicting Unspecified Spatial Relations

It has been motivated above (cf. Section 4.4.3.2) that it is sensible not to complement all spatial relations in the relational completion process. In this case, the spatial relations not specified in the enriched representation have to be determined by the image specification process. This is especially interesting with respect to the shapes of extended entities, when the default shapes are too restrictive to construct a mental image representation.

Figure 4.13a shows an enriched representation in which some spatial relations are left unspecified (indicated by the 'unspec' label in the figure). The topological and orientation relations between the point entities are still completely specified, whereas the extended entities are not assigned any shapes. Moreover, the topological relations involving the extended entities have not been complemented by orientation relations, i.e. the position of the point entities in the extended entities are not specified beyond their topological relationship. Consequently, relative positions of the point entities can be assigned in a straightforward manner in the image specification process. For the extended entities however, the given qualitative spatial relations have to be evaluated by the image specification process.

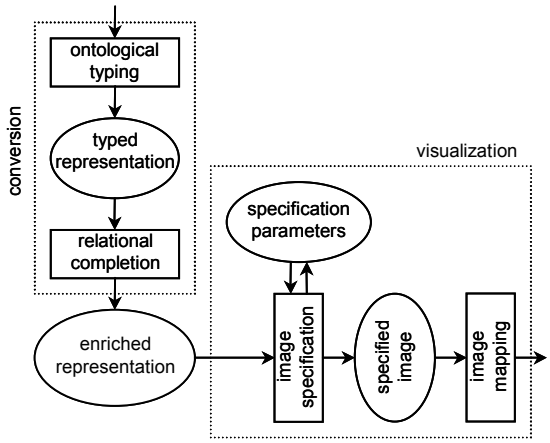

Fig. 4.12. Extension of the MIRAGE visualization system to enable modifications in the image specification process. Image specification modifies its specification parameters to enable varying interpretations of spatial relations

In Fig. 4.13b an image specification using the available pieces of knowledge is sketched out. On the basis of the topological relationships between the point entities and the extended entities, as well as based on the orientation relationship between the extended entities, partial representations of the extended entities are specified (indicated by the thin lines). These partial representations are subsequently used to construct the tentative shape of the extended entities. The resulting representation of the two extended entities is shown in Fig. 4.13c.

4.5 Summary

In this chapter I discussed several strategies that can be applied in MIRAGE's image construction subsystem when a straightforward image specification is not possible. MIRAGE has been extended by the structures and processes necessary to perform these alternative image construction strategies. Figure 4.14 shows the extended architecture of MIRAGE's image construction subsystem with all processes and representation structures.

Compared to the initial system (cf. Fig. 4.1), the conversion process has been subdivided into the *ontological typing* subsystem and the *relational completion* subsystem. The *typed representation* is the intermediate representation provided by the ontological typing subsystem. The visualization process has been subdivided into the *image specification* subsystem and the *image mapping* subsystem (with the *specified image* as intermediate representation). The *activated long-term memory reduction* system can remove spatial knowledge fragments retrieved from long-term memory from the activated long-term memory representation, which

yields the *reduced activated long-term memory representation*. For the relational completion and the image specification subsystems a representation of *completion parameters* and *specification parameters* has been introduced, respectively. These sets of parameters control the respective processes and can both be modified by the image specification process.

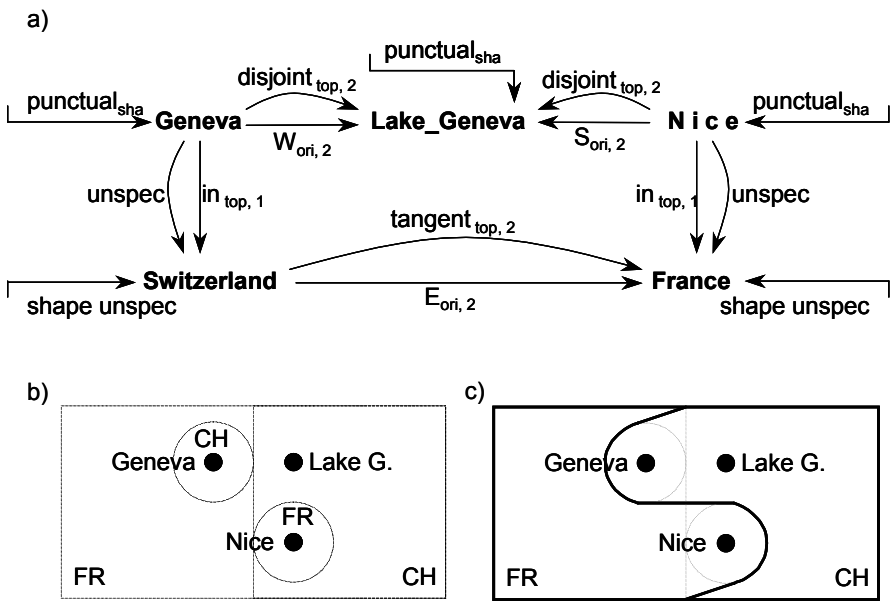

Fig. 4.13. An enriched representation with unspecified spatial relations (a); depiction of specified image on the basis of available spatial relations (b); construction of complex shapes for extended entities on the basis of partial representations (c)

These extensions of MIRAGE allow for a number of strategies to be employed when a straightforward image construction is not possible due to conflicts in the image specification process. These strategies are the following (see Fig. 4.15):

- assess the spatial knowledge fragments represented in the activated long-term memory representation and remove the least reliable facts from this representation; continue with conversion process (ontological typing);
- modify relational completion, or reduce qualitative spatial relations or default shapes added by relational completion; continue with image specification;
- vary the interpretation of qualitative spatial relations in the image specification process; determine positions of entities not restricted by qualitative spatial relations due to relaxation in the conversion process;
- ignore conflicting facts in the enriched representation during image specification and map the partial image in the visual buffer for inspection;

- abort image specification when computational or time resources do not allow for further image generation attempts, reject image.

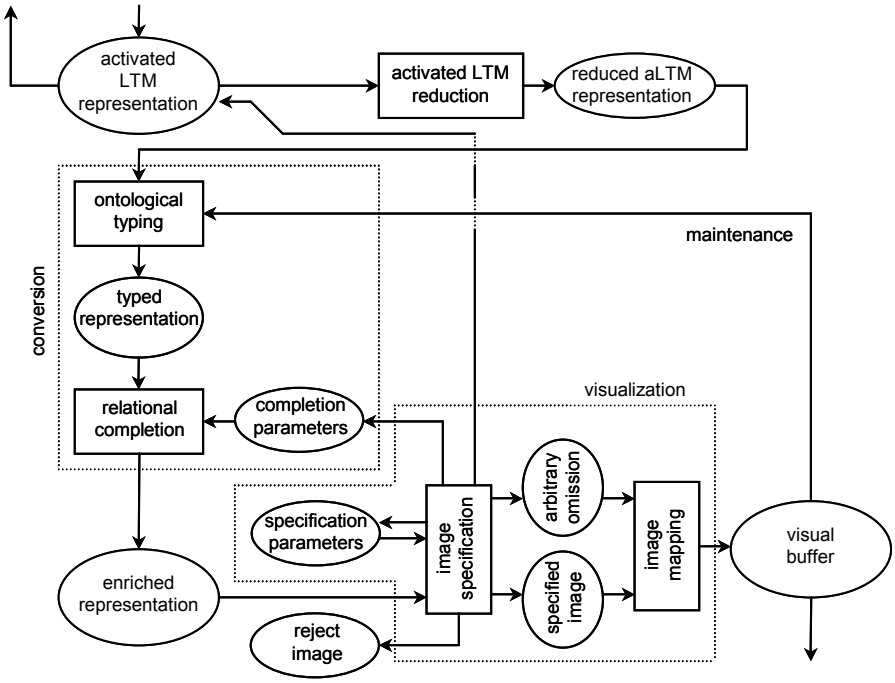

Fig. 4.14. MIRAGE's extended image construction subsystem

This chapter closes with some remarks about the presented options for image constructions. First, the strategies developed have different effects on the resource requirements of the image construction task. Whereas the variation of relational completion is neutral with respect to the resources required in the image construction task, the relaxation of relational completion causes more complicated operations in the image specification process. Also, varying the interpretation of relational completion towards treating cardinal directions as sectors, for instance, increases the complexity in image specification. The omission of spatial facts from activated long-term memory or during image specification, on the other hand, reduces the complexity of the problem at the cost of some spatial information not being realized in the constructed mental image.

Second, the strategies developed in this chapter are not independent of each other. For example, a reduction of relational completion causes more complex operations in the image specification, since spatial entities are less constrained by qualitative spatial descriptions. Or, variations in the relational completion may require certain types of interpretations in the image specification process.

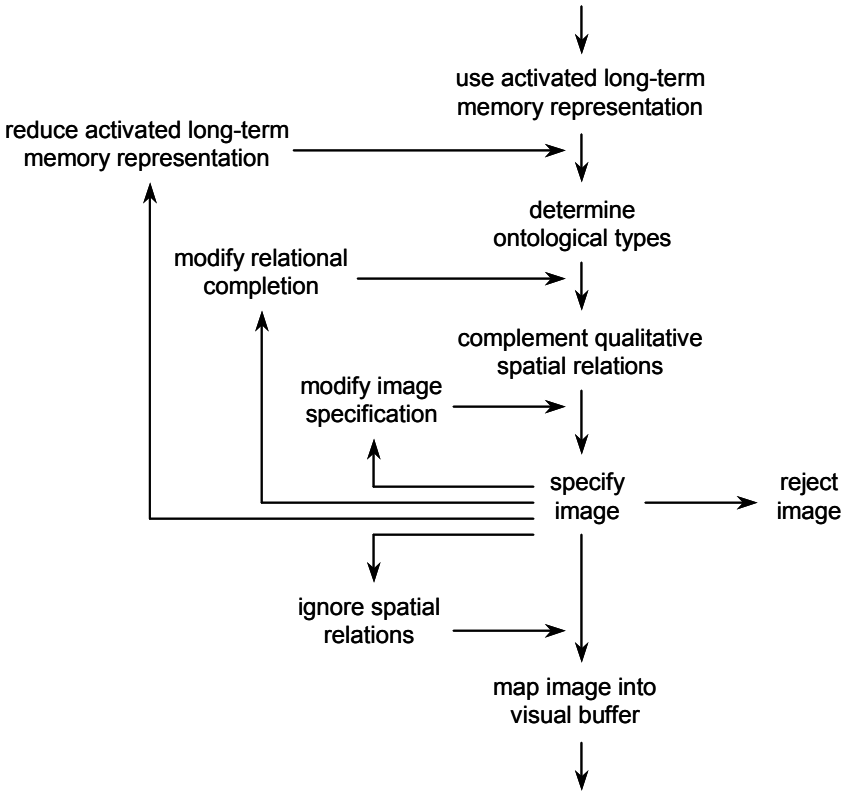

Fig. 4.15. Overview of strategies that can be employed when a straightforward image construction fails due to inconsistencies in the image specification process

Third, there is no determinism in MIRAGE that specifies which strategies are applied in which order or combination. The control of possible combinations of the image construction strategies is not part of the model and can vary depending on:

- types of situations or domains,
- inter-individual differences,
- resource requirements involved in the task to be performed, or
- additional knowledge used to solve the spatial problem under consideration.

In the following chapter I will show how MIRAGE has been realized in a prototypical implementation.

5 MIRAGE Implementation

In this chapter I will describe the prototypical implementation of MIRAGE. This implementation is intended as a first approach to demonstrate how mental images can be constructed in form of a computational model on the basis of underdetermined geographic knowledge from long-term memory. MIRAGE is realized as a diagrammatic reasoning system that can be used for experimentation in a computational modeling environment. The implementation of MIRAGE is based on the diagrammatic reasoning framework SIMSIS.

First, I will describe the SIMSIS system. Second, I will show how MIRAGE has been realized on the basis of the conception of SIMSIS, and I will explain how the characteristics claimed in Chapter 3 have been realized. Third, I will describe the behavior of the implemented model using the exemplary 'Reno – San Diego' scenario introduced in Chapter 3.

5.1 Computational Tools for Modeling: SIMSIS

In this section I will describe the SIMSIS[49] system (Klann, 1998) that is used as the basis for the implementation of MIRAGE. We developed SIMSIS in the research project *Spatial Structures in Aspect Maps* (Spatial Cognition priority program, DFG[50]). SIMSIS is a toolkit for modeling interpretation and construction of map-like depictions of spatial information (*aspect maps*). It is implemented in Common Lisp using the object-oriented Common Lisp Object System (CLOS) as modeling paradigm.

---

[49] SIMSIS stands for System for the development of Intelligent Methods for Spatial Information Systems (see http://www.informatik.uni-hamburg.de/WSV/karten/SIMSIS_manual/manual.html).

[50] Deutsche Forschungsgemeinschaft (German Science Foundation)

5.1.1 The Idea of SIMSIS

Map-like representations (like route maps, tourist city maps, or topographic reference maps) represent spatial information in a spatio-analogical way. Due to space restrictions in the representation medium, they must be restricted to a number of *aspects* they are intended to convey. Further aspects of the represented environment have to be omitted. As pictorial representations, map-like representations are over-specific in so far as they contain information that has no correspondence with the environment they represent (see below). Thus, certain types of information legitimately may be read off a given map, whereas others must not. The SIMSIS system has been developed to model the construction and the interpretation of *aspect maps.*

5.1.1.1 The Aspect Map Model

An *aspect map* "is a formal description of a map that allows to distinguish between intended or representational pieces of information and information that can be read off the map due to the pictorial property of over-representation" (Barkowsky & Freksa, 1997). Figure 5.1 shows an example of an aspect map. The figure shows a part of the public transportation network map of Hamburg, Germany. This map depicts spatial aspects that may be read off the map (for example the succession of stations on a specific underground line), and others that would result in an over-interpretation of this map and therefore must not be read off (for example cardinal directions between pairs of stations).

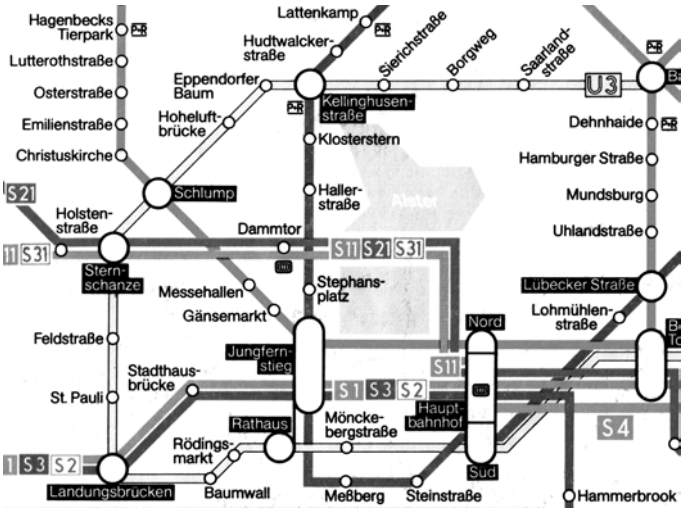

Fig. 5.1. Example of an aspect map: detail of the Hamburg public transportation network map

Figure 5.2 shows how spatial knowledge is processed in the aspect maps model. In the figure it is illustrated how information from two different aspect maps is used to generate some new piece of information, which is visualized in a new aspect map. The original aspect maps are first evaluated to obtain their legitimate information content. The legitimate information content is processed, i.e., spatial inference is performed to derive the required result from the available information. The result is then represented in another aspect map, which visualizes the inferred pieces of information.

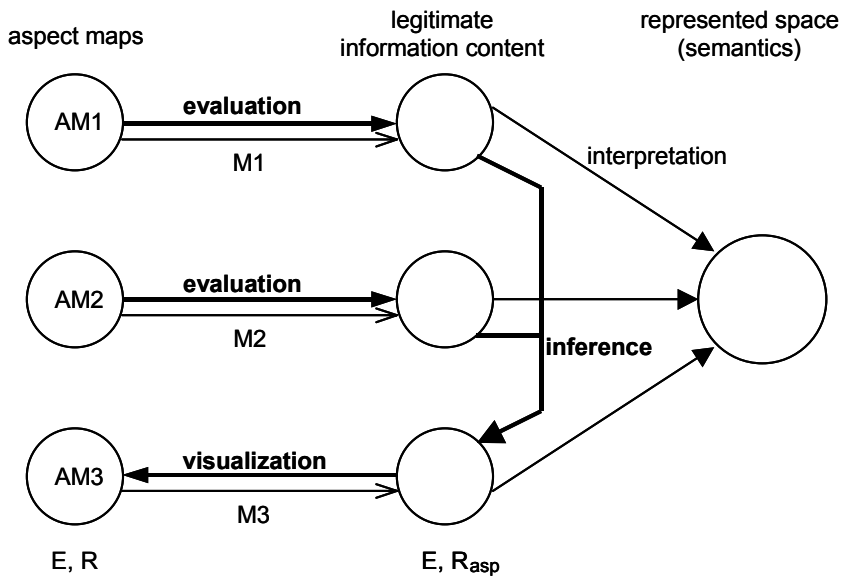

Fig. 5.2 Processing spatial knowledge in aspect maps. Legitimate information content of two different aspect maps is evaluated. Resulting information is obtained by spatial inferences. This resulting information is visualized in another aspect map (Berendt et al., 1998a: 323)

All involved aspect maps need to have a common spatial semantics to enable the combination of their content. So every piece of spatial information involved needs to be linked to a common interpretation process that ensures the mutual correspondence of the involved information.

