What's in a mental model?
On conceptual models in reasoning with spatial descriptions

Goran Hagert

UPMAIL
Uppsala Programming Methodology and Artificial Intelligence Laboratory
Computing Science Department
Uppsala University, P.O. Box 2050, S-750 02 Uppsala, Sweden

Abstract

A frequently used concept in Cognitive Science research is the notion of a mesial *model* Several very different areas of research use the concept as an explanatory construct. In this paper we explore one meaning of the concept developed for explaining cognitive reasoning with spatial concepts. Several experiments have been interpreted as showing that an abstract propositional representation is not sufficient for understanding reasoning in this domain. Another form of representation is the mental model which has been proposed to explain the obtained results. In contrast to this, we discuss some steps towards a computational theory founded on the notions of a *conceptua! model* and *levels of reasoning* which is able to reproduce the experimental results within a single abstract form of representation. Some empirical consequences are outlined and tested. A tentative conclusion is that the experimental results do not reflect fundamental differences in representational form but strategy differences in how to solve the tasks.

1. Introduction

An interesting explanatory concept in current Cognitive Science research is the notion of mesial *models.* The concept has been used in such diverse research areas as *rentoning* (e.g., |8|), *problem tolling* (e.g., [2]), and *human-computer interaction* (e.g., [9]). There must, of course, be several similarities between the different uses. However, no precise definition has been put forward as to what it means for something to be a mental model. That is, there is no definition which is supposed to hold among the different areas of research and it is even hard to find a definition that is not ambiguous within a particular area. So, one can really ask what it is that makes something a mental model or to put it differently, "what's *in a mental model"?*

A common theme within the three research areas is the emphasis on *knowledge repretcntation,* or perhaps we should say *mental repretentation.* Of particular interest in this respect is the ongoing debate about the status of *images and their* relation to mental models.

Some researchers claim that images correspond to a basic and concrete form of mental representation that is essential for the brain's way of representing the environment (e.g., [10]). Others hold the view that images can be analyzed into a more fundamental and abstract representation. In this view, images are seen as products of processes operating on an abstract representation (e.g., [14]). Thus, mental models as images could be regarded as either a fundamental format of representation or as a result of a strategy-driven process.

The concept of mental models as developed by Johnson-Laird [8] is very interesting in this perspective. In discussing how *meaning* is mentally represented, Johnson-Laird emphasized the distinction between a statement's *eztension* (or denotation) in contrast to its *istention* (or sense) and concluded that current theories of knowledge representation in Cognitive Science (e.g., semantic networks, lexical decomposition etc.) cannot give a satisfactory account of the relation between intension and extension. Instead, what is needed is a concept of a *mental moid* in order to explain how meaning comes into mind.

For example, when a reasoner heart a sentence like,

The knife it to the right of the fork,

it is encoded in an abstract propositional (or conceptual) representation, e.g.,

Right(Knife,Fork),

but its meaning is not reached until the reasoner builds a model of it, e.g.,

[Fork Knife]

It is clear that for many domains a mental model corresponds to an image. This is particularly relevant when the domain consists of spatial descriptions as the above example. Thus, a mental model in this conception is basically realized through an introduction of a new representational format. Moreover, it seems as if the notion is based on the assumption that a model, by definition, should be concrete, i.e., the model should have stru*ctural similarities* to the real world object.

In this paper we intend to discuss the concept of a mental model as developed by Johnson-Laird and discussed in |3), [4], and [5]. Naturally, we are also interested in generalizing the theoretical work to mental models in general. The fundamental concept we explore is the notion of a *conceptual model* showing many resemblances with *conceptual ttructuret* [15] and s*tructural detcriptiont* [7|. The computational theory we try to build is essentially an attempt to realize the "mental model phenomenon" in a conceptual form of representation. The emphasis is on processes operating on an abstract representation and thus creating an "image-like" structure. Let us start by a brief review of some experiments which gives the discussion a concrete content.

2. Experiments with spatial descriptions

The task domain we use as a means for the discussion has been studied by, for example, [12] and [8].

Description 1. *The spoon it on the left of the knife. The plate it on the right of the knife. The fork it in front of the tpoon. The cup it in front of the knife.*

Description 2. *The tpoon it on the left of the knife. The plate it on the right of the tpoon. The fork it in front of the tpoon. The cup it in front of the knife.*

Figure 1. Examples of spatial descriptions.

Figure 1 shows two spatial descriptions used in the experiments in [12]. The task is to read a description and then (i) to draw a scene, (ii) to recognize a scene, or (iii) to recognize the sentences among a set. In [3] we reviewed five experiments and presented steps towards a computational theory able to reproduce the empirical findings. In the following we concentrate on the notion of s*patial determinacy,* which is the difference between the descriptions in Figure 1.

2.1 Spatial determlnacy and mental models

The experimental situation for descriptions 1 and 2 is as follows [12]. You are instructed to read a set of sentences and afterwards you might be presented with

- a diagram and asked to decide whether or not the diagram corresponds to the description jnst read (e.g., Diagram 1) or
- four descriptions each containing three sentences and asked to rank order the sets in terms of their resemblance to the original set.

The four sets could be composed of the original description, a paraphrase set or a set containing an inference, and two confusions sets.

spoon	knife	plate
fork	cup	

Diagram 1

The difference between descriptions 1 and 2 is easily seen if Diagram 1 is compared to Diagram 2. Description 2 is ambiguous since both of these diagrams satisfy the description which is not the case for description 1. Description 1 is a *determinate description* whereas 2 is an *indeterminate description* [12].

spoon	plate	knife
fork		cup

Diagram 2

It is argued that a pure conceptual or propositional representation is unable to catch the difference between an indeterminate and a determinate description. This argument is verified if the descriptions are represented in a semantic network (c.f., [1]). That is, there is no structural difference between the two propositional representations which corresponds to the difference in the underlying meaning.

If this conception is cognitively relevant, human reasoners should be able to process and remember the descriptions equally well. On the other hand, if the network is further processed into a mental model format - for instance, a mental diagram as shown in Diagram 1 or 2 - then the differences between determinate and indeterminate descriptions should show up.

The experiments showed that human reasoners were able to recall diagrams as consistent with the descriptions more correctly if they were determinate. Also, the reasoners were more likely to recognize the original descriptions if they were indeterminate. This was interpreted in [12] and [8] as indicating two quite different processes. When the description is determinate, reasoners build a mental diagram of the scene which makes it possible to mix original and inferred sentences. When the indeterminate description is processed it leads only to a verbal or propositional encoding.

Thus, these experiments together with others have been interpreted as not supporting a pure propositional form of representation, but a notion of mental models [8]. However, in the next two sections we show that this interpretation is not completely without problems. We start by introducing a general view of reasoning processes.

3. Reasoning and conceptual models

The steps towards a computational framework of reasoning have been discussed in [3], [4], and [5]. In this section we review some of the central aspects of this framework.

3.1 Designing and running conceptual models

The fundamental concept in the theory is the notion of a *conceptual model* Briefly, we consider a reasoning process to be composed of a *design process* in which a conceptual model of the task is built. In this process the reasoner tries to interpret the information making up the task and activate knowledge that can help the design process. When a model has been developed it can be run. Thus, we can understand a conceptual model as a program which can be interpreted and executed. By *running the model* it is possible to generate inferences that, depending on the task, might lead to a terminating state. In order to execute the model, we need a control structure, which is discussed below. The next phase is a *refinement process* in which the model is modified according to results of the execution or to some other knowledge source. The model can then be run again, and so on. A more detailed description can be found in [4] and [5].

Moreover, we want to explore a *flexible control structure* which can be modified during a reasoning process. Of course, this is to move the "real" control somewhere else. What we want to reach is the ability of executing the same conceptual model according to different control strategies. For example, the model might be executed by a depth-first search strategy through a forward inference rule or by a breadth-first search through a backward inference rule. We discuss this further in Section 3.2.

Finally, we intend to model the above processes within a logic modeling framework [4] [5] meaning that we emphasize logic as a representational language. It also means that we use computational techniques developed in the area of Logic Programming (11). This is in line with proposals and claims raised by, for example, [13].

3.2 Levels of reasoning and conceptual models

In logic it is a well-founded idea that a reasoning system can be viewed from two perspectives. Usually these are referred to as the *object level* and the *meta level,* respectively. For instance, the meta-level has most often been discussed in terms of how to control a line of reasoning. What inference should be made? How is the search through the space conducted? These are of course important questions for any reasoning system.

In discussing different logical levels of problem solving Sterling [16] identified three different levels and called them *domain level, methods level* and *planning level* In a sense, the latter two concern the way a representation can be manipulated and reasoned about, i.e., a meta-level. We use the same terminology as in [16] except for the methods level which is called *strategy level* here.