In an empirical study we investigated how people can make use of two given maps in constructing a new representation that exhibits spatial information not contained in either of the two given maps (Berendt et al., 1998b). The various possible spatial reasoning strategies that may be applied in this task (both legitimately and illegitimately, i.e., in the sense of over-interpretation) have been demonstrated using the SIMSIS framework (Klann, 1998).

5.1.1.2 Modeling Aspect Maps in SIMSIS

SIMSIS distinguishes two general types of data structures: *pictures*, which are used to represent aspect maps, and *scenarios*, which represent corresponding spatial information in non-pictorial form. The spatial information contained in a scenario is either the result of an analysis of a SIMSIS picture (*picture analysis*), or it is the result of inference processes operating on a given SIMSIS scenario. On the other hand, scenarios can also contain information that is used to construct a new picture that visualizes the facts it holds (*picture synthesis*). Figure 5.3 presents an overview of the SIMSIS architecture.

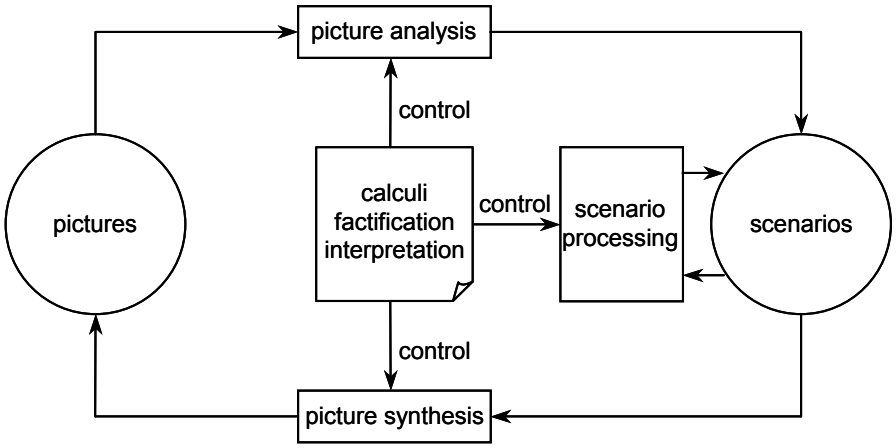

Fig. 5.3. Outline of the SIMSIS architecture (cf. Klann, 1998: 12)

Picture analysis is the interpretation of a given picture yielding spatial facts that hold between the entities represented in the picture. This process is also called *factification* in SIMSIS, as it produces facts by making information implicitly contained in the picture explicit for representation in a scenario. The process of *interpretation*, though commonly used for evaluating the content of diagrammatic representations, is conceived in SIMSIS in a broader sense: every transformation between representation systems based on semantic assumptions is called an interpretation (see Section 5.1.2.3). So picture analysis, picture synthesis, as well as transformations between scenarios are modeled as a form of interpretation in SIMSIS.

5.1.2 Depictions, Scenarios, and Interpretations

As explained above, processing map-like representations in SIMSIS is based on three major data structures and processing instances:

1. *pictures*, which contain aspect maps,
2. *scenarios*, which are based on *facts*, and
3. *interpretations* between the former two.

I will explain these three instances in the following.

5.1.2.1 SIMSIS Pictures

The SIMSIS data structure *picture* is used to describe a depiction by specifying its elements in terms of their respective graphic coordinates. Pictures are built up from the basic elements *polyline*, *circle*, *rectangle*, *polygon*, and *text*. Besides the specification of picture elements and their positions, the graphical extension of the picture as a whole is specified in the picture description. As picture elements potentially can be infinite (for instance, in the case of open polygons), the picture's extension is required for the visualization of the picture on a finite output device (for instance, in a window on a computer screen). As an example, Fig. 5.4 shows the description of a picture representing a detail of the Hamburg public transportation network together with the image that is generated by this description in SIMSIS.

```
(defparameter tmap
  (picture "topographic net map"
    (list
      (polyline (list (point 0 1.65)
                      (point 6.4 1.9)
                        (point 8.4 0))
                        :linecolor *grey*
                        :linewidth 0.1)
[...]
      (circle (point 8.1 0.3) 0.3
                      :linestyle 'none
                      :fillcolor *green*)
      (text "S" (point 8.1 0.3) 0.3 0.3
                        :fillcolor *white*)
      (rectangle (point 1.5 2) 0.5 0.5
                        :linestyle 'none
                        :fillcolor *blue*)
      (text "U" (point 1.5 2) 0.3 0.3
                      :fillcolor *white*)
[...]
      (text "Dammtor" (point 8.1 0.8) 1.8 0.3
                      :fillcolor *black*))
    (point 0 10) (point 10 0)))
```

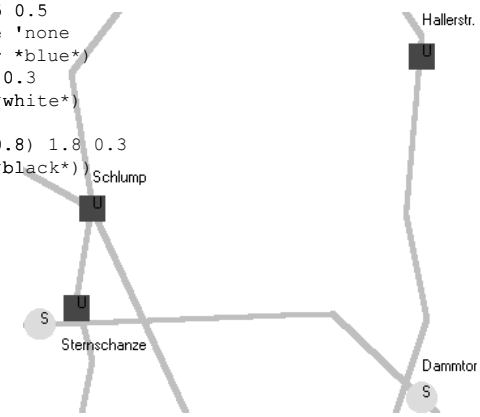

Fig. 5.4. The definition of a SIMSIS picture ("[...]" indicates omissions) together with the image output it generates (cf. Klann, 1998: 23)

Processing facilities for picture data structures implemented in SIMSIS comprise methods for adding elements to a picture, as well as methods for accessing

all or selected picture elements. Picture transformations, interpretations, and visualizations are performed using the SIMSIS *interpretation* concept, which I will explain in Section 5.1.2.3.

5.1.2.2 SIMSIS Facts and Scenarios

Facts in SIMSIS are the data structures that are used to represent the information gained by analyzing a picture, or that are used for being visualized in a picture. Facts comprise both predicates and (n-ary) relations between spatial entities. SIMSIS predicates (or unary relations) are labels (which are used to refer to an object by a given name) or position relations (which assign a point position to an entity). SIMSIS relations are, for instance, (quantitative or qualitative) orientation relations between two entities.

Facts that describe properties of and relations between entities are organized in scenarios. So all facts that are obtained as the result of a picture analysis or that are to be visualized in a picture are comprised in a common scenario. Note that one given picture may yield diverse scenarios (depending on the way it is interpreted); a given scenario, in turn, can be used to produce various pictures, depending on how the facts contained in this scenario are visualized.

Besides the facts themselves, every scenario organizes a list of the spatial entities it is about (i.e., the entities that are referred to in the facts represented in the scenario). Scenarios are operated on by processes for storing (*register*, *register-named*) and retrieving objects and facts (*getobj*, *getfact*). The methods *findobjects* and *findfacts* are used for retrieving objects and facts that refer to a required piece of information. Figure 5.5 shows an extract of an exemplary description of a SIMSIS scenario.

```
#<GEOMETRIC-SCENARIO "Geometric scenario of topographic map"
  scenario-objects:
    1.) #<U-BAHN-HALTESTELLE: Hallerstraße>
    2.) #<U-BAHN-HALTESTELLE: Dammtor>
 [...]
  scenario-facts:
    1.) #<LABEL "Hallerstraße" denotes
          #<U-BAHN-HALTESTELLE: Hallerstraße>>
    2.) #<LABEL "Dammtor" denotes
          #<U-BAHN-HALTESTELLE: Dammtor>>
 [...]
    5.) #<POSITION-RELATION:
          #<U-BAHN-HALTESTELLE: Hallerstraße>
            is at (POINT 8.10 6.80)>
    6.) #<POSITION-RELATION:
          #<U-BAHN-HALTESTELLE: Dammtor>
            is at (POINT 8.10 0.30)>
 [...]
>
```

Fig. 5.5. Example of a SIMSIS scenario description

5.1.2.3 SIMSIS Interpretations and Meaning Systems

The concept of *interpretation* is used in SIMSIS for every transformation between data structures, i.e., for the analysis of entities represented in pictures, for processing facts represented in scenarios, as well as for the visualization of facts in a picture. Even the transformation of a given picture structure to produce an image on an output device (for example on a computer screen or in an output file) is done by interpretation processes.

Interpretation in SIMSIS is based on theoretical concepts known from semiotics. Semiotics is the discipline that examines the relationship between symbols and symbol systems on the one hand and the meaning underlying them on the other (e.g., Eco, 1976; Noeth, 1990). In a simple scheme according to Richards and Ogden (1923) meaning is assigned to symbols by the relationship between three components: the symbol, the reference, and the referent (Fig. 5.6). The symbol is interpreted using the reference (i.e., the correspondence relation between the symbol and the object it denotes) to refer to the referent (the object it denotes). The relation between these three components is called a *signification system* (Eco, 1976).

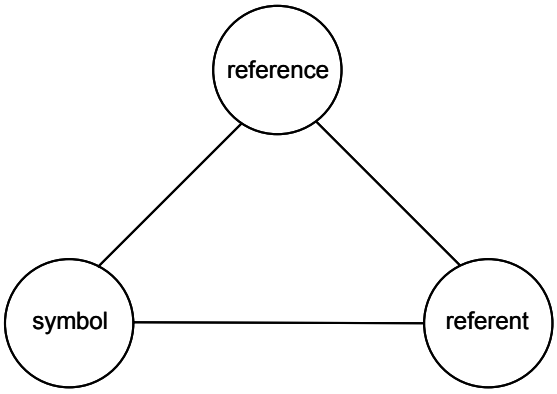

Fig. 5.6. Schematic depiction of a *signification system*: the relationships between *symbol*, *reference*, and *referent* constitute meanings of symbols

In this conception of a signification system, it is important to note that the representation of the reference can only be done by providing another symbol. This means that the meaning of a symbol always is explained by another symbol, which leads to an infinite recourse of using symbols for explaining the meaning of other symbols.

Many authors have applied semiotic concepts to explain the interpretation and processing of diagrammatic and map-like representations (for an overview see Head, 1991). In SIMSIS, the above semiotic conception is modeled using *meaning*

systems, which control interpretation processes. The function of a meaning system lies in defining the meaning of symbols. As the meaning of a symbol can only be given by another symbol, the interpretation of a given symbol is done by transferring this symbol from one meaning system to another one. So besides the symbol to be interpreted the interpretation process requires the meaning system the symbol originally belongs to, as well as the new meaning system in which the symbol will be interpreted. From a technical (i.e., implementation-related) point of view, meaning systems are used to select the appropriate SIMSIS interpretation method (*interpret*) for a given object in the required interpretation context.

To enable the interpretation of pictures with respect to their geometrical content (i.e., the factification of picture elements) and vice versa (i.e., the visualization of scenario facts in a picture) specific meaning systems are predefined in SIMSIS (*picture-mean-sys* and *geometric-mean-sys*).

5.2 Realization of the Model

In this section I will explain how the concepts of the SIMSIS system are used to implement the MIRAGE model. Figure 5.7 gives an overview of how the components of MIRAGE are modeled in types of SIMSIS structures and processes. The figure is arranged according to the schematic depiction of the SIMSIS architecture presented in Fig. 5.3. The long-term memory structures of hierarchical long-term memory, activated long-term memory, and the enriched representation are modeled through SIMSIS scenarios. All processes are interpretation methods in SIMSIS. The visual buffer is realized on the basis of a picture data structure. Differing from Fig. 5.3, SIMSIS facts are explicitly depicted in Fig. 5.7 because they are used for modeling the spatial knowledge fragments yielded by the access process, as well as for the results of the inspection process. All structures and processes are further explained in the following.

5.2.1 MIRAGE Structures

5.2.1.1 Entities, Relations, and Spatial Knowledge Fragments

Spatial knowledge fragments encode pieces of information about geographic entities. More specifically, spatial knowledge fragments are used to store topological relations, cardinal directions, and shape information in MIRAGE.

For modeling geographic entities, the Lisp class *geo_entity* has been defined. An instance of *geo_entity* is characterized by the name of the entity it represents. A geographic entity is created through the Lisp method *geo_entity*, which is called with the name of the object to create as its parameter:

```
> (geo_entity "California")

#<GEO_ENTITY: California>
```

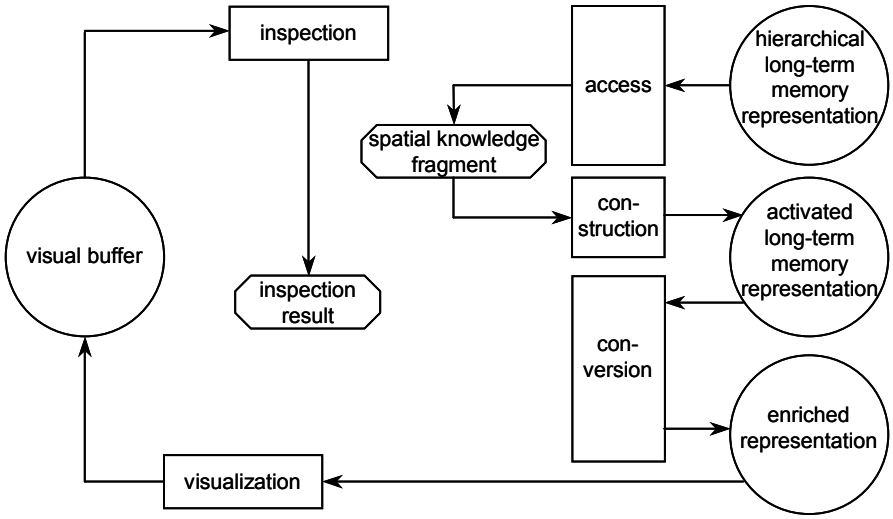

Fig. 5.7. Overview of how the MIRAGE components are modeled through SIMSIS structures and processes. Processing starts with the hierarchical long-term memory representation and returns the inspection result

Spatial knowledge fragments are modeled by the Lisp class *SKF*, which is a direct subclass of the SIMSIS class *fact*. The class *fact* has one slot, which encodes the level of resolution of the spatial knowledge fragment. The specific spatial relations to be represented in the spatial knowledge fragments, i.e. topological relations, cardinal directions, and shape properties, are each modeled by their own class: the classes *topological-relation*, *cardinal-direction*, and *shape* are defined as subclasses of *SKF*. Each of these classes has its own method to create an instance of the respective type of spatial knowledge fragment, for instance:

```
> (topological-relation (geo_entity "San Diego") "in"
(geo_entity "California") :resolution 1)

#<SKF: #<GEO_ENTITY: San Diego> in #<GEO_ENTITY:
California> (TOPOLOGICAL-RELATION, res 1)>


> (cardinal-direction (geo_entity "Nevada") "E" (geo_entity
"California") :resolution 2)

#<SKF: #<GEO_ENTITY: Nevada> E #<GEO_ENTITY: California>
(CARDINAL-DIRECTION, res 2)>


> (shape (geo_entity "Nevada") '((point  6 19)
                                 (point 15 19)
                                 (point 15  6)
```

```
                                    (point  6 13)
                                    (point  6 19))
                                    :resolution 1)

#<SKF: SHAPE of #<GEO_ENTITY: Nevada> is ((POINT 6 19)
(POINT 15 19) (POINT 15 6) (POINT 6 13) (POINT 6 19))
(resolution 1)>
```

Besides for encoding specific shapes of geographic entities the class *shape* is also used for representing default shapes of extended entities or for encoding an entity as punctual:

```
#<SKF: SHAPE of #<GEO_ENTITY: Nevada> is square_shape
(resolution NIL)>

#<SKF: SHAPE of #<GEO_ENTITY: San Diego> is punctual
(resolution NIL)>
```

Figure 5.8 gives an overview of the class hierarchy of the classes predefined in SIMSIS and additionally defined in MIRAGE. So far, the subclass *SKF* of the SIMSIS class *fact* has been described together with its subclasses. The SIMSIS class *label*, which is also a subclass of *fact* is used for registering names of geographic entities in scenarios. I will explain MIRAGE's scenarios in the following two subsections.

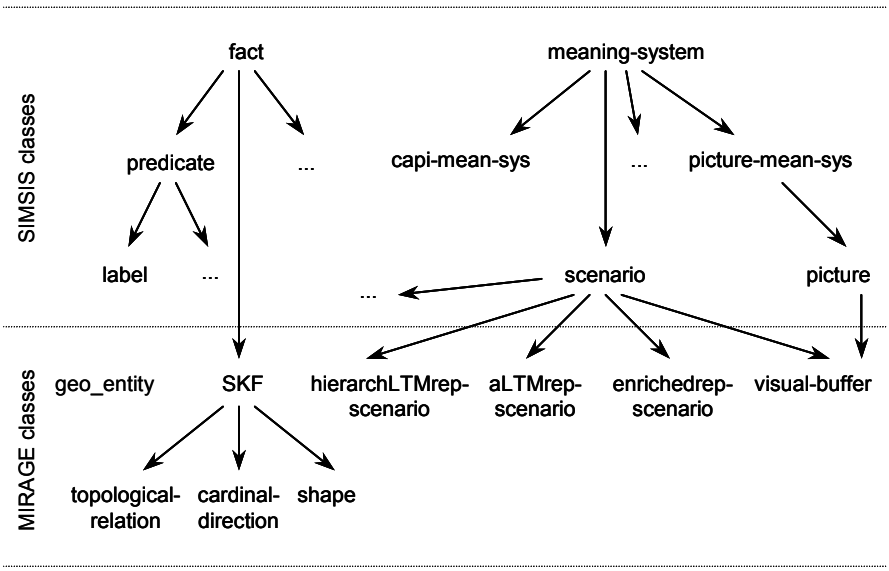

Fig. 5.8. Hierarchical structure of the Lisp classes predefined in SIMSIS and additionally defined in MIRAGE

5.2.1.2 The Long-Term Memory Representations

All long-term memory representation structures (except the spatial knowledge fragments explained above) are SIMSIS scenarios. The classes *hierarchLTMrep-scenario*, *aLTMrep-scenario*, and *enrichedrep-scenario* are direct subclasses of the SIMSIS class *scenario* (which itself is a SIMSIS meaning system, see Section 5.1.2.3). From the SIMSIS class *scenario* the classes inherit the slots *name*, *objects*, and *facts*. So each long-term memory structure contains the list of objects it is about, together with the list of facts (i.e., spatial knowledge fragments) that are stored in the structure.

As an example I present the scenario that represents the hierarchical long-term memory structure discussed in Chapter 3 (cf. Fig. 3.6):

```
#<HIERARCHLTMREP-SCENARIO "hierarchical LTM representation"
  scenario-objects:
     1.) #<GEO_ENTITY: Reno>
     2.) #<GEO_ENTITY: Nevada>
     3.) #<GEO_ENTITY: San Diego>
     4.) #<GEO_ENTITY: California>
     5.) #<GEO_ENTITY: U.S.>
  scenario-facts:
     1.) #<LABEL "Reno" denotes #<GEO_ENTITY: Reno>>
     2.) #<LABEL "Nevada" denotes #<GEO_ENTITY: Nevada>>
     3.) #<LABEL "San Diego" denotes #<GEO_ENTITY: San
         Diego>>
     4.) #<LABEL "California" denotes #<GEO_ENTITY:
         California>>
     5.) #<LABEL "U.S." denotes #<GEO_ENTITY: U.S.>>
     6.) #<SKF: #<GEO_ENTITY: Nevada> in #<GEO_ENTITY:
         U.S.> (TOPOLOGICAL-RELATION, res 1)>
     7.) #<SKF: #<GEO_ENTITY: California> in #<GEO_ENTITY:
         U.S.> (TOPOLOGICAL-RELATION, res 1)>
     8.) #<SKF: #<GEO_ENTITY: Reno> in #<GEO_ENTITY:
         Nevada> (TOPOLOGICAL-RELATION, res 1)>
     9.) #<SKF: #<GEO_ENTITY: San Diego> in #<GEO_ENTITY:
         California> (TOPOLOGICAL-RELATION, res 1)>
    10.) #<SKF: #<GEO_ENTITY: Nevada> tangent #<GEO_ENTITY:
         California> (TOPOLOGICAL-RELATION, res 2)>
    11.) #<SKF: #<GEO_ENTITY: Nevada> NE #<GEO_ENTITY:
         California> (CARDINAL-DIRECTION, res 3)>
    12.) #<SKF: #<GEO_ENTITY: Nevada> E #<GEO_ENTITY:
         California> (CARDINAL-DIRECTION, res 2)>
    13.) #<SKF: #<GEO_ENTITY: Reno> W #<GEO_ENTITY: Nevada>
         (CARDINAL-DIRECTION, res 2)>
    14.) #<SKF: #<GEO_ENTITY: San Diego> S #<GEO_ENTITY:
         California> (CARDINAL-DIRECTION, res 2)>
    15.) #<SKF: SHAPE of #<GEO_ENTITY: California> is
         ((POINT -0.50 19.50) (POINT 6.00 19.50) (POINT
         6.00 13.00) (POINT 15.00 6.00) (POINT 16.00 1.00)
         (POINT 11.00 0.00) (POINT 6.00 3.50) (POINT 0.00
         13.00) (POINT -0.50 19.50)) (resolution 1)>
    16.) #<SKF: SHAPE of #<GEO_ENTITY: Nevada> is ((POINT
```

```
6.00 19.50) (POINT 15.00 19.50) (POINT 16.00 8.00)
(POINT 15.00 6.00) (POINT 6.00 13.00) (POINT 6.00
19.50)) (resolution 1)>
```

Examples for the other long-term memory representations will be given in Section 5.2.2 where the processes operating on these representation structures will be explained.

5.2.1.3 The Visual Buffer

As can be seen in Fig. 5.8, the *visual-buffer* data structure is a subclass of both the SIMSIS class *scenario* and the SIMSIS class *picture*. This is done for being able to link geographic entities via their names to the picture elements represented in the visual buffer. The visual buffer data structure contains as its slots the list of objects it contains and the list of facts (both slots are inherited from the class *scenario*). From the class *picture* the *visual-buffer* class inherits the slot that contains the list of the picture elements. The SIMSIS fact *label* is used to link the picture elements to the geographic entities they depict.

As an example for a visual buffer I present the representation of the image generated in Section 3.4.2.4 (cf. Fig. 3.11):

```
VISUAL-BUFFER "MIRAGE Visual Buffer"
  objects:
    1.) #<GEO_ENTITY: Reno>
    2.) #<GEO_ENTITY: Nevada>
    3.) #<GEO_ENTITY: California>
    4.) #<GEO_ENTITY: San Diego>
  facts:
    1.) #<LABEL "Reno" denotes #<GEO_ENTITY: Reno>>
    2.) #<LABEL "Nevada" denotes #<GEO_ENTITY: Nevada>>
    3.) #<LABEL "California" denotes #<GEO_ENTITY:
        California>>
    4.) #<LABEL "San Diego" denotes #<GEO_ENTITY: San
        Diego>>
  picture-elements:
    1.) <CIRCLE with centre (POINT 50.00 50.00) and radius
        1>
    2.) <RECTANGLE through (POINT 40.00 40.00)
                           (POINT 60.00 40.00)
                           (POINT 60.00 60.00)
                           (POINT 40.00 60.00)
                           (POINT 40.00 40.00)>
    3.) <RECTANGLE through (POINT 10.00 40.00)
                           (POINT 30.00 40.00)
                           (POINT 30.00 60.00)
                           (POINT 10.00 60.00)
                           (POINT 10.00 40.00)>
    4.) <CIRCLE with centre (POINT 20.00 50.00) and radius
        1>
```

The picture elements occur in the same order as the *label* facts that provide the link between the geographic entities and their proper names. Observe that the picture elements are all spatially extended entities (i.e., point entities are represented as circles in the visual buffer).

For adding further entities to the visual buffer structure, as well as for retrieving entities from the visual buffer specific methods have been defined in MIRAGE (*add2vb* and *findpicelem*, respectively).

5.2.2 MIRAGE Processes

This section describes the processes operating in MIRAGE. First, I will describe the access and the construction processes of the long-term memory activation subsystem. Second, I will describe the conversion and the visualization processes of the image construction subsystem. Third, the image inspection process will be described.

5.2.2.1 The Long-Term Memory Activation Processes

The access process performs a graph search on the hierarchical long-term memory structure. It comprises the functions *initialize-access* and the *access* process proper. *Initialize-access* computes the sequence of spatial knowledge fragments that are retrieved from long-term memory based on cost criteria as described in Section 3.4.1.3. *Access* yields the retrieved spatial knowledge fragments, one at each time it is called. Initialize-access is called with the two entities to be related, the required type of spatial knowledge fragment, and the hierarchical long-term memory representation to be used.

To retrieve the spatial knowledge fragments necessary to elaborate on the cardinal direction of Reno with respect to San Diego using the above hierarchical long-term memory representation (Section 5.2.1.2) the following expression may be used:

```
> (initialize-access (findobj "Reno" hltmrep)
                     (findobj "San Diego" hltmrep)
                     'cardinal-direction
                     hltmrep)
```

According to the example, this function call results in the following list of nine spatial knowledge fragments:

```
(#<SKF: #<GEO_ENTITY: San Diego> in #<GEO_ENTITY:
California> (TOPOLOGICAL-RELATION, res 1)>

#<SKF: #<GEO_ENTITY: Nevada> E #<GEO_ENTITY: California>
(CARDINAL-DIRECTION, res 2)>

#<SKF: #<GEO_ENTITY: Reno> in #<GEO_ENTITY: Nevada>
(TOPOLOGICAL-RELATION, res 1)>
```

```
#<SKF: #<GEO_ENTITY: Nevada> NE #<GEO_ENTITY: California>
(CARDINAL-DIRECTION, res 3)>

#<SKF: #<GEO_ENTITY: San Diego> S #<GEO_ENTITY: California>
(CARDINAL-DIRECTION, res 2)>

#<SKF: #<GEO_ENTITY: Reno> W #<GEO_ENTITY: Nevada>
(CARDINAL-DIRECTION, res 2)>

#<SKF: #<GEO_ENTITY: Nevada> tangent #<GEO_ENTITY:
California> (TOPOLOGICAL-RELATION, res 2)>

#<SKF: SHAPE of #<GEO_ENTITY: California> is ((POINT -0.50
19.50) (POINT 6.00 19.50) (POINT 6.00 13.00) (POINT 15.00
6.00) (POINT 16.00 1.00) (POINT 11.00 0.00) (POINT 6.00
3.50) (POINT 0.00 13.00) (POINT -0.50 19.50)) (resolution
1)>

#<SKF: SHAPE of #<GEO_ENTITY: Nevada> is ((POINT 6.00
19.50) (POINT 15.00 19.50) (POINT 16.00 8.00) (POINT 15.00
6.00) (POINT 6.00 13.00) (POINT 6.00 19.50)) (resolution
1)>)
```

This list is used by *access* to provide the knowledge fragments one after the other:

```
> (access)

#<SKF: #<GEO_ENTITY: San Diego> in #<GEO_ENTITY:
California> (TOPOLOGICAL-RELATION, res 1)>

> (access)

#<SKF: #<GEO_ENTITY: Nevada> E #<GEO_ENTITY: California>
(CARDINAL-DIRECTION, res 2)>

> …
```

The construction process *construct* is implemented on the basis of the SIMSIS method *register*, which stores a fact in a scenario. The method *construct* is called with a spatial knowledge fragment and an activated long-term memory representation scenario as its parameters. The construction process first checks whether there is already a spatial knowledge fragment of the same type involving the same set of entities represented in the activated long-term memory representation. If this is the case, the spatial knowledge fragment is replaced; otherwise, the new spatial knowledge fragment is just added. The method *construct* has no return value and directly modifies the *scenario-facts* list in the activated long-term memory representation it is called with.

For instance the expression

```
> (construct current_skf aLTMrep)
```

stores the spatial knowledge fragment `current_skf` to the activated long-term memory representation `aLTMrep`.

A resulting activated long-term memory representation according to the example in Section 3.4.1.5 (cf. Fig. 3.9a) after integrating three spatial knowledge fragments is:[51]

```
#<ALTMREP-SCENARIO "aLTMrep"
  scenario-objects:
    1.) #<GEO_ENTITY: Reno>
    2.) #<GEO_ENTITY: Nevada>
    3.) #<GEO_ENTITY: San Diego>
    4.) #<GEO_ENTITY: California>
    5.) #<GEO_ENTITY: U.S.>
  scenario-facts:
    1.) #<LABEL "Reno" denotes #<GEO_ENTITY: Reno>>
    2.) #<LABEL "Nevada" denotes #<GEO_ENTITY: Nevada>>
    3.) #<LABEL "San Diego" denotes #<GEO_ENTITY: San
        Diego>>
    4.) #<LABEL "California" denotes #<GEO_ENTITY:
        California>>
    5.) #<LABEL "U.S." denotes #<GEO_ENTITY: U.S.>>
    6.) #<SKF: #<GEO_ENTITY: San Diego> in #<GEO_ENTITY:
        California> (TOPOLOGICAL-RELATION, res 1)>
    7.) #<SKF: #<GEO_ENTITY: Nevada> E #<GEO_ENTITY:
        California> (CARDINAL-DIRECTION, res 2)>
    8.) #<SKF: #<GEO_ENTITY: Reno> in #<GEO_ENTITY: Nevada>
        (TOPOLOGICAL-RELATION, res 1)>>
```

5.2.2.2 The Image Construction Processes

The *conversion* process is based on a SIMSIS interpretation method. It comprises the partial processes of *ontological typing* and *relational completion*. The parameters of the *conversion* method are an *aLTMrep-scenario* (see Section 5.2.1.2) and a set of *completion parameters*. As the result of the conversion process, an instance of the class *enrichedrep-scenario* is returned.

The enriched representation to be constructed by the conversion process contains all objects and facts that are represented in the underlying activated long-term memory representation. Therefore, all geographic entities and all facts contained in the activated long-term memory representation are copied to the enriched representation that is created.

---

[51] The first four *label* facts in this representation structure do not result from a construction operation. Rather, they are adopted from the underlying long-term memory representation when the activated long-term memory representation is instantiated. This is done to be able to refer to the same entities in both representations using the proper names of the geographic entities involved.

In the *ontological typing* process, an ontological type is assigned to all geographic entities in the representation structure. More specifically, for all entities for which no specific (polygonal) shape information is represented in the activated long-term memory representation, the shape information 'square_shape' is stored for entities that need to be extended entities. This is the case for all geographic entities A that are in one of the following topological relations:

contains (A, B); contains-at-border (A, B);
in (B, A); in-at-border (B, A);
tangent (A, B); tangent (B, A);
overlaps (A, B); or overlaps (B, A),

where A and B are geographic entities. All other entities are represented as punctual.

In the *relational completion* process every topological relation is completed by a cardinal direction between the respective pair of geographic entities, when no cardinal direction is specified in the underlying activated long-term memory representation, and vice versa. The relations are assigned using an association list which in the default case is arranged according to Table 3.4 (Section 3.4.2.2).

The enriched representation that results from the conversion process using the activated long-term memory representation shown above (Section 5.2.2.1) and the default set of completion parameters is:

```
#<ENRICHEDREP-SCENARIO "enriched rep"
  scenario-objects:
     1.) #<GEO_ENTITY: Reno>
     2.) #<GEO_ENTITY: Nevada>
     3.) #<GEO_ENTITY: San Diego>
     4.) #<GEO_ENTITY: California>
     5.) #<GEO_ENTITY: U.S.>
  scenario-facts:
     1.) #<LABEL "Reno" denotes #<GEO_ENTITY: Reno>>
     2.) #<LABEL "Nevada" denotes #<GEO_ENTITY: Nevada>>
     3.) #<LABEL "San Diego" denotes #<GEO_ENTITY: San
         Diego>>
     4.) #<LABEL "California" denotes #<GEO_ENTITY:
         California>>
     5.) #<LABEL "U.S." denotes #<GEO_ENTITY: U.S.>>
     6.) #<SKF: #<GEO_ENTITY: San Diego> in #<GEO_ENTITY:
         California> (TOPOLOGICAL-RELATION, res 1)>
     7.) #<SKF: #<GEO_ENTITY: Nevada> E #<GEO_ENTITY:
         California> (CARDINAL-DIRECTION, res 2)>
     8.) #<SKF: #<GEO_ENTITY: Reno> in #<GEO_ENTITY:
         Nevada> (TOPOLOGICAL-RELATION, res 1)>
     9.) #<SKF: SHAPE of #<GEO_ENTITY: Reno> is punctual
         (resolution NIL)>
    10.) #<SKF: SHAPE of #<GEO_ENTITY: Nevada> is
         square_shape (resolution NIL)>
    11.) #<SKF: SHAPE of #<GEO_ENTITY: San Diego> is
         punctual (resolution NIL)>
    12.) #<SKF: SHAPE of #<GEO_ENTITY: California> is
         square_shape (resolution NIL)>
```

```
13.) #<SKF: #<GEO_ENTITY: San Diego> neutral
     #<GEO_ENTITY: California> (CARDINAL-DIRECTION, res
     NIL)>
14.) #<SKF: #<GEO_ENTITY: Reno> neutral #<GEO_ENTITY:
     Nevada> (CARDINAL-DIRECTION, res NIL)>
15.) #<SKF: #<GEO_ENTITY: Nevada> disjoint
     #<GEO_ENTITY: California> (TOPOLOGICAL-RELATION,
     res NIL)>>
```

The *visualization* process is also based on a SIMSIS interpretation. It obtains an enriched representation as parameter and yields an instance of a visual buffer structure, which is returned. Visualization uses an image specification function that starts with positioning the first entity it is called with in the visual buffer; after that, all other entities represented in the enriched representation are located recursively. Using the example *enrichedrep-scenario*, visualization yields the visual buffer structure shown in Section 5.2.1.3.

5.2.2.3 Image Inspection

The image inspection process *inspection* analyses a visual buffer structure with respect to a given type of spatial relation, a pair of geographic entities, and an intended degree of resolution. Thus, inspection is called with the required SKF type, the names of two geographic entities, the required resolution, and the visual buffer to be inspected as its parameters. As result, inspection yields a spatial knowledge fragment.