The distinction illustrates the levels we want to emphasize in the present context. Earlier we proposed the term *flexible control structure* in order to emphasize the ability of a reasoning system - be it a simulation model or an AI program - to modify and redirect its reasoning. Mechanisms for such processes can be explored if we make a distinction between different levels of reasoning.

A conceptual model should include all three levels in order to be executable. In some domains it might not be necessary to develop a very sophisticated planning level since no essential problem solving is involved. That is, the planning level corresponds to reasoning about how to design a strategy to use at the domain level. The levels indicate that it is a possible for a reasoner to carry out a line of reasoning on the domain level or at the planning level. In our conception, the latter means that the reasoner tries to design *or* refine the model so that it can be executed at the domain level.

With respect to the mental representation issue it is important to consider the domain level of a conceptual model. That is, how is the problem actually represented? To take our current example, what basic representation and process can underlie the processing of spatial descriptions?

In the remainder of this paper we will focus on the domain and strategy levels. An important point we are going to make is that it is not possible to predict a domain level without considering a strategy level since experimental differences might not correspond to differences at the considered level.

4. Conceptual models in spatial reasoning

Let us assume a domain level in which the basic form of representation is a proposition as found in different semantic or propositional network theories, c.f., (1). For example, description 1 in Figure 1 is then represented as follows,

Left(Spoon,Knife) (1)
Right(Plate,Knife) (2)
Front(Fork,Spoon) (3)
Front(Cup,Knife) (4)

We use basically first-order predicate logic with an implicit ∀ quantification. Variables will denoted by x, y, z, etc. and constants by a beginning upper case letter. Function and terms begin with a lower case letter.

In [3] we proposed domain-knowledge to be represented as s*chems* built up through a set of rules. In this context, we assume a schema to belong to the domain level The following set of rules represents a hypothetical part of possible domain-knowledge.

Left(x,y) ∧ Left(y,z) → Left(x,left(y,z)) (5)

Left(x,y) ∧ Front(z,x) → Left(front(z,x),y) (6)

Left(x,y) ∧ Front(z,y) → Left(x,front(z,y)) (7)

Left(x,left(y,z)) → Left(x,y) ∧ Left(x,z) ∧ Left(y,z) (8)

Left(front(x,y),z) → Left(x,z) ∧ Left(y,z) ∧ Front(x,y) (9)

Left(x,front(y,z)) → Left(x,y) ∧ Left(x,z) ∧ Front(y,z) (10)

Left(x,y) ↔ Right(y,x) (11)

Left(x,y) → ¬ Right(x,y) (12)

Right(x,y) → ¬ Left(x,y) (13)

Behind(x,y) ↔ Front(y,x) (14)

The rale in (5) can be translated to: *If tie on the left of y* and y *i$ on the left of* 1, *then z to the left of* a *composite object eon$i$ting of y end z tuck thnt y is to the left of z*. The rule in (10) says: x *i$ on the left of y* and *of z nni in front of* y, *if * is on the left of the composite object in which p i$ in front of z*

This rule-schema can compute a *composite repretentntion* or stru*ctural description* from a set of sentences. Furthermore, it can also decompose a structural description into its elementary constitutents. The different usages are dependent on the particular design *of* the strategy level.

Executing a conceptual model consisting of the above schema and a strategy level based on a forward inferencing and a depth-first search on description 1 we get the following edited trace showing assertions and where they come from.

Left(Spoon,Knife) (1)

Right(Plate,Knife) (2)

Left(Knife,Plate) (11)

Left(Spoon,left(Knife,Plate)) (5)

Front(Fork,Spoon) (3)

Left(front(Fork,Spoon),left(Knife,Plate)) (6)

Front(Cup,Knife) (4)

Left(front(Fork,Spoon),Knife) ∧ Left(front(Fork,Spoon),Plate) ∧
Left(Knife,Plate) (8)

Left(front(Cup,Knife),Plate) (6)

Left(front(Fork,Spoon),front(Cup,Knife)) (7)

Left(front(Fork,Spoon),left(front(Cup,Knife),Plate)) (5)

Let us also give an example trace when the description b indeterminate. In this case, the execution of the conceptual model will not lead to a single composite representation. The process terminates with a few partially combined structures which cannot be further processed into a single composite representation given the design of the schema. We can say that we have a *partial composite representation.*

Left(Spoon,Knife) (1)