For example, when using the above visual buffer structure (Section 5.2.1.3), inspection can be used in the following way to inspect the spatial orientation between ‘Reno’ and ‘San Diego’:

```
> (inspection 'cardinal-direction "Reno" "San Diego" 3 vb)

#<SKF: #<GEO_ENTITY: Reno> E #<GEO_ENTITY: San Diego>
       (CARDINAL-DIRECTION, res 3)>
```

Depending on the resolution chosen, the result of the inspection process may vary. When using a visual buffer representation like the one depicted in Fig. 3.13, calling inspection with different resolutions yields different results (cf. Table 3.5):

```
> (inspection 'cardinal-direction "San Diego" "Reno" 3 vb)

#<SKF: #<GEO_ENTITY: San Diego> S #<GEO_ENTITY: Reno>
       (CARDINAL-DIRECTION, res 3)>


> (inspection 'cardinal-direction "San Diego" "Reno" 4 vb)

#<SKF: #<GEO_ENTITY: San Diego> SSE #<GEO_ENTITY: Reno>
       (CARDINAL-DIRECTION, res 4)>
```

5.3 Operation and Behavior of MIRAGE

In this section I will describe a prototypical reasoning process in MIRAGE using the 'Reno – San Diego' example. As underlying knowledge the hierarchical long-term memory representation shown in Section 5.2.1.2 will be used (cf. Fig. 3.6). The construction of the working memory representation is done to obtain the spatial orientation of Reno with respect to San Diego.

The first spatial knowledge fragment that is provided by the access process states that San Diego is located in California:

```
#<SKF: #<GEO_ENTITY: San Diego> in #<GEO_ENTITY:
California> (TOPOLOGICAL-RELATION, res 1)>
```

The conversion process determines San Diego as a point entity, California as extended, and it adds a neutral orientation relation that locates San Diego in the center of California. The image produced by the visualization process on the basis of the enriched representation is depicted in Fig. 5.9a.[52] As the two entities whose orientation with respect to each other is to be determined are not contained in this visual buffer representation, the inspection process does not yield an orientation relation.

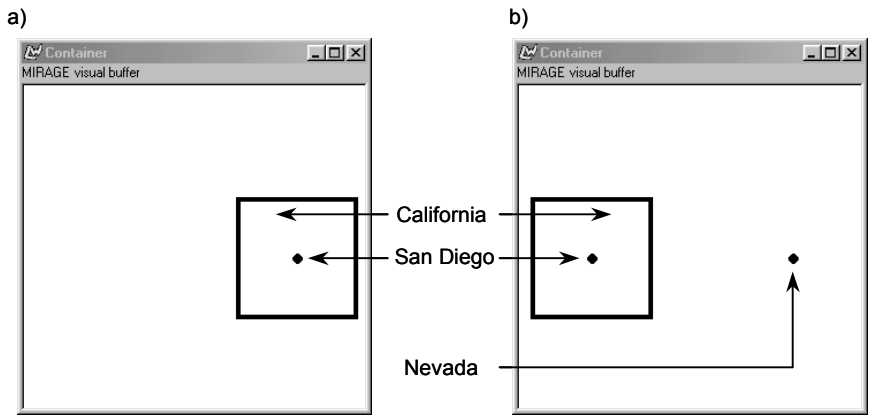

Fig. 5.9. Visual buffer representations resulting from the image construction based on (a) the first and (b) the first two spatial knowledge fragments retrieved from long-term memory

In the next access step Nevada is stated to be east of California:

```
#<SKF: #<GEO_ENTITY: Nevada> E #<GEO_ENTITY: California>
(CARDINAL-DIRECTION, res 2)>
```

[52] All figures in this section show output generated by the implemented model.

The conversion process is performed as described above. Additionally, Nevada is defined as a point entity by the ontological typing process. Although this seems odd at first sight, it is nevertheless sensible: the property of Nevada as extended is not yet required by any spatial relation in the activated long-term memory representation at the current processing stage. So for reasons of cognitive economy it is sufficient to envisage Nevada as punctual. The topological relationship between California and Nevada is determined as disjoint by the conversion process. The resulting visual buffer representation if shown in Fig. 5.9b. Again, the inspection process does not yield an orientation relation since 'Reno' is not yet contained in the visual buffer representation.

The spatial knowledge fragment that is retrieved and included next in the activated long-term memory representation is the first one that provides knowledge about Reno:

```
#<SKF: #<GEO_ENTITY: Reno> in #<GEO_ENTITY: Nevada>
(TOPOLOGICAL-RELATION, res 1)>
```

Other than in the reasoning steps above, Nevada is now modeled as an extended entity to fit the containment relationship with Reno. The resulting visual buffer representation is shown in Fig. 5.10. Now that both entities under consideration (Reno and San Diego) are represented in the visual buffer, the inspection yields a cardinal direction as result. Reno is visualized straight to the right of San Diego in the visual buffer. Therefore, inspection yields the information that Reno is east of San Diego. This result is returned at every degree of resolution that is used for the inspection:

```
#<SKF: #<GEO_ENTITY: Reno> E #<GEO_ENTITY: San Diego>
(CARDINAL-DIRECTION, res 3)>
```

Observe that in contrast to the visual buffer representations above that do not contain 'Reno' (Fig. 5.9), the image mapping process can now focus on the two entities of 'Reno' and 'Nevada'. As a consequence, the extended entities 'California' and 'Nevada' are no longer entirely contained in the visual buffer.

The spatial knowledge fragment retrieved next states that Nevada is north-east of California:

```
#<SKF: #<GEO_ENTITY: Nevada> NE #<GEO_ENTITY: California>
(CARDINAL-DIRECTION, res 3)>
```

Therefore, it replaces the 'east' relation between 'Nevada' and 'California' in the activated long-term memory representation. The resulting visual buffer representation is shown in Fig. 5.11a.

The next two spatial knowledge fragments that are retrieved determine the spatial orientations of San Diego within California, and of Reno within Nevada:

```
#<SKF: #<GEO_ENTITY: San Diego> S #<GEO_ENTITY: California>
(CARDINAL-DIRECTION, res 2)>
```

```
#<SKF: #<GEO_ENTITY: Reno> W #<GEO_ENTITY: Nevada>
(CARDINAL-DIRECTION, res 2)>
```

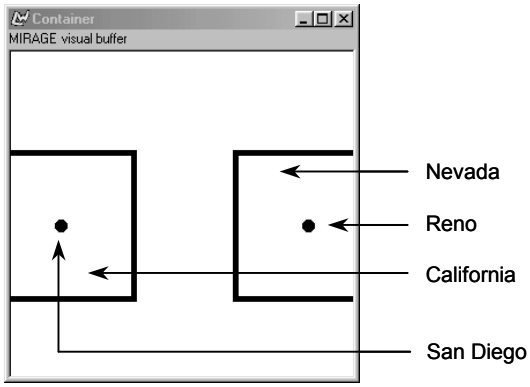

Fig. 5.10. With the third spatial knowledge fragment retrieved from long-term memory, knowledge about all four relevant entities is available. California and Nevada are now modeled as extended entities

Next, Nevada is found to be tangent to California:

```
#<SKF: #<GEO_ENTITY: Nevada> tangent #<GEO_ENTITY:
California> (TOPOLOGICAL-RELATION, res 2)>
```

The visual buffer representations that result from image constructions based on the last three knowledge fragments are depicted in Fig. 5.11b through d. Since all topological relations and all cardinal directions between the object pairs 'Reno – Nevada', 'San Diego – California', and 'Nevada – California' are explicitly retrieved from long-term memory, there is no default knowledge added by the relational completion process. The shapes of the extended entities, however, are still the same default shapes. The results of the inspection of the visual mental image representation depicted in Fig. 5.11d vary between north (at resolution 2) and north-north-east (at resolution 4) depending on the degree of resultion used in the image inspection:

```
#<SKF: #<GEO_ENTITY: Reno> N #<GEO_ENTITY: San Diego>
(CARDINAL-DIRECTION, res 2)>

#<SKF: #<GEO_ENTITY: Reno> NE #<GEO_ENTITY: San Diego>
(CARDINAL-DIRECTION, res 3)>

#<SKF: #<GEO_ENTITY: Reno> NNE #<GEO_ENTITY: San Diego>
(CARDINAL-DIRECTION, res 4)>
```

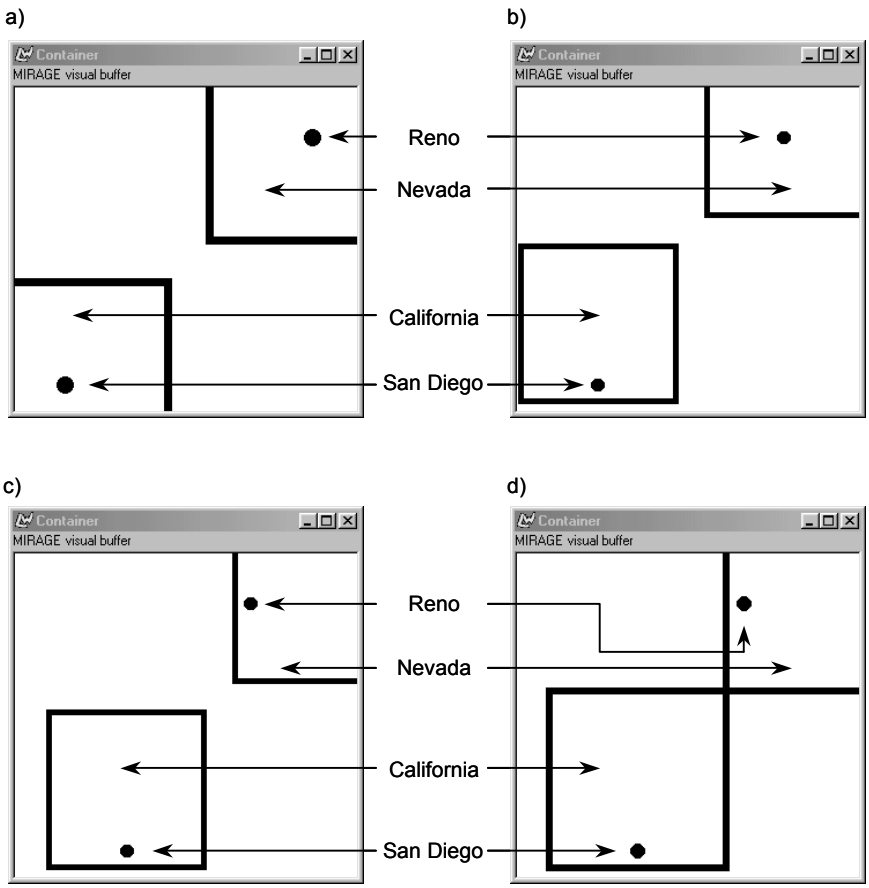

Fig. 5.11. The visual buffer representations resulting from the fragments that state that Nevada is north-east of California (a), that San Diego is located south in California (b), that Reno is west in Nevada (c), and that Nevada and California are tangent (d)

The next spatial knowledge fragment gives an explicit shape for California:

```
#<SKF: SHAPE of #<GEO_ENTITY: California> is
((POINT -0.50 19.50) (POINT 6.00 19.50) (POINT 6.00 13.00)
(POINT 15.00 6.00) (POINT 16.00 1.00) (POINT 11.00 0.00)
(POINT 6.00 3.50) (POINT 0.00 13.00) (POINT -0.50 19.50))
(resolution 1)>
```

Consequently, no default shape has to be assigned in the conversion process. For Nevada, however, there is still the square default shape used. The resulting visual buffer representation is depicted in Fig. 5.12a.

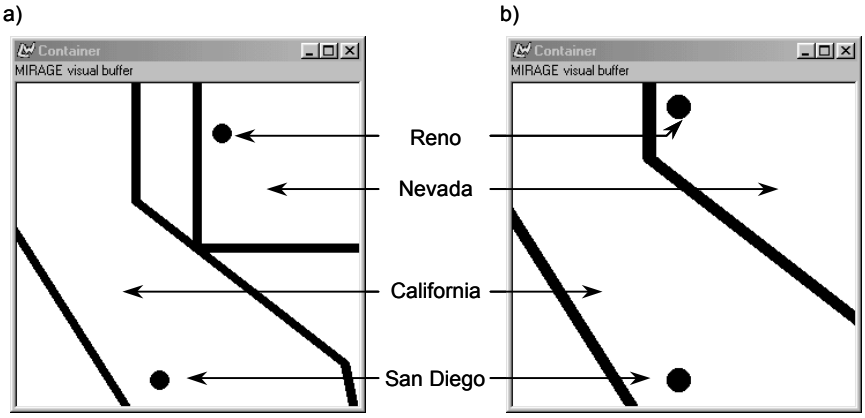

Fig. 5.12. Use of explicit shapes from long-term memory in the mental image construction: (a) Nevada is still represented as default square shape; (b) both California and Nevada are based on their explicit shape representations from long-term memory

The last fragment that is available from long-term memory to answer the spatial question at hand is the explicit shape information about Nevada:

```
#<SKF: SHAPE of #<GEO_ENTITY: Nevada> is
((POINT 6.00 19.50) (POINT 15.00 19.50) (POINT 16.00 8.00)
(POINT 15.00 6.00) (POINT 6.00 13.00) (POINT 6.00 19.50))
(resolution 1)>
```

Now, both extended entities are represented with their proper shapes in the visual buffer. The inspection process yields the result that Reno is north of San Diego at all resolutions.

6 Conclusion and Outlook

This chapter presents the conclusion of the work described in this thesis and provides an outlook on future investigations. First, I will give a summary of this work that features the characteristics of the MIRAGE model. Second, the results of this work are reflected with respect to the theses, the research questions, and the approach described in Chapter 1; I will relate the work to the research reported in Chapter 2. Third, I will point to open issues to be elaborated in future work in AI and in cognitive psychology. The chapter ends with an outlook on possible applications of the work described here.

6.1 Summary

In this work, I developed the MIRAGE model that describes the construction of spatial configurations about large scale space in visual mental images. More specifically, the mental construction is based on the spatial aspects of topological relations, of orientation relations (cardinal directions), and of shape information. It involves both point entities and extended regions. The knowledge available from memory is assumed to be underdetermined with respect to the question to be answered through the configuration constructed in the mind. As MIRAGE models the construction of visual mental images, the constructed representation is in a spatio-analogical representation format. It is evaluated by an inspection process.

MIRAGE is subdivided into a number of sub-processes and representation structures that reflect theoretical aspects of human memory known from cognitive psychology. Processing starts with a hierarchical knowledge representation in long-term memory that is formed by spatial knowledge fragments. Spatial knowledge fragments are the elementary representation structures for spatial facts that hold between geographic entities. A spatial knowledge fragment is characterized by the type of spatial knowledge it encodes and by the degree of resolution of the represented spatial relation. These two characteristics induce a twofold hierarchical representation structure in long-term memory.

This hierarchical representation structure is accessed in a sequential manner. The order in which spatial information is retrieved from memory is controlled by the type of spatial knowledge and by the degree of resolution. The spatial knowledge fragments retrieved from long-term memory are used to construct the activated long-term memory representation. The activated long-term memory representation is a representation structure in working memory that holds the pieces of knowledge used to construct the mental image of a spatial configuration. With respect to the question to be answered, the activated long-term memory representation is underdetermined.

The image construction proper is based on the activated long-term memory representation and is performed in two phases. First, the activated long-term memory representation is completed in a qualitative manner using default knowledge: the ontological types of the involved entities are determined, and missing spatial relationships are assigned. Second, the geographic entities are specified metrically, and they are localized in the visual buffer. The visual buffer is a quasi-pictorial representation structure that is used by the inspection process to obtain the required result. The resulting spatial relation between a pair of entities is obtained in the form of a spatial knowledge fragment.

The image construction is performed in a step-by-step manner. Spatial knowledge fragments are successively integrated into the activated long-term memory representation. The three subsystems of long-term memory activation, image construction, and image inspection operate in parallel. Thus, an image can already be evaluated at an early processing stage, while the working memory representation is further refined. In Chapter 3, I demonstrated how the systems operate in a reasoning problem using the 'Reno – San Diego' scenario.

In Chapter 4, I focused on the image construction subsystem, i.e. the sub-processes that lead from the activated long-term memory representation to the mental image in the visual buffer. Using the more demanding 'Geneva – Nice' scenario, it has been demonstrated that the construction of a visual mental image in MIRAGE cannot always be performed in as straightforward a manner as described in the 'Reno – San Diego' scenario. The spatial knowledge fragments in the activated long-term memory representation together with the default knowledge employed in the image construction can lead to inconsistencies that only permit partial solutions; not all pieces of knowledge can be integrated.

I have argued that in cognitive systems it is usually not feasible to solve a spatial constraint satisfaction problem by directly computing its solutions. Rather, due to resource restrictions in the human mind, it is necessary to reduce a problem's complexity by eliminating degrees of freedom in the problem space at early stages of the problem solving process. However, when a solution to a problem cannot be found in a straightforward way, alternative options must be considered. I analyzed the consequences for the image construction and extended the MIRAGE model. I described a number of image revision strategies that utilize the degrees of freedom in the image construction process. These image revision strategies allow for more flexible image constructions that can be employed when the straightforward image construction fails.

In the model I identified and discussed three options for more elaborate image construction strategies:

1. the reduction of spatial constraints represented in the enriched representation,
2. the modification of the relational completion that leads to the enriched representation, and
3. variations in the interpretation of the relations represented in the enriched representation during the visualization process.

Accordingly, three types of image revision strategies have been included in MIRAGE.

In the *omission of facts* strategy the number of spatial relations to be included in the visual mental image is reduced. This can either be done by eliminating arbitrary pieces of information: in this case the image construction proceeds until a conflict is detected; then it stops. Or spatial relations are excluded from the activated long-term memory representation in an informed manner by the activated long-term memory reduction system: in this case the spatial knowledge fragments in the activated long-term memory representation are reconsidered; the represented relations are assessed with respect to their reliability, and the least reliable pieces of information are excluded.