Right(Plate,Spoon) (2)

Left(Spoon,Plate) (11)

Front(Fork,Spoon) (3)

Left(front(Fork,Spoon),Knife) (6)

Left(front(Fork,Spoon),Plate) (6)

Front(Cup,Knife) (4)

Left(front(Fork,Spoon),front(Cup,Knife)) (7)

If we change the strategy level it is possible for the model to derive sentences which were not included in the original description. Notice that we do not change the domain level, but only *how the domain level is used.* For example, if we have the structure

Left(front(Fork,Spoon),front(Cup,Kfiift))

we can through the rules in (9) and (10) infer that

Left(Fofk,Cup)

This short illustration should have indicated some of the properties we find especially interesting with respect to the notion of a conceptual model and levels of reasoning. It is also clear that this theory can explain the results discussed in Section 2 [3]. In the next section, we will discuss some empirical observations of human reasoners engaged in the same tasks as were formally treated above.

5. Conceptual models end empirical relevance

Suppose a human reasoner is asked to read a set of sentences describing a scene (as in Figure 1) and then to

- draw the scene described in the sentences or
- rank order a set of four descriptions according to their verbal similarity

What can then be predicted if the difference between the descriptions lies in how spatially determined they are? What consequences follow the notion of conceptual models as described above?

The major and essential difference between processing a determinate description versus an indeterminate description as found in the present theory, is the degree to which the process leads to a single composite conceptual representation. In the latter case, we end up with a partial composite representation whereas in the former case we end up with a single composite representation or structural description. Let us discuss the tasks of drawing a scene from memory and recognizing a description from memory given this conceptualisation.

In Section 4 we showed how the execution contained at least two partial representations when description 2 was processed. They were as follows,

Left(front(Fork,Spoon),front(Cup,nife)) (15)

Left(front(Fofk,Spoon),Plate) (16)

Let us assume a drawing process which simply takes a composition as input and outputs the objects accordingly. It is easy to show bow this can be done using the same domain model but changing the strategy level to a backward inference mechanism (3). We get Diagram 3 when (16) is taken as input.

Diagram 3

Next, we execute the process with (16) as input and Diagram 3 is extended to include one more object and we get Diagram 4,

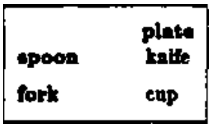

Diagram 4

Diagram 4 should indicate (in bold face) what objects are processed and that the plate and the knife "share" the same position relative the spoon.

The drawing process for a determinate task is straightforward. This exercise shows that both processes are very similar, but that *processing effort it greater when the task* is *indeterminate.* The very same general prediction holds for the task of recognizing the original sentences among a set. Thb is in contrast to predictions found in, for example, (8) and (12).

5.1 Methodological considerations

Fifteen subjects (Ss) participated in this experiment. Each of the Ss solved 12 spatial reasoning tasks of which 6 were indeterminate and 6 were determinate (examples are found in Figure 1). Afterwards, the Ss were asked to either (i) draw the scene described or (ii) rank order four alternatives in terms of their resemblance to the original description.

In each group of descriptions (i.e., determinate or indeterminate) 2 of the descriptions lead to the task of drawing the scene, 2 contained the task of rank ordering where one alternative contained a paraphrase (i.e., if the original description contained a sentence like, *The fork* it *to the left of the spoon,* then it was replaced by, *The tpoon it to right of fork),* and the remaining 2 consisted in a rank-ordering task where one alternative contained and inferred sentence (i.e., a sentence not included in the original description but which was true irrespective of whether or not the original was indeterminate or determinate). The content of the tasks varied, i.e., each task described the relations among five objects where the same objects could not be found in another task.

Each S was tested individually and the whole session was tape-recorded enabling us to record reading-times as well as solution-times for each task. No time-limit was set. The 12 descriptions were randomized for each subject.

5.2 Some empirical results

We only give some very brief comments on the results that are of special interest in the present context. A more detailed account is given in |6).

Drawing a scene. A drawing was judged to be correct if it was consistent with the corresponding description. The somewhat surprising result was that 16 out of 30 drawings were correct for both types of descriptions. It was also observed that there was no systematic difference in the types of errors the Ss made.

Of the 15 correct drawings when the description was indeterminate, 12 could be classified as being similare. in structure to Diagram 4. That is, where two objects were placed in the same position relative a third object.

Table 1. Mean reading time and mean drawing time for both descriptions (sec).