In the *revision of relational completion* strategy the spatial relations included in the relational completion process either are modified, or one or more spatial relations are left unspecified. In the first case, the representation of the completion parameters that controls the relational completion process is modified; this results in a varied set of default knowledge employed. In the second case, the completion of the missing relations is left to the interpretation of the image specification process.

In the *revision of image specification* strategy the interpretation of qualitative spatial relations by the visualization process is relaxed. If spatial relations are left unspecified by the relational completion process, they are constructed in the image specification step.

In Chapter 5, I described the prototypical implementation of MIRAGE. The implementation was done in Common Lisp. The implementation of MIRAGE is based on SIMSIS, a system for the construction and interpretation of map-like representations. SIMSIS has been developed to model processing spatial knowledge in aspect maps. Aspect maps are a conception for describing interpretation strategies that can be applied to map-like spatial knowledge representations to obtain spatial information. The two basic forms of representations in SIMSIS are *pictures*, which represent the pictorial information of a map, and *scenarios*, which are used to hold the spatial facts that either result from evaluating a picture or that are represented and processed prior to their representation in a picture data structure. The central concept for extracting information of a representation structure in SIMSIS is the interpretation. Any information transfer between two forms of representation is conceived as an interpretation.

The MIRAGE representation structures and processes are designed on the basis of SIMSIS. Spatial knowledge fragments are modeled based on SIMSIS facts. The long-term memory representation structures of hierarchical long-term memory, activated long-term memory, and enriched representation are modeled as SIMSIS

scenarios. The visual buffer is based on a SIMSIS picture data structure for the pictorial representation content. The processes operating on these structures are based on SIMSIS interpretations.

The implemented model shows that MIRAGE yields a visual buffer representation at every refinement step. As soon as the two entities to be related with each other are represented in the visual buffer, the image representation can be inspected by the image inspection process. The result of the image inspection depends on how many knowledge fragments have been retrieved from long-term memory (i.e., the degree of refinement of the image representation) and on the degree of resolution at which the inspection is performed on the visual buffer.

6.2 Results and Discussion

Before discussing the results of this work, I need to make a remark concerning MIRAGE. As a cognitive model, MIRAGE is a computational modeling *framework* rather then the embodiment of a specific psychological theory. In Chapter 3 I provided the metadescription of the model in the sense of Kosslyn (1980). This metadescription determines the cognitive principles of the architecture and its characteristics with respect to its functionality (cf. Section 1.4.2).

I outlined the model, identified the subsystems and characterized their interaction with each other. Section 3.3 and Section 3.4 provided a specification of the entities used in MIRAGE and the subsystems constituting the model, respectively. I demonstrated the behavior of model in the simple 'Reno – San Diego' scenario. Chapter 4 further specified the image construction subsystem and extended the corresponding structures and processes. However, the subject matter MIRAGE deals with is very comprehensive. Many aspects of the model needed to be dealt with in an exemplary manner; many extensions are conceivable, and in partial aspects many design decisions only provide possible modeling solutions within the overall system. Therefore, I will now address the questions,

1. what has been reached with the model,
2. which aspects of the model allow for variations in design, and
3. which aspects require extensions of the model in future work.

In Section 6.2.1 I will discuss how the theses claimed in Chapter 1 have been realized in the model, and I will relate them to the cognitive and AI principles reported in Chapter 2. In Section 6.2.2 I will describe the parameters of the model, i.e. how the operation of the model depends on the data and how varying modes of operation can be described in MIRAGE. Section 6.2.3 concludes this work with respect to how the approach taken in this thesis has been embodied by the model and how the results relate to the research questions and goals of this work.

6.2.1 Reflecting the Theses

MIRAGE is based on four core theses (cf. Section 1.2):

1. geographic knowledge representations are constructed in the mind on demand;

2. geographic knowledge available for the construction is underdetermined;
3. geographic knowledge is fragmented and hierarchically organized in long-term memory; and
4. the construction in the mind is performed in visual mental images.

6.2.1.1 Spatial Knowledge Construction

The central aspect of MIRAGE is the step-by-step construction of knowledge representations. This construction leads to a realization of an image representation in the visual buffer. The several stages during this construction process reflect the structure of human memory as conceived in cognitive psychology. The whole construction is based on elementary pieces of knowledge stored in long-term memory (cf. Section 2.2.2). The knowledge construction itself is performed in working memory, more specifically in the visuo-spatial part of working memory (cf. Section 2.2.1). The access of the long-term memory knowledge is described as an activation process. Consequently, working memory consists both of the long-term memory contents that have been activated and the short-term memory representation that holds the image proper and that has been constructed in the visual buffer.

The complete construction is performed on demand depending on the spatial information required. This construction characteristic is related to the theory of spatial mental models (cf. Section 2.1.3). The actual content of working memory is not based on a spatial state of affairs that has been perceived and stored. Rather, it is the result of how a spatial state of affairs is envisaged by a human in a given situation.

From an AI point of view, MIRAGE is a *hybrid* and an *integrated* diagrammatic reasoning (DR) architecture (cf. Section 2.5.1 and Section 2.5.2, respectively):

1. It is a hybrid DR system as it involves both propositional and pictorial pieces of knowledge for the construction task. The elementary representations of topological relations and cardinal directions are given in a propositional form, whereas the information about the shapes of extended entities is represented pictorially. Both representational forms are stored in spatial knowledge fragments. The visual buffer representation of the image components is completely pictorial, whereas the result of the image inspection again is a spatial knowledge fragment.
2. MIRAGE is an integrated DR system as it relies both on relational and on positional knowledge. The representation of the image components in the visual buffer is a positional representation, whereas the underlying representations in the activated long-term memory (i.e., the activated long-term memory representation and the enriched representation) are relational representations.

6.2.1.2 Underdeterminacy in Long-Term Memory

The knowledge in long-term memory used for the working memory construction is characterized in MIRAGE as underdetermined or lean knowledge (cf. Section 1.2.2). Many spatial relationships between geographic entities are not represented explicitly in the mind (scarce knowledge). Therefore, required relations have to be

inferred from pieces of knowledge available from long-term memory. Moreover, available pieces of knowledge are often represented qualitatively (coarse knowledge), rather than as precise metric values. Coarse knowledge can be available at different levels of granularity. Different levels of granularity allow for constructing a mental representation at the granularity level suitable for solving a given spatial problem. In MIRAGE, mental constructions are successively refined when pieces of knowledge of a higher resolution are retrieved from long-term memory and included in the construction process.

MIRAGE demonstrates how the construction of a working memory representation can be performed on the basis of underdetermined knowledge. To this end, available pieces of information are completed by default knowledge (cf. Section 3.1.2). Default knowledge is used to construct representations specific enough to answer a given question at a required level of granularity. The use of default knowledge in the construction process results in distortions in the working memory representation. Compared to how a spatial configuration could have been acquired from a secondary knowledge source (say, a geographic map), mental representations tend to exhibit simplifications and schematizations (cf. Section 2.1.2).

6.2.1.3 Fragmentation and Hierarchical Organization

MIRAGE models the processing of fragmented spatial knowledge organized in hierarchical structures. Pieces of spatial knowledge in long-term memory are organized in spatial knowledge fragments in MIRAGE. The model operates on the basis of highly fragmented knowledge. I decided for this form of representation to provide a uniform basis for the use of long-term memory knowledge in MIRAGE. However, in Section 6.3.1.2 I will motivate that partially aggregated knowledge structures are a valid option for extending MIRAGE. The knowledge encoded in spatial knowledge fragments is explicit (or declarative) knowledge that forms the semantic memory in the human mind (cf. Section 2.2.2).

The spatial knowledge fragments form a graph structure in long-term memory. They are organized in a twofold hierarchy, which is given by the degree of resolution of the knowledge fragments and the type of spatial knowledge encoded. The type of spatial knowledge is ordered according to the expressiveness of the spatial knowledge (cf. Section 2.4). Both aspects of the hierarchy are used to control the order of access in long-term memory. Spatial knowledge fragments of lower resolution and less expressive knowledge types are retrieved first (cf. Section 3.4.1.3). The retrieval of knowledge from long-term memory is modeled as a graph search process which is related to spreading activation processes in the human mind (cf. Section 2.2).

Both the knowledge fragmentation and the hierarchical structure is in line with the atlas metaphor and the GIS metaphor for human memory, albeit the atlas metaphor suggests partially coherent representations of spatial knowledge (cf. Section 2.1.2).

6.2.1.4 Visual Mental Imagery

MIRAGE models the construction of spatial configurations in mind in visual mental images. A visual buffer data structure is used to hold the image content. MIRAGE exhibits a number of essential characteristics of mental imagery (cf. 2.3.2): images are constructed on demand from organized pieces of knowledge from long-term memory. Both pictorial and propositional information is involved in the image construction. The use of the constructed image to explore a spatial configuration refers to the implicit encoding principle, which is also crucial for diagrammatic reasoning (cf. Section 2.5). The spatio-analogical visual buffer structure realizes the spatial equivalence principle of mental imagery.

MIRAGE models how both the underlying (propositional) working memory structure and the pictorial representation in the visual buffer are constructed, maintained, and modified. As such, the model has the overall structure of the computational imagery approach by Glasgow and Papadias (1992; cf. Section 2.5.3). However, as MIRAGE is a psychologically motivated approach, it is more closely related to the mental imagery conception by Kosslyn (1994a) in important points (cf. Section 2.3.3.2).

6.2.2 The Parameters of the Model

The purpose of MIRAGE is to describe the mental operations that are performed by a person reasoning about a spatial configuration. The operation of MIRAGE with respect to a required spatial configuration depends on the spatial facts available and on the image construction strategies employed in the construction process. In this section I will give an overview of the parameters that determine the model's performance. These parameters comprise data that are explicitly provided to MIRAGE and those, which are implicitly contained in MIRAGE's partial processes. The first type of parameters can be directly changed to evoke an alternative behavior of the model. The latter type of parameters is either determined by how the processes have been designed, or how the partial processes are controlled in MIRAGE. The variation of the implicit parameters requires modifications in the implementation of the respective processes or modifications in the control structure of the model.

Explicit parameters are

- the representation structure in long-term memory,
- the entities under consideration,
- the spatial relation required,
- the list of completion parameters employed in the relational completion process, and
- the image specification parameters.

Implicit parameters are

- the assessment of the hierarchy in long-term memory for the access process,
- the default shapes of extended entities and the metric values used for the image specification,

- the control of the activated long-term memory reduction, the modification of the completion parameters, and the modification of the image specification parameters, and
- the overall control of the various image construction strategies.

6.2.2.1 Explicit Parameters

Image constructions depend on the pieces of knowledge represented in long-term memory. The knowledge represented in the hierarchical long-term memory is explicitly instantiated in MIRAGE (cf. Section 5.2.1.2). A given hierarchical long-term memory representation models the individual knowledge of a person who performs the reasoning process. As a precondition for employing MIRAGE, the required spatial relation is not explicitly represented in memory. Varying pieces of knowledge available from an individual's memory influence the reasoning process performed: both the operation of the model and the final result of the image construction depend on the facts available from long-term memory and on the degree of resolution they are represented at.

Also the geographic entities a person wants to relate spatially and the type of spatial relation required (together with the intended degree of resolution) are explicitly provided to MIRAGE. Both parameters are needed for the access process and for the image inspection. The access process retrieves suitable spatial knowledge fragments from the hierarchical long-term memory representation. So it determines the pieces of knowledge that are used for the image construction. The image inspection yields the required spatial relation at the intended degree of resolution. Although in principle arbitrary relations between any pairs of entities can be inspected from a constructed mental image (as long as the respective entities are contained in the image), the construction and the inspection of the mental image refer to the same pair of entities and to the same type of spatial relation in MIRAGE.

In the conversion process, the relational completion is based on the set of completion parameters. In Section 3.4.2.2 I described which completion parameters are used as default in MIRAGE. These completion parameters are explicitly encoded in the model and are used when the relational completion process is called. The predefined completion parameters model the default knowledge a person employs when missing spatial relations are complemented in the image construction process. Varying definitions of these parameters reflect varying image construction strategies, either with respect to a person's preferences or with respect to specific situations or reasoning problems.

Also, how image specification is performed when an enriched representation is visualized in the image specification process is explicitly defined by the image specification parameters. These image specification parameters determine how qualitative spatial relations contained in the enriched representation are interpreted in the image specification. In MIRAGE, distinct positions are used for the interpretation of cardinal directions as default. This default can be modified to model other default reasoning strategies in a person. For example, an expert trained in reasoning about spatial configuration may use areal interpretations of cardinal

directions as default to allow for more complex configurations in the image construction process.

6.2.2.2 Implicit Parameters

In Section 3.4.1.3 I described how the hierarchical structure in long-term memory is used to control the access process that extracts spatial knowledge fragments from long-term memory. The way the access process is performed is controlled implicitly. I proposed an assessment strategy both for the type of spatial knowledge fragment and the degree of resolution, and I described which order of retrieval results from this assessment. Both characteristics, however, can be conceived otherwise, which would result in a different retrieval behavior. Whether the proposed retrieval strategy is sensible, whether there are variations between different persons or between different situations, and what extensions are required in the model are questions to be investigated empirically. Modeling varying access strategies in MIRAGE requires a modification of the access process.

In the conversion process, square shapes are assigned to extended entities for which no explicit shape information has been retrieved from memory. It is not yet clear however, what kind of default shape can be assumed in the construction of visual mental images. The use of other shapes or alternative shapes requires an extension of the conversion process in MIRAGE.

The same holds for the metric values used in the image specification process. To construct an image in the visual buffer, the size of default shapes and the distances between entities has to be determined. These parameters are part of the image specification process. As MIRAGE does not explicitly deal with distance information, these parameters cannot be modified in the model.

In Chapter 4, the alternative image construction strategies have been presented that can be employed in MIRAGE when a straightforward image construction fails: the reduction of spatial knowledge fragments in the activated long-term memory representation, the modification of the completion parameters, and the modification of the image specification parameters. However, the overall control of how these image construction strategies are applied is not determined in the MIRAGE model. Rather, the control of these strategies is specified by varying sequences of function calls to employ the different strategies in MIRAGE. Thus, the overall behavior of the model depends on the choice of a specific image construction strategy and the order in which the processes are called. A general modeling of overall control strategies of image constructions is beyond the scope of MIRAGE.

6.2.3 Conclusions

The goal of the work described in this thesis was to develop a computational model of the construction of geographic knowledge representations in the human mind based on lean knowledge. To this end, the MIRAGE model has been developed. The development of MIRAGE has been performed in three steps:

- First, I worked out a conceptual model for the processing of geographic knowledge in mental images. I designed a metadescription of the model that identifies subsystems involved in the mental image construction process. I described the roles of these subsystems in the image construction task.
- Second, I specified the components of the model with respect to their functional properties and the subsystems with respect to the representation structures that hold the spatial knowledge at the respective processing stages. I developed the processes operating on these knowledge structures, and I showed how these processes can be controlled with respect to each other to obtain the intended model behavior.
- Third, I implemented the basic functionality of the model to demonstrate its overall behavior. I showed how MIRAGE's components can be realized in an implemented computational model and how the model behaves in a reasoning scenario.

In the development of MIRAGE I referred to metaphorical conceptions of geographic knowledge processing. I used results from empirical studies that point to representational and reasoning characteristics of geographic knowledge in the human mind. And I related the work to existing models from AI and from cognitive psychology. Through these methods I designed a model that comprises aspects of human reasoning about large scale spaces from the representation of geographic knowledge in long-term memory, via its use in working memory, up to its final result aimed at by the reasoning process.

I employed the method of experimental computational modeling to provide a unique interpretation of the phenomena involved in human reasoning about geographic space: MIRAGE allows for observing and assessing the conceptions developed in this work.

Another objective of this work was to identify further issues to be worked on with respect to the research question. In the following section I will point to future work that will lead to further iteration cycles in the interdisciplinary research of human spatial cognition.

6.3 Future Work

In this last section I will discuss future work to be done

1. with respect to extensions of the MIRAGE model,
2. with respect to empirical investigations to be performed to clarify a number of psychological aspects touched by MIRAGE, and
3. with respect to application perspectives that can profit from the work described here.

6.3.1 Extending MIRAGE

First, I will from an AI perspective point to a number of issues that extend MIRAGE. These points cover:

1. the types of geographic entities and spatial relations that can be represented and processed in MIRAGE,
2. the operation on partially aggregated knowledge structures, both in long-term memory and in working memory,
3. extensions of the mental imagery functionality, and
4. the explicit realization of implicit parameters in the model.

6.3.1.1 Geographic Entities and Spatial Relations

MIRAGE allows for reasoning with point entities and with extended entities. The ontological type of an entity is decided during the image construction process. However, linear entities also play an important role in reasoning about geographic entities. For instance, a river separates two areas, a path may run alongside a coast line, or a mountain range structures a region. Therefore, it is desirable to include reasoning about linear entities as ontological object type in MIRAGE as well. Together with linear entities as ontological types, new interpretations of spatial relations are also required. For instance, linear entities influence the topological relationships that hold between geographic objects. Moreover, the interpretation of orientation relations with respect to linear entities has to be modeled to be able to construct visual mental images.

Regarding the types of spatial relations it is sensible to include (qualitative) distances in MIRAGE. Qualitative distances directly influence the construction of visual mental images as they provide knowledge that helps determine the relative positions of entities within the mental image. When two or more entities are already contained in the image, it is important to know how the distances between those entities comparatively relate to the tentative position of another entity to be included in the image.

MIRAGE has been developed under the precondition that spatial knowledge available from memory is coarse, or qualitative. However, pieces of spatial knowledge may also be of the precise metric type. For instance, someone might know a precise distance between two locations, or an exact orientation between a pair of entities. In this case, a mental image may be partially based on precise knowledge and partially on coarse knowledge. For this reason, it would be desirable to allow for precise spatial information to be employed in the image construction process, i.e., to combine precise knowledge with coarse pieces of knowledge. For this purpose, it is necessary to map some geographic dimension, say, a distance given in miles to the pictorial space of the visual buffer.

6.3.1.2 Partially Aggregated Knowledge Structures

In Section 3.1.1 I explained why spatial knowledge fragments are used in MIRAGE as the basic form of representing spatial information in long-term memory: I decided that I did not want to deal with knowledge structures of varying complexity. However, it is sensible to assume that there are partially aggregated knowledge structures represented in long-term memory. For instance, someone might have learnt a detail from an external map very well. This would result in a complex representation in memory that comprises several geographic entities. Or

a mental image constructed in a specific situation might have been memorized in long-term memory for later use to be recalled. Therefore, MIRAGE should be extended to allow for dealing with partially aggregated pieces of knowledge. It is important to note that this idea does not conflict with the general idea of image constructions in working memory. Aggregated knowledge structures can be combined in an image construction with other pieces of knowledge, and they even may be decomposed to be used in parts to construct new mental images.

Besides in (non-activated) long-term memory, also in working memory aggregated knowledge structures are sensible. It has been reported in Section 2.3.3.2 that the possible degree of complexity in a mental image depends on chunking strategies applied to pieces of knowledge to reduce the number of distinct pieces of knowledge to be maintained at a time. So also from the viewpoint of the complexity of the representation in activated long-term memory, it is sensible to have complex partial knowledge structures that can be employed in the image construction process to allow for more complex images.

6.3.1.3 Mental Imagery Functionality

Another option for extending MIRAGE lies in the integration of further functional aspects of mental imagery. Especially the following functional properties have not been considered in detail yet:

- the fading out of the visual buffer content: although the image maintenance loop that periodically refreshes the visual buffer is part of MIRAGE, the physical property of fading of the visual buffer content is not modeled. So the model does not cover temporal aspects like, for instance, how much time some intermediate processing may take before image contents are lost.
- size and resolution properties of the visual buffer: by means of the image mapping process, it is modeled that the visual buffer content is focused on the relevant parts of the entities contained. However, the visual buffer has not been restricted with respect to its extent, and with respect to its resolution.[53]
- image modification facilities to reason about alternative spatial configurations when an image has already been successfully constructed: so far, the aim of MIRAGE is to provide a possible image construction on the basis of a given activated long-term memory representation. However, in reasoning contexts it can be sensible to search for further possible representations to check for spatial alternatives (cf. the theory of preferred mental models, Section 4.2).

6.3.1.4 Parameters of MIRAGE

As a last point for possible extensions of MIRAGE, I would like to point to the implicit parameters discussed in Section 6.2.2.2. These implicit parameters are determined by the design decisions taken to enable a specification of the respective processes that allows for a computer implementation. However, it is desirable to

[53] Remember that the resolution of the visual buffer decreases towards the periphery (cf. Section 2.3.3.1).

make them explicit to test for potential alternatives and to gain a more powerful experimentation tool.

6.3.2 Empirical Investigations

MIRAGE points to a number of questions that require further investigation through empirical experiments. The results obtained from these experiments will help provide more profound grounding for the design of the model's sub-components. Important questions to be empirically investigated are:

1. How is missing information compensated for in working memory?
2. How is the overall control of the image construction system performed and how do partial processes communicate with each other?
3. How complex may the mental operations be that are employed in constructing a visual mental image, and how complex may the mental image itself be?
4. How can mental processing capacity be augmented by the technique of chunking?
5. How can visual and propositional reasoning strategies be combined to find a solution to a given problem?

6.3.2.1 Use of Default Knowledge

In MIRAGE, missing information is completed in the conversion and in the visualization process. In the conversion process, missing qualitative spatial relations are compensated in the activated long-term memory representation by default knowledge (cf. Section 3.4.2.2). From a psychological point of view it is an interesting question which qualitative relations are employed as defaults when no explicit knowledge is available from long-term memory. From the theory of preferred mental models (cf. Section 4.2) we know that mental representations are not selected randomly when several alternative constructions are possible. So the objective of this investigation would be to identify preferences in the use of spatial relations.

In the same way, default knowledge is used when the representation in activated long-term memory is interpreted by the image specification process. For instance, to construct an image representation, default shapes of extended entities must be selected, qualitative orientation relations must be manifested by distinct angles, and distances between entities must be fixed. It needs to be investigated how qualitative relations between entities are interpreted and visualized in a mental image construction.

The options of how default knowledge is employed in the mental reasoning task are related to the complex image construction strategies developed in Chapter 4. All strategies deal with the modification of default assumptions during the image construction process. By considering these strategies, the behavior of the model can be compared to human reasoning under experimental conditions. I will

briefly sketch how default knowledge affects the diverse reasoning strategies and how these strategies can be empirically investigated:[54]

- When the number of spatial constraints represented in the working memory representation is reduced to achieve a mental image construction, this is either done randomly or in an informed manner. When constraints are randomly excluded from the image specification, the resulting image only contains pieces of information included prior to the detection of the conflict. Thus, the resulting image depends on the order in which entities and relations have been considered. As demonstrated in the example in Section 4.3.1, spatial reasoning situations can be constructed such that the result of the reasoning process reveals which relations have been used and which have been excluded. This reasoning strategy can be induced in experiments under time pressure. When the image reduction is performed in an informed manner, we can test whether the expected reduction strategy has been (successfully) applied. In the experiment, the assessment of the reliability of spatial facts must be controlled. Experimental results can be interpreted from errors produced in the reasoning task.
- In the modification of relational completion strategy, default relations employed in the relational completion process are exchanged by spatial alternatives. The first question to be empirically investigated is which spatial relations are used as defaults. Preferred relational completions can be revealed by distortions in the resulting images, when reasoning conditions are appropriately designed: alternative defaults must lead to varying valid solutions that can be contrasted. After having detected which relations are used as defaults, experiments can be designed that contrast reasoning tasks that can be performed on the basis of default relations with reasoning tasks in which defaults have to be modified to obtain a solution. Reaction times and error characteristics in the results can refer to the relational completion that has been chosen.
- The investigation of the variations in the interpretation of qualitative relations in the image specification process seems to be most demanding. However, it is possible to construct reasoning situations in which complex and variable interpretations are required to reach a visual mental image construction (cf. the example depicted in Fig. 4.13). In the experiment it can be checked whether complicated solutions are found. To track the employed mental reasoning strategies, methods like sketch drawing or verbal protocols can be used.

6.3.2.2 Control of Image Construction

In MIRAGE, complex image construction strategies are available when a straightforward image construction is impossible. However, it is not clear which of these strategies is applied in a given situation, in which order alternative image con-

[54] In empirical investigations the pieces of knowledge the reasoning task is based on must be carefully controlled. To control the preconditions of the mental reasoning task, geographic information can be provided to the participants in the form of fictitious maps (e.g., Kosslyn et al., 1978; Stevens & Coupe, 1978) or in the form of verbal information.

struction strategies are used, and how the different strategies are combined. Moreover, when image strategies are used that consist of partial processes that depend on each other, it is not yet clear how these strategies are coordinated (for instance, is there a superordinate controlling instance or do processes autonomously communicate with each other?).

From a psychological point of view the influencing factors for the selection of the strategy can be investigated. How is a given image construction situation assessed with respect to the knowledge involved? For instance, when the reliability of spatial knowledge fragments has to be assessed in the image construction process, this assessment may be based on the form of representation in long-term memory, on the reasoning processes used to compute intermediate results, on simplification strategies employed during knowledge processing, etc. Further empirical results can help to design a more adequate model of the interacting subtasks in the image construction process.

6.3.2.3 Processing Capacity for Mental Images

The image construction strategies developed in Chapter 4 can become rather complex. Especially, when spatial relationships are left unspecified in the relational completion (cf. Section 4.4.3.2) or when orientation relations are interpreted as sectors (cf. Section 4.4.4.1) to allow for more flexible image construction options in the image specification process, the image construction becomes more and more complex. This is particularly true, when image construction strategies are combined in the attempt to find a suitable visual mental image: the amount of information to be simultaneously dealt with in working memory increases.

The processing capacity of human working memory (specifically: the number of entities that can be dealt with at a time) is restricted to a small number of items. This number has been estimated at about seven (Miller, 1956) or at about four (Cowan, 2001). Thus, an important issue to be investigated is how the limitations in the working memory capacity restrict the options of complex image construction strategies (cf. Just et al., 1996). As the mental image construction process in MIRAGE is done by a number of sub-processes, an interesting question is, how the distribution of the image construction task on these sub-processes relates to the possible complexity of the image construction process. It would be interesting to investigate which of the partial processes provide limiting factors for the complexity of the image construction process.

Besides the restrictions in the processing capacity of working memory for *constructing* a visual mental image, also the complexity of the image itself is an issue to be investigated. The image complexity is influenced by two limiting characteristics: first, by the mere size of the visual buffer, and second, by the amount of information that can be simultaneously maintained. It is known that the visual buffer is restricted both in extension and in resolution. Therefore, the amount of spatial information that can be mapped to the visual buffer is restricted. On the other hand, images need to be periodically refreshed to prevent them from fading. Therefore, the number of items that can be held in the visual buffer is also a time-critical factor.

6.3.2.4 Use of Chunking Facilities

It is known that more complex structures can be dealt with in working memory through the technique of chunking (cf. Section 2.3.3.2): two or more entities are combined into a chunk, which then is treated as a single entity in the mind. From a psychological point of view it is interesting to investigate how chunks are formed in the image construction process, and how chunking can increase the complexity of the image construction process.

This issue is related to modeling partially aggregated knowledge structures in MIRAGE (cf. Section 6.3.1.2). Chunks can be formed when a partial image has been formed. For example, when two or more geographic entities have been successfully combined in a mental image, they may be aggregated to focus on another problem. The spatial relationship of this chunk with respect to further spatial entities can be considered in the image construction. Insights about how chunks can be formed (and, if necessary, decomposed again) would provide interesting options for extending the MIRAGE conception.

6.3.2.5 Combination of Propositional and Image-Based Reasoning

In MIRAGE, a solution to a given spatial problem is found by constructing a visual mental image. However, reasoning strategies that are not based on mental images but are performed on propositional representation structures also can be used in spatial problem solving. For instance, an inference about spatial orientation can be purely based on transitivity properties: if a location A is west of a location B, and another location C is west of A, it may be directly concluded on the basis of the transitivity of the 'west of' relations that location C is also west of location B. Also non-visual spatio-analogical representation structures have been investigated in cognitive psychology (Knauff & Johnson-Laird, 2000; Knauff et al., 2001). These investigations showed that constructing visual mental images for spatial reasoning processes can even impair the reasoning process.

It is interesting to investigate in which situations visual mental images are used, in which cases non-visual reasoning strategies are used, and whether also hybrid reasoning strategies are employed (i.e., strategies that involve combinations of image-based and propositional strategies).

6.3.3 Application Perspectives

The research reported in this thesis and the development of MIRAGE contribute to interdisciplinary basic research in mental processing of large scale spatial knowledge. A number of application perspectives can be identified that can profit from the work reported in this thesis. These perspectives refer to:

1. the cognitively adequate presentation of visuo-spatial information and
2. the support of visual thinking in mental images by external representations.

6.3.3.1 Adequate Presentation of Visual Information

Diagrammatic representations are an important means of conveying information. This information can be inherently spatial information, or it can be non-spatial in-

formation that has been spatially encoded. For example, to visualize structural characteristics of a subject-matter, a diagrammatic representation can be produced. In both cases, the diagrammatic form of representation is chosen to ease comprehension, memorizing, and processing of the given information.

From this point of view a profound understanding of human spatial reasoning capabilities can help to decide for adequate forms of visuo-spatial representations. A computational model of human spatial reasoning with visual mental images can be adopted for a better conception of how spatial information is processed in a human user under application perspectives. Possible applications are in the user interface of geographic information systems (GISs), in assistance systems that support spatial tasks (like wayfinding, e.g. Casakin et al., 2000), or in tutorial systems that convey spatial information.

An important task for GIS research and development is to provide user interfaces that are compatible with the representation and processing of geographic information in humans (Taylor & Tversky, 1995; cf. Section 1.3.2). An implemented model of human large scale spatial reasoning may be used to test alternatives in representing spatial information with respect to their cognitive adequacy. From a cognitive perspective it is sensible to process information in a GIS like humans think about large scale spaces, for example, when a common sense conception of a geographic setting is to be processed. The conception of large scale spaces in human reasoning may significantly deviate from reality (cf. the ideas of naive geography, Egenhofer & Mark, 1995b).

Another aspect can be the interactive presentation of geographic information provided by a GIS. A user can be assisted interactively by a system that helps her visualize a spatial state of affairs similar to processing in visual mental images: the visualization task is driven by how the human mind would tend to envisage a spatial configuration. This may help improve the understanding of complex spatial phenomena. In addition, a system that embodies a model for human reasoning about large scale spaces can help avoid typical mistakes in a mental construction and can point to alternatives not preferably considered by the human user.

6.3.3.2 External Support of Reasoning in Mental Images

It is known that the spatial reasoning capacity based on visual mental images is restricted by the processing restrictions in human working memory. However, many tasks performed by human experts rely on these mental imagery facilities. For example, spatial configuration problems in design and architecture, or planning tasks in urban and regional planning are based on envisaging spatial configurations. Frequently is is desirable to assist humans in carrying out spatial reasoning tasks and to help them overcome the restrictions set by the capacity limitations of the human working memory.

It has been reported that (internal) visual mental imagery and the visual perception of the external world are widely based on the same systems in the human mind (cf. Section 2.3). Therefore, it is sensible to include external visuo-spatial representations into the mental reasoning process about spatial configurations. Similar to the way designers and architects use pencil sketches on paper to en-

hance their imagination (cf. Suwa et al., 1999; Goldschmidt, 1999), a software system might be used to augment the human visuo-spatial reasoning capacities. An implemented model of human spatial reasoning with mental images can provide the basis for a spatial task assistance system.

The main advantage of such a system would be that external representations (for example, on an interactive computer screen) can be manipulated by the human designer and by the system. The human can manipulate the external representation according to operations performed in his mind, whereas the assistance system can operate on its representation according to its underlying model of the human mental imagery functionality.

This application of a model for human mental image processing is related to the theory of visual languages (Marriott & Meyer, 1998). The issue of visual language theory is to provide a formal foundation for the use of visual notations (e.g., Haarslev, 1998), i.e., to provide a means to mediate between a diagrammatic representation and a corresponding formal specification, for example for the use in spatial queries (e.g., Egenhofer, 1996; Wessel & Haarslev, 1998).

A crucial aspect of a system for the external support of spatial reasoning in visual mental images would be the development of a suitable interface between the assisting computer system and the human user: the information exchange between the interacting partners, the human and the computer must be as smooth as possible in order not to disturb the overall reasoning process. Ideal interaction modalities may be natural language interfaces and/or the interaction with gestures (Hauptmann & McAvinney, 1993; Fröhlich & Wachsmuth, 1998).

References

Alexander, C. (1982). A city is not a tree. In Kaplan, S. & Kaplan, R. (Eds.), Humanscape - Environments for people (pp. 377-402). Ann Arbor, MI: Ulrich's Books Inc.

Allen, J. F. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM, 26* (11), 832-843.

Anderson, J. R. (1978). Arguments concerning representations for mental imagery. *Psychological Review, 85* (4), 249-277.

Anderson, M. (Chair) (1998). *Reasoning with diagrammatic representations - Papers from the 1997 AAAI Fall Symposium* (Technical Report No. FS-97-02). Menlo Park, CA: The American Association for Artificial Intelligence.

Appleyard, D. (1970). Styles and methods of structuring a city. *Environment and Behavior, 2*, 100-118.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation (Vol. 2)*. New York: Academic Press.

Baddeley, A. D. (1986). *Working memory*. New York: Oxford University Press.

Barkowsky, T. (2001). Mental processing of geographic knowledge. In D. R. Montello (Ed.), *Spatial Information Theory - Foundations of Geographic Information Science* (pp. 371-386). Berlin: Springer.

Barkowsky, T., & Freksa, C. (1997). Cognitive requirements on making and interpreting maps. In S. Hirtle & A. Frank (Eds.), *Spatial information theory: A theoretical basis for GIS* (pp. 347-361). Berlin: Springer.

Barkowsky, T., Latecki, L. J., & Richter, K.-F. (2000). Schematizing maps: Simplification of geographic shape by discrete curve evolution. In C. Freksa, W. Brauer, C. Habel, & K. F. Wender (Eds.), *Spatial Cognition II - Integrating abstract theories, empirical studies, formal models, and practical applications* (pp. 41-53). Berlin: Springer.

Barkowsky, T., Röhrig, R., & Freksa, C. (1996). Operationalizing diagrammaticity. In W. Maaß & P. McKevitt (Eds.), *ECAI-96 Workshop on Representations and Processes between Vision and Natural Language* (pp. 27-34). Budapest, Hungary.

Barr, A., & Feigenbaum, E. A. (Eds.) (1981). *The handbook of artificial intelligence, Vol. 1*. Stanford, CA: HeurisTech Press.

Barwise, J., & Etchemendy, J. (1994). *Hyperproof*. New York: Cambridge University Press.