	Indeterminate	Determinate
Reading	104.8	105.3
Drawing	57.3	41.1

From Table 1 we can see that there is no difference in the time taken to read the descriptions. However, drawing the scene requires more time when the description is indeterminate.

Recognising a description. The percentage of rank-orderings in which the original and the paraphrase were ranked prior the confusion descriptions was 80% for the determinate descriptions and 82% for the indeterminate descriptions.

The corresponding percentages when the task included a description with an inferred sentence were 87% for the determinate cases and 83% for the indeterminate ones. No major difference in reading times or in times for doing the rank-ordering was observed.

In short, the pattern that emerges is very similar for both tasks. The sufficient point to make here is that the experiment does not reject the framework developed in this paper. Of course, it is possible that the result is in line with other theories as well. It is interesting to note, though that Ss did draw a scene in the manner outlined earlier in this section. Only in 3 cases did Ss draw a hypothetical scene, i.e., a scene in which the indeterminacy was resolved by a "guess" and where the drawing is similar to Diagram 1 or 2 rather than 4.

0. Concluding remarks

We have tried to discuss the mental model concept within a conceptual frame of reference. Rather than postulating different forms of representations we have assumed a basic abstract form embedded in a conceptual model. We emphasized that the model includes different levels of reasoning. The focus here have been on the domain and strategy level. It is obvious that we need more analyses and studies, both computational and experimental in order to be more precise as to what the levels might contain and how they interact. In particular, how can modification of conceptual models be understood in this perspective?

Taken together, the experiments briefly described in Section 2 and in Section 5, seem to indicate that in order to further develop a computational theory we need to consider not only the domain level but also the strategy level. It is on this and the planning level we find a reasoner's search for a strategy that is effective relative the task demands and the goals. In a sense, this is only a reformulation of the thesis discussed by Anderson [1| in which the claim was that it is necessary to consider both *repretentation and process* in order to develop a computational theory.

Acknowledgment. This work is supported by the National Swedish Board for Technical Development (STU). I thank Anna Malmsjo for collecting data described in Section 5 and Ake Hansson for several valuable discussions.

6. References

|I] Anderson, JR. Language, Memory, and Thought Hillsdale, New Jersey: Erlbuam, 1976.

[2] Gentner, D. *k* Stevens, A.L. Mental Models. Hillsdale, New Jersey: Erlbaum, 1983.

|3] Hagert, G. Modelling mental models: Experiments in cognitive modelling of spatial reasoning. Proc. of the Sixth European Conference on Artificial Intelligence, ECAI-84, Pisa, Italy, 1984.

[4] Hagert, G. *k* Hansson, A. Logic modelling of cognitive reasoning. Proc. of the Eighth International Joint Conference on Artificial Intelligence, IJCAI-83, Karlsruhe, West Germany, 1983.

[5] Hagert, G. *k* Hansson, A. Reasoning models within a logical framework. Technichal Report No. 25, UP MAIL, Uppsala University, 1984.

[6| Hagert, G. *k* Malmsjo, A. Mental representation and leveb of reasoning. In preparation, 1985.

[7] Hinton, G. Some demonstrations of the effects of structural descriptions in mental imagery.Cognitive Science, 3,1979.

|8] Johnson-Laird, P.N. Mental Models. London: Cambridge University Press, 1983.

[9] Kieras, D.E. *k* Bovair, S. The role of a mental model in learning to operate a device. Cognitive Science, 8, 1984.

[10] Kosslyn, S.M. *k* Pomerantz, J.R. Imagery, propositions, and the form of internal representations. Cognitive Psychology, 9, 1977.

[11] Kowalski, R. Logic for Problem Solving. New York: North-Holland, 1979.

[12] Mani, K. *k* Johnson-Laird, P.N. The mental representation of spatial descriptions. Memory *it,* Cognition, 10,1982.

[13] Moore, R.C. The role of logic in representation and commonsense reasoning. Proc. of The National Conference on Artificial Intelligence, AAAI-82, Pittsburgh, USA, 1982.

[14] Pylyshyn, Z. Computation and cognition. Cambridge, Massachusetts: MIT Press, 1984.

[15] Sowa, J.F. Conceptual structurest Information processing in mind and machine. Amsterdam, The Netherlands: Addison-Wesley, 1984.

[16] Sterling, L. Logical leveb of problem solving. Proc. of the Second International Logic Programming Conference, Uppsala, Sweden, 1984.