Barwise, J., & Etchemendy, J. (1995). Heterogeneous logic. In J. Glasgow, N. H. Narayanan, & B. Chandrasekaran (Eds.), *Diagrammatic reasoning: Computational and cognitive perspectives* (pp. 211-234). Menlo Park, CA: AAAI Press.

Berendt, B., Barkowsky, T., Freksa, C., & Kelter, S. (1998a). Spatial representation with aspect maps. In C. Freksa, C. Habel, & K. F. Wender (Eds.), *Spatial cognition - An interdisciplinary approach to representing and processing spatial knowledge* (pp. 313-336). Berlin: Springer.

Berendt, B., Rauh, R., & Barkowsky, T. (1998b). Spatial thinking with geographic maps: an empirical study. In H. Czap, H. P. Ohly, & S. Pribbenow (Eds.), *Herausforderungen an die Wissensorganisation: Visualisierung, multimediale Dokumente, Internetstrukturen* (pp. 63-73). Würzburg: ERGON Verlag.

Braitenberg, V. (1984). *Vehicles - Experiments in synthetic psychology*. Cambridge, MA: MIT Press.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretative approach. *Cognitive Psychology, 3*, 193-209.

Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence, 47*, 139-159.

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). *Readings in information visualization: Using vision to think*. San Francisco, CA: Morgan Kaufman.

Carriero, C. M., Futrelle, R. P., Nikolakis, N., & Tselman, M. (1992). *Informational diagrams in scientific documents*. N. H. Narayanan (Ed.). AAAI Spring Symposium on Reasoning with Diagrammatic Representations - Working Notes.

Casakin, H., Barkowsky, T., Klippel, A., & Freksa, C. (2000). Schematic maps as way-finding aids. In C. Freksa, W. Brauer, C. Habel, & K. F. Wender (Eds.), *Spatial Cognition II - Integrating abstract theories, empirical studies, formal models, and practical applications* (pp. 54-71). Berlin: Springer.

Clementini, E., & Di Felice, P. (1995). A comparison of methods for representing topological relationships. *Information Sciences, 3*, 149-178.

Clementini, E., & Di Felice, P. (1997). A global framework for qualitative shape description. *Geoinformatica, 1* (1), 11-27.

Clementini, E., Di Felice, P., & Hernández, D. (1997). Qualitative representation of positional information. *Artificial Intelligence, 95*, 317-356.

Clementini, E., Di Felice, P., & van Oosterom, P. (1993). A small set of formal topological relationships suitable for end-user interaction. In D. Abel, & B. C. Ooi (Eds.), *Advances in spatial databases - Third International Symposium, SSD'93, Singapore* (pp. 277-295). Berlin: Springer.

Cohn, A. G. (1995). A hierarchical representation of qualitative shape based on connection and convexity. In A. U. Frank & W. Kuhn (Eds.), *Spatial information theory - A theoretical basis for GIS* (pp. 311-326). Berlin: Springer.

Cohn, A. G. (1997). Qualitative spatial representation and reasoning techniques. In G. Brewka, C. Habel, & B. Nebel (Eds.), *KI-97: Advances in Artificial Intelligence* (pp. 1-30). Berlin: Springer.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*, 407-428.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8*, 240-247.

Cooper, L. A., & Shepard, R. N. (1973). Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), *Visual information processing* (pp. 76-176). New York: Academic Press.

Couclelis, H. (1992). People manipulate objects (but cultivate fields): Beyond the raster-vector debate in GIS. In A. U. Frank, I. Campari, & U. Formentini (Eds.), *Theories and methods of spatio-temporal reasoning in geographic space* (pp. 65-77). Berlin: Springer.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24* (1), 87-114.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.

Downs, R. M., & Stea, D. (Eds.) (1973). *Image and environment: Cognitive mapping and spatial behavior*. Chicago, IL: Aldine.

Downs, R. M., & Stea, D. (1977). *Maps in minds: reflections on cognitive mapping*. New York: Harper & Row.

Eco, U. (1976). *A theory of semiotics*. Bloomington, IN: Indiana University Press.

Edgar, G. A. (1990). *Measure, topology, and fractal geometry*. New York: Springer.

Egenhofer, M. (1991). Reasoning about binary topological relations. In O. Gunther & H.-J. Schek (Eds.), *Second Symposium on Large Spatial Databases, Zurich, Switzerland* (pp. 143-160). Berlin: Springer.

Egenhofer, M. J. (1996). Spatial-query-by-sketch. *1996 IEEE Symposium on Visual Languages* (pp. 60-67). IEEE Computer Society Press.

Egenhofer, M., & Franzosa, R. (1991). Point-set topological spatial relations. *International Journal of Geographical Information Systems, 5* (2), 161-174.

Egenhofer, M. J., & Mark, D. M. (1995a). Modeling conceptual neighborhoods of topological line-region relations. *International Journal of Geographical Information Systems, 9* (5), 555-565.

Egenhofer, M. J., & Mark, D. M. (1995b). Naive geography. In A. U. Frank & W. Kuhn (Eds.), *Spatial information theory. A theoretical basis for GIS* (pp. 1-15). Berlin: Springer.

Farah, M. J. (1988). Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychological Review, 95*, 307-317.

Finke, R. A. (1989). *Principles of mental imagery*. Cambridge, MA: MIT-Press.

Finke, R. A. (1990). *Creative imagery: Discoveries and inventions in visualization*. Hillsdale, NJ: Lawrence Erlbaum.

Finke, R. A. (1992). *Creative inference in imagery and invention*. AAAI Spring Symposium on Reasoning with Diagrammatic Representations, Stanford University, California.

Frank, A. U. (1991). Qualitative spatial reasoning with cardinal directions. *Proceedings of the Seventh Austrian Conference on Artificial Intelligence, Vienna* (pp. 157-167). Berlin: Springer.

Frank, A. U. (1992). Qualitative spatial reasoning about distances and directions in geographic space. *Journal of Visual Languages and Computing, 3*, 343-371.

Frank, A. U. (1996). Qualitative spatial reasoning: Cardinal directions as an example. *IJGIS, 10* (3), 269-290.

Frank, A. U., & Raubal, M. (1999). Formal specification of image schemata - a step towards interoperability in geographic information systems. *Spatial Cognition and Computation, 1*, 67-101.

Freksa, C. (1988). Intrinsische vs. extrinsische Repräsentation zum Aufgabenlösen oder die Verwandlung von Wasser in Wein. In G. Heyer, J. Krems & G. Görz (Eds.), *Wissensarten und ihre Darstellung* (pp. 155-165). Berlin: Springer.

Freksa, C. (1991). Qualitative spatial reasoning. In D. M. Mark & A. U. Frank (Eds.), *Cognitive and linguistic aspects of geographic space* (pp. 361-372). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Freksa, C. (1992a). Temporal reasoning based on semi-intervals. *Artificial Intelligence, 54* (1-2), 199-227.

Freksa, C. (1992b). Using orientation information for qualitative spatial reasoning. In A. U. Frank, I. Campari, & U. Formentini (Eds.), *Theories and methods of spatio-temporal reasoning in geographic space* (pp. 162-178). Berlin: Springer.

Freksa, C., & Barkowsky, T. (1996). On the relation between spatial concepts and geographic objects. In P. Burrough & A. Frank (Eds.), *Geographic objects with indeterminate boundaries* (pp. 109-121). London: Taylor & Francis.

Freksa, C., & Barkowsky, T. (1999). On the duality and on the integration of propositional and spatial representations. In G. Rickheit & C. Habel (Eds.), *Mental models in discourse processing and reasoning* (pp. 195-212). Amsterdam: Elsevier.

Freksa, C., Barkowsky, T., & Klippel, A. (1999). Spatial symbol systems and spatial cognition: A computer science perspective on perception-based symbol processing. *Behavioral and Brain Sciences, 22* (4), 616-617.

Freksa, C., Moratz, R., & Barkowsky, T. (2000a). Schematic maps for robot navigation. In C. Freksa, W. Brauer, C. Habel, & K. F. Wender (Eds.), *Spatial Cognition II – Integrating abstract theories, empirical studies, formal models, and practical applications* (pp. 100-114). Berlin: Springer.

Freksa, C., Moratz, R., & Barkowsky, T. (2000b). Robot navigation with schematic maps. In E. Pagello, F. Groen, T. Arai, R. Dillmann, & A. Stentz (Eds.), *Intelligent Autonomous Systems 6 (IAS-6)* (pp. 809-816). Amsterdam: IOS Press.

Freksa, C., & Röhrig, R. (1993). Dimensions of qualitative spatial reasoning. In N. P. Carreté & M. G. Singh (Eds.), *Qualitative reasoning and decision technologies, Proc. QUARDET'93* (pp. 483-492). Barcelona.

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 126-132.

Friedman, A., & Brown, N. R. (2000). Reasoning about geography. *Journal of Experimental Psychology: General, 129* (2), 193-219.

Fröhlich, M., & Wachsmuth, I. (1998). Gesture recognition of the upper limbs - From signal to sybol. In I. Wachsmuth & M. Fröhlich (Eds.), *Gesture and sign language in human-computer interaction* (pp. 173-184). Berlin: Springer.

Forbus, K. D., Nielsen, P., & Faltings, B. (1991). Qualitative spatial reasoning: the CLOCK project. *AI, 51* (1-3), 417-471.

Funt, B. (1980). Problem-solving with diagrammatic representations. *Artificial Intelligence, 13*, 201-230.

Furnas, G. W. (1992). Reasoning with diagrams only. In N. H. Narayanan (Ed.), *AAAI Spring Symposium on Reasoning with Diagrammatic Representations*. Stanford, CA: AAAI.

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (1998). *Cognitive neuroscience - The biology of the mind*. New York: Norton.

Gentner, D., & Stevens, A. L. (Eds.) (1983). *Mental models*. Hillsdale, NJ: Lawrence Erlbaum.

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 5*, 351-360.

Glasgow, J., Narayanan, H., & Chandrasekaran, B. (Eds.) (1995). *Diagrammatic reasoning: Computational and cognitive perspectives*. Cambridge, MA: MIT-Press.

Glasgow, J., & Papadias, D. (1992). Computational imagery. *Cognitive Science, 16*, 355-394.

Glinert, E. P. (Ed.) (1990). *Visual programming environments*. Washington, D.C.: IEEE Computer Society Press.

Gogel, W. C. (1978). The adjacency principle in visual perception. *Scientific American, 238*, 126-139.

Goldschmidt, G. (1999). The backtalk of self-generated sketches. In J. S. Gero & B. Tversky (Eds.), *Visual and spatial reasoning in design* (pp. 163-184). University of Sydney, Australia: Key Centre of Design Computing and Cognition.

Golledge, R. G. (1992). Do people understand spatial concepts: the case of first-order primitives. In A. U. Frank, I. Campari, & U. Formentini (Eds.), *Theories and methods of spatio-temporal reasoning in geographic space* (pp. 1-21). Berlin: Springer.

Gould, P., & White, R. (1986). *Mental maps* (2nd ed.). Winchester, MA: Allen & Unwin.

Grigni, M., Papadias, D., & Papadimitriou, C. (1995). Topological inference. *Proceeddings of the 14th Int. Joint Conf. of Artificial Intelligence (IJCAI)* (pp. 901-907). Montreal, Canada: AAAI Press.

Haarslev, V. (1998). A fully formalized theory for describing visual notations. In K. Marriott & B. Meyer (Eds.), *Visual language theory* (pp. 261-292). New York: Springer.

Habel, C. (1990). Propositional and depictorial representations of spatial knowledge: the case of path concepts. In R. Studer (Ed.), *Natural language and logic* (pp. 94-117). Berlin: Springer.

Habel, C., Pribbenow, S., & Simmons, G. (1995). Partonomies and depictions: a hybrid approach. In J. Glasgow, N. H. Narayanan, & B. Chandrasekaran (Eds.), *Diagrammatic reasoning: computational and cognitive perspectives* (pp. 625-651). Cambridge, MA: MIT-Press.

Hart, R. A., & Moore, G. T. (1973). The development of spatial cognition: A review. In R. M. Downs & K. E. Boulding (Eds.), *Image and evironment: Cognitive mapping and spatial behavior* (pp. 246-288). Chicago, IL: Aldine.

Hauptmann, A. G., & McAvinney, P. (1993). Gestures with speech for graphic manipulation. *Int. Journal of Man-Machine Studies, 38*, 231-249.

Hayes, P. (1978). The naive physics manifesto. In D. Michie (Ed.), *Expert systems in the microelectronic age* (pp. 242-270). Edinburgh, Scotland: Edinburgh University Press.

Hayes, P. (1985). The second naive physics manifesto. In J. Hobbs & R. Moore (Eds.), *Formal theories of the commonsense world* (pp. 1-36). Norwood, NJ: Ablex.

Head, C. G. (1991). Mapping as language or semiotic system: Review and comment. In D. M. Mark & A. U. Frank (Eds.), *Cognitive and linguistic aspects of geographic space* (pp. 237-262). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Hegarty, M. (1992). Mental animation: inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18* (5), 1084-1102.

Hegarty, M. (2000). Capacity limits in diagrammatic reasoning. In M. Anderson, P., Cheng., & V. Haarslev (Eds.), *Theory and application of diagrams* (pp. 194-206). Berlin: Springer.

Hernández, D. (1994). *Qualitative representation of spatial knowledge*. Berlin: Springer.

Hernández, D., Clementini, E., & Di Felice, P. (1995). Qualitative distances. In A. U. Frank & W. Kuhn (Eds.), *Spatial information theory - A theoretical basis for GIS* (pp. 45-57). Berlin: Springer.

Hirtle, S. C. (1998). The cognitive atlas: using GIS as a metaphor for memory. In M. Egenhofer & R. Golledge (Eds.), *Spatial and temporal reasoning in geographic information systems* (pp. 267-276). Oxford University Press.

Hirtle, S. C., & Heidorn, P. B. (1993). The structure of cognitive maps: Representations and processes. In T. Gärling & R. G. Golledge (Eds.), *Behavior and environment: Psychological and geographical approaches* (pp. 170-192). Amsterdam: North-Holland.

Hirtle, S. C., & Jonides J. (1985). Evidence of hierarchies in cognitive maps. *Memory & Cognition, 13* (3), 208-217.

Holyoak, K. J., & Mah, W. A. (1982). Cognitive reference points in judgements of symbolic magnitude. *Cognitive Psychology, 14*, 328-352.

Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review, 98* (3), 352-376.

Intraub, H., Gottesman, C. V., Willey, E. V., & Zuk, I. J. (1996). Boundary extension for briefly glimpsed photographs: Do common perceptual processes result in unexpected memory distortions? *Journal of Memory and Language, 35*, 118-134.

Intraub, H., & Hoffman, J. E. (1992). Reading and visual memory: Remembering scenes that were never seen. *American Journal of Psychology, 105* (1), 101-114.

Isli, A., Haarslev, V., & Möller, R. (2001). *Combining cardinal direction relations and relative orientation relations in qualitative spatial reasoning* (Technical report No. FBI-HH-M-304/01). University of Hamburg: Department for Informatics.

Isli, A., Museros Cabedo, L., Barkowsky, T., & Moratz, R. (2000). A topological calculus for cartographic entities. In C. Freksa, W. Brauer, C. Habel, & K. F. Wender (Eds.), *Spatial Cognition II - Integrating abstract theories, empirical studies, formal models, and practical applications* (pp. 225-238). Berlin: Springer.

Johnson, M. (1987). *The body in the mind: the bodily basis of meaning, imagination, and reason*. Chicago, IL: University of Chicago Press.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, MA: Harvard University Press.

Johnson-Laird, P. N. (1990). Mental models. In M. I. Posner (Ed.), *Foundations of cognitive science (2nd ed.)* (pp. 469-499). Cambridge, MA: MIT Press.

Jungert, E. (1993). Symbolic spatial reasoning on object shapes for qualitative matching. In A. U. Frank & I. Campari (Eds.), *Spatial information theory - A theoretical basis for GIS* (pp. 444-462). Berlin: Springer.

Just, M. A., Carpenter, P. A., & Hemphill, D. D. (1996). Constraints on processing capacity: Architectural or implementational? In D. M. Steier & T. M. Mitchell (Eds.), *Mind matters - A tribute to Allen Newell* (pp. 141-178). Mahwah, NJ: Lawrence Erlbaum.

Kaplan, S. (1973). Cognitive maps in perception and thought. In R. M. Downs & D. Stea (Eds.), *Image and environment: Cognitive mapping and spatial behavior* (pp. 63-78). Chicago, IL: Aldine.

Kaup, B., Kelter, S., & Habel, C. (1999). Taking the functional aspect of mental models as a starting point for studying discourse comprehension. In G. Rickheit & C. Habel (Eds.), *Mental models in discourse processing and reasoning* (pp. 93-112). Amsterdam: Elsevier.

Khenkhar, M. (1991). Object-oriented representation of depictions on the basis of cell matrices. In O. Herzog & C.-R. Rollinger (Eds.), *Text understanding in LILOG* (pp. 645-656). Berlin: Springer.

Klann, M. (1998). *Ein System zur Entwicklung intelligenter Methoden für raumbasierte Informationssysteme*. Studienarbeit, Fachbereich Informatik, Universität Hamburg.

Knauff, M., & Johnson-Laird P. N. (2000). Visual and spatial representations in relational reasoning. *Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society* (pp. 759-765). Mahwah, NJ: Lawrence Erlbaum.

Knauff, M., Jola, C., & Strube, G. (2001). Spatial reasoning: No need for visual information. In D. R. Montello (Ed.), *Spatial information theory - Foundations of geographic information science* (pp. 447-457). Berlin: Springer.

Knauff, M., Rauh, R., & Schlieder, C. (1995). Preferred mental models in qualitative spatial reasoning: A cognitive assessment of Allen's calculus. *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* (pp. 200-205). Mahwah, NJ: Lawrence Erlbaum.

Koedinger, K. R. (1992). *Emergent properties and structural constraints: Advantages of diagrammatic representations for reasoning and learning*. Working notes of the 1992 AAAI Spring Symposion on Reasoning with Diagrammatic Representations, Stanford University, March 27-29.

Kong, Y. T., & Rosenfeld, A. (Eds.) (1996). Special issue on topology and geometry in computer vision. *Journal of Mathematical Imaging and Vision, 6*, 105-308.

Kosslyn, S. M. (1973). Scanning visual images: Some structural implications. *Perception and Psychophysics, 14*, 90-94.

Kosslyn, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.

Kosslyn, S. M. (1987). Seeing and imagining in the cerebral hemispheres: a computational approach. *Psychological Review, 94*, 148-175.

Kosslyn, S. M. (1994a). *Image and brain - The resolution of the imagery debate*. Cambridge, MA: MIT Press.

Kosslyn, S. M. (1994b). *Elements of graph design*. New York: Freeman.

Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 47-60.

Kosslyn, S. M., & Pomerantz, J. P. (1977). Imagery, propositions and the form of internal representations. *Cognitive Science, 9*, 52-76.

Kosslyn, S. M., & Shin, L. M. (1994). Visual mental images in the brain: Current issues. In M. J. Farah & G. Ratcliff (Eds.), *The neuropsychology of high-level vision* (pp. 269-296). Hillsdale, NJ: Lawrence Erlbaum.

Kuhn, T. S. (1963). *The structure of scientific revolutions (2nd impr.).* Chicago, IL: University of Chicago Press.

Kuhn, W. (1993). Metaphors create theories for users. In A. U. Frank & I. Campari (Eds.), *Spatial information theory - A theoretical basis for GIS* (pp. 366-376). Berlin: Springer.

Kuhn, W., & Frank, A. (1991). A formalization of metaphors and image-schemas in user interfaces. In A. Frank & D. Mark (Eds.), *Cognitive and linguistic aspects of geographic space* (pp. 419-434). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Kuipers, B. (1982). The 'map in the head' metaphor. *Environment and Behavior, 14* (2), 202-220.

Kuipers, B. (1994). *Qualitative reasoning*. Cambridge, MA: MIT Press.

Kuipers, B. (2000). The spatial semantic hierarchy. *Artificial Intelligence, 119*, 191-233.

Kulik, L., & Klippel, A. (1999). Reasoning about cardinal directions using grids as qualitative geographic coordinates. In C. Freksa & D. M. Mark (Eds.), *Spatial information theory - Cognitive and computational foundations of geographic information science* (pp. 205-220). Berlin: Springer.

Kulpa, Z. (1994). Diagrammatic representation and reasoning. *Machine Graphics & Vision, 3* (1-2), 77-103.

Lakoff, G. (1987). *Women, fire, and dangerous things: what categories reveal about the mind*. Chicago, IL: University of Chicago Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*, 65-99.

Latecki, L. J., & Lakämper, R. (1999a). Convexity rule for shape decomposition based on discrete contour evolution. *Computer Vision and Image Understanding, 73* (3), 441-454.

Latecki, L. J., & Lakämper, R. (1999b). Contour-based shape similarity. *Proc. of Int. Conf. on Visual Information Systems, Amsterdam, June 1999*. Berlin: Springer.

Latecki, L. J., & Lakämper, R. (2000). Shape similarity measure based on correspondence of visual parts. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22* (10), 1185-1190.

Lee, T. R. (1968). Urban neighborhood as a socio-spatial schema. *Human Relations, 21*, 241-268.

Ligozat, G. (1998). Reasoning about cardinal directions. *Journal of Visual Languages and Computing, 9*, 23-44.

Logie, R. H. (1995). *Visuo-spatial working memory*. Hillsdale, NJ: Lawrence Erlbaum.

Logie, R. H. (2001). Working memory: A mental space for design and discovery. In J. S. Gero, B. Tversky, & T. Purcell (Eds.), *Visual and spatial reasoning in design II* (pp. 223-235). University of Sydney, Australia: Key Centre of Design Computing and Cognition.

Lynch, K. (1960). *The image of the city*. Cambridge, MA: MIT Press.

Marks, J. (1991). A formal specification for network diagrams that facilitates automated design. *Journal of Visual Languages and Computing, 2* (4), 395-414.

Marriott, K., & Meyer, B. (1998). *Visual language theory*. New York: Springer.

McNamara, T. P. (1986). Mental representations of spatial relations. *Cognitive Psychology, 18*, 87-121.

McNamara, T. (1991). Memory's view of space. *The Psychology of Learning and Motivation: Advances in Research and Theory, 27*, 147-186.

McNamara, T. P., Hardy, J. K., & Hirtle, S. C. (1989). Subjective hierarchies in spatial memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 15* (2), 211-227.

Meyer, D. E., & Schvanefeldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*, 227-234.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for information processing. *Psychological Review, 63*, 81-97.

Mishkin, J., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends in Neuroscience, 6*, 414-417.

Miyake, A., & Shah, P. (Eds.) (1999). *Models of working memory - Mechanisms of active maintenance and executive control*. Cambridge, UK: Cambridge University Press.

Moar, I., & Bower, G. H. (1983). Inconsistency in spatial knowledge. *Memory and Cognition, 11* (2), 107-113.

Montello, D. R. (1992). The geometry of environmental knowledge. In A. U. Frank, I. Campari, & U. Formentini (Eds.), *Theories and methods of spatio-temporal reasoning in geographic space* (pp. 136-152). Berlin: Springer.

Montello, D. R. (1993). Scale and multiple psychologies of space. In A. Frank & I. Campari (Eds.), *Spatial information theory: A theoretical basis for GIS* (pp. 312-321). Berlin: Springer.

Montello, D. R. (1998). A new framework for understanding the acquisition of spatial knowledge in large-scale environments. In M. J. Egenhofer & R. G. Golledge (Eds.), *Spatial and temporal reasoning in geographic information systems* (pp. 143-154). New York: Oxford University Press.

Montello, D. R., & Freundschuh, S. M. (1995). Sources of spatial knowledge and their implications for GIS: An introduction. *Geographical Systems, 2*, 169-176.

Myers, B. A. (1990). Taxonomies of visual programming and program visualization. *Journal of Visual Languages and Computing, 1* (1), 97-123.

Narayanan, N. H. (Ed.) (1992). *AAAI Spring Symposium on Reasoning with Diagrammatic Representations - Working Notes*.

Narayanan, N. H., & Hegarty, M. (1998). On designing comprehensible interactive hypermedia manuals. *Int. J. Human-Computer Studies, 48*, 267-301.

Newell, A. (1980). Physical symbol systems. *Cognitive Science, 4*, 135-183.

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

Nilsson, N. J. (1980). *Principles of artificial intelligence*. Palo Alto, CA: Tioga Publ. Co.

Noeth, W. (1990). Handbook of semiotics. Bloomington, IN: Indiana University Press.

O'Rourke, J. (1998). *Computational geometry in C* (2nd edition). Cambridge: Cambridge University Press.

Olivier, P. (2001). Diagrammatic reasoning: An artificial intelligence perspective. *Artificial Intelligence Review, 15*, 63-78.

Osherson, D. (Ed.) (1990). *An invitation to cognitive science (3 vols.).* Cambridge, MA: MIT Press.

Paivio, A. (1971). Imagery and language. In S. J. Segal (Ed.), *Imagery: Current cognitive approaches* (pp. 7-32). New York: Holt, Rinehart & Winston.

Paivio, A. (1986). *Mental representations - A dual coding approach*. Oxford: Oxford University Press.

Palmer, S. E. (1978). Fundamental aspects of cognitive representation. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 259-303). Hillsdale, NJ: Lawrence Erlbaum.

Peterson, M. (1995). *Interactive and animated cartography*. Englewood Cliffs, NJ: Prentice Hall.

Portugali, J. (Ed.) (1996a). *The construction of cognitive maps*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Portugali, J. (1996b). Inter-representation networks and cognitive maps. In J. Portugali (Ed.), *The construction of cognitive maps* (pp. 11-43). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Posner, M. (Ed.) (1989). *Foundations of cognitive science*. Cambridge, MA: MIT Press.

Preparata, F. P., & Shamos, M. J. (1985). *Computational geometry*. New York: Springer.

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin, 80*, 1-24.

Pylyshyn, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review, 87*, 16-45.

Pylyshyn, Z. W. (1990). Computing in cognitive science. In M. I. Posner (Ed.), *Foundations of cognitive science (2nd ed.)* (pp. 51-91). Cambridge, MA: MIT Press.

Pylyshyn, Z. W. (in press). Mental imagery: In search of a theory. *Behavioral and Brain Sciences*.

Randell, D. A., Cui, Z., & Cohn, A. G. (1992). A spatial logic based on regions and connection. *Proc 3rd Int. Conf. on Knowledge Representation and Reasoning* (pp. 165-176). San Mateo, CA: Morgan Kaufman.

Ratcliff, R. A., & McKoon, G. (1981). Does activation really spread? *Psychological Review, 88*, 454-462.

Raubal, M. M. (1997). *Structuring wayfinding tasks with image schemata*. M.S. Thesis, University of Maine.

Rauh, R., Schlieder, C., & Knauff, M. (1997). Präferierte mentale Modelle beim räumlich-relationalen Schließen: Empirie und kognitive Modellierung. *Kognitionswissenschaft, 6*, 21-34.

Rehkämper, K. (1995). Analoge Repräsentationen. In K. Sachs-Hombach (Ed.), *Bilder im Geiste: Zur kognitiven und erkenntnistheoretischen Funktion piktorieller Repräsentationen* (pp. 63-105). Amsterdam, Atlanta: Rodopi.

Retz-Schmidt, G. (1988). Various views on spatial prepositions. *AI Magazine, 9* (2), 95-105.

Richards, I. A., & Ogden, C. K. (1923). *The meaning of meaning*. London: Routledge and Kegan Paul.

Rickheit, G., & Habel, C. (Eds.) (1999). *Mental models in discourse processing and reasoning*. Amsterdam: Elsevier.

Rock, I., & Palmer, S. (1990). The legacy of gestalt psychology. *Scientific American, 263* (6), 48-61.

Roediger, H. L. (1980). Memory metaphors in cognitive psychology. *Memory & Cognition, 8* (3), 231-246.

Rumelhart, D. E., McClelland, J. L., & The PDP Research Group (1986). *Parallel distributed processing - Explorations in the microstructure of cognition. Volume 1: Foundations*. Cambridge, MA: MIT Press.

Rumelhart, D. E., & Norman, D. A. (1985). Representation of knowledge. In A. M. Aitkenhead & J. M. Slack (Eds.), *Issues in cognitive modeling* (pp. 15-62). Hillsdale, NJ: Lawrence Erlbaum.

Sadalla, E. K., Burroughs, W. J., & Staplin, L. J. (1980). Reference points in spatial cognition. *Journal of Experimental Psychology: Human Learning and Memory, 6* (5), 516-528.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 501-518.

Schacter, D. L., Cooper, L. A., Delaney, S. M., Peterson, M. A., & Tharan, M. (1991). Implicit memory for possible and impossible objects: Constraints on the construction of structural descriptions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 3-19.

Schlieder, C. (1996). Qualitative shape representation. In P. Burrough & A. Frank (Eds.), *Geographic objects with indeterminate boundaries* (pp. 123-140). London: Taylor & Francis.

Schlieder, C. (1998). *Diagrammatische Repräsentation - Ein algorithmischer Zugung mit kognitiven Implikationen*. Habilitation, University of Hamburg, Dept. for Informatics.

Schlieder, C. (1999). The construction of preferred mental models in reasoning with interval relations. In G. Rickheit & C. Habel (Eds.), *Mental models in discourse processing and reasoning* (pp. 333-357). Amsterdam: North-Holland.

Schlieder, C., & Berendt, B. (1998). Mental model construction in spatial reasoning: A comparison of two computational theories. In U. Schmid, J. F. Krems & F. Wysotzki (Eds.), *Mind modelling: A cognitive science approach to reasoning, learning and discovery* (pp. 133-162). Lengerich: Pabst Science Publishers.

Shamos, M. I. (1978). *Computational geometry*. Ph.D. Thesis, Dept. of Computer Science, Yale University

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*, 701-703.

Siegel, A. W., & White, S. H. (1975). The development of spatial representations of large-scale environments. In H. W. Reese (Ed.), *Advances in child development and behavior* (pp. 9-55). New York: Academic Press.

Simon, H. A. (1978). On the forms of mental representations. In C. W. Savage (Ed.), *Minnesota studies in the philosophy of science, Vol. ix: Perception and cognition: Issues in the foundations of psychology*. Minneapolis, MI: University of Minnesota Press.

Sloman, A. (1971). Interactions between philosophy and artificial intelligence: The role of intuition and non-logical reasoning in intelligence. *Artificial Intelligence, 2*, 209-225.

Sloman, A. (1975). Afterthoughts on analogical representations. *Proc. Theoretical Issues in Natural Language Processing (TINLAP-1)* (pp. 164-168). Cambridge, MA.

Sloman, A. (1994). Explorations in design space. In A. G. Cohn (Ed.), *Proceedings of the 11th Conference on Artificial Intelligence (ECAI'94)* (pp. 578-582). Chichester: Wiley.

Steinhauer, J. H., Wiese, T., Freksa, C., & Barkowsky, T. (2001). Recognition of abstract regions in cartographic maps. In D. R. Montello (Ed.), *Spatial Information Theory - Foundations of Geographic Information Science* (pp. 306-321). Berlin: Springer.

Stevens, A., & Coupe. P. (1978). Distortions in judged spatial relations. *Cognitive Psychology, 10*, 422-437.

Strube, G., Becker, B., Freksa, C., Hahn, U., Opwis, K., & Palm, G. (Eds.) (1996). *Wörterbuch der Kognitionswissenschaft*. Stuttgart: Klett-Cotta.

Sulin, R. A., & Dooling, D. J. (1974). Intrusion of a thematic idea in retention of prose. *Journal of Experimental Psychology, 103*, 255-262.

Suwa, M. Gero, J., & Purcell, T. (1999). Unexpected discoveries: How designers discover hidden features in sketches. In J. S. Gero & B. Tversky (Eds.), *Visual and spatial reasoning in design* (pp. 145-162). University of Sydney, Australia: Key Centre of Design Computing and Cognition.

Taylor, H. A., & Tversky, B. (1992). Spatial mental models derived from survey and route descriptions. *Journal of Memory and Language, 31*, 261-292.

Taylor, H. A., & Tversky, B. (1995). Assessing spatial representation using text. *Geographical Systems, 2*, 235-254.

Thorndyke, P. W. (1981). Distance estimation from cognitive maps. *Cognitive Psychology, 13*, 526-550.

Tolman, E. C. (1948). Cognitive maps in rats and men. *The Psychological Review, 55* (4), 189-208.

Trowbridge, C. C. (1913). On fundamental methods of orientation and imaginary maps. *Science, 38* (990), 888-897.

Tufte, E. R. (1998). *Envisioning information*. Cheshire, CT: Graphics Press.

Tversky, B. (1981). Distortions in memory for maps. *Cognitive Psychology, 13*, 407-433.

Tversky, B. (1991). Spatial mental models. *The Psychology of Learning and Motivation, 27*, 109-145.

Tversky, B. (1992). Distortions in cognitive maps. *Geoforum, 23* (2), 131-138.

Tversky, B. (1993). Cognitive maps, cognitive collages, and spatial mental models. In A. Frank & I. Campari (Eds.), *Spatial information theory - A theoretical basis for GIS* (pp. 14-24). Berlin: Springer.

Tversky, B., Franklin, N., Taylor, H. A., & Bryant, D. J. (1994). Spatial mental models from descriptions. *Journal of the American Society for Information Science, 45* (9), 656-668.

Vieu, L. (1997). Spatial representation and reasoning in AI. In O. Stock (Ed.), *Spatial and temporal reasoning* (pp. 5-41). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review, 72*, 89-104.

Weld, D. S., & De Kleer, J. (Eds.) (1990). *Readings in qualitative reasoning about physical systems*. San Mateo, CA: Morgan Kaufman.

Wessel, M., & Haarslev, V. (1998). VISCO: Bringing visual spatial querying to reality. *Proceedings 1998 IEEE Symposium on Visual Languages* (pp. 170-177). Los Alamitos, CA: IEEE Computer Society.

Yi-Luen Do, E., & Gross, M. D. (2001). Thinking with diagrams in architectural design. *Artificial Intelligence Review, 15*, 135-149.

Zimmermann, K. (1993). Enhancing qualitative spatial reasoning - Combining orientation and distance. In A. U. Frank & I. Campari (Eds.), *Spatial information theory - A theoretical basis for GIS* (pp. 69-76). Berlin: Springer.

Zimmermann, K. (1995). Measuring without measures - The Delta-Calculus. In A. U. Frank & W. Kuhn (Eds.), *Spatial information theory - A theoretical basis for GIS* (pp. 59-67). Berlin: Springer.

Zimmermann, K., & Freksa, C. (1996). Qualitative spatial reasoning using orientation, distance, and path knowledge. *Applied Intelligence, 6*, 49-58.

Subject Index