Preferences in Cardinal Direction

Marco Ragni (ragni@cognition.uni-freiburg.de)
Department of Cognitive Science, University of Freiburg
Germany

Benedikt Becker (beckerb@informatik.uni-freiburg.de)
Department of Cognitive Science, University of Freiburg
Germany

Abstract

How do we reason with imprecise spatial descriptions? Do reasoners typically prefer one conclusion (over another) consistent with the imprecise descriptions? Based on empirical findings we are able to give a positive answer for the second question for spatial reasoning with cardinal direction relations. Analyzing further the pattern of the preferred conclusion reflects the idea of informativeness of the description. In consequence, we briefly explain heuristics and present a Bayesian model representing subjective belief of the reasoner.

Keywords: Probabilistic Reasoning; Preferential Reasoning; Qualitative Reasoning

Introduction

Reasoning with spatial information requires sometimes to reason with incomplete information. Take for example,

Berlin is north-east of Paris.
Paris is north-west of Rome.

You can (based on this information alone, e.g. no background knowledge, no map) easily infer that Berlin must be north of Rome. But you cannot infer (based on this information alone) if Berlin is eastern or western of Rome. But if you have to reason without having assumptions about geographic positions – do we prefer certain relations? The question on how humans solve such deduction problems is at the core of qualitative reasoning. In other words, how do we infer new knowledge (a *conclusion*) from given knowledge, and moreover, what are the differences to formal approaches in artificial intelligence?

Formally, there are two main approaches in AI on how such reasoning problems can be solved: By the application of (transitivity) rules or by the construction and inspection of models. Principally, both approaches are equivalent (Russell & Norvig, 2003), i.e. it is not possible to derive more information with each of these methods. This equivalence, however, makes it harder to distinguish which method(s) is applied by humans while solving such problems. Nonetheless, a number of empirical studies investigates this research question by psychological means. The most prominent and best supported theory with respect to the number of effects that can be accounted for is the theory of mental models (MMT) (Johnson-Laird & Byrne, 1991) (to name only a few: the indeterminacy effect (Johnson-Laird & Byrne, 1991), the form of premises and the figural effect (Knauff, Rauh, Schlieder, & Strube, 1998), the wording of conclusions (Van der Henst & Schaeken, 2007), etc.). According to the MMT, linguistic processes are relevant to transfer information from the premises into a spatial array and back again, but the reasoning process itself relies on model manipulation only. A *mental model* is an internal representation of objects and relations in spatial working memory, which matches the state of affairs given in the premises. The semantic theory of mental models is based on the mathematical definition of deduction, i.e. a propositional statement φ is a consequence of a set of premises $\mathcal{P}$, written $\mathcal{P} \models \varphi$, if in each model $\mathcal{A}$ of $\mathcal{P}$, the conclusion φ is true.

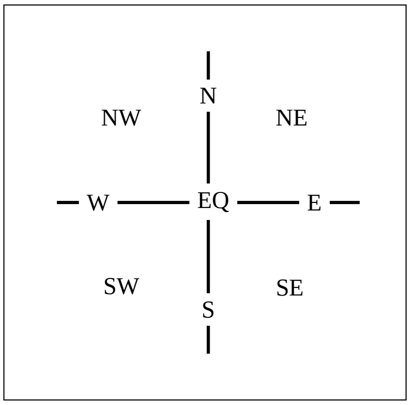

Figure 1: The nine base relations of the cardinal direction calculus in the projection based representation. Other representations are cone-based representations (Ligozat, 1998)

An interesting finding is the so-called preference effect, i.e. in multiple model cases (nearly always) one preferred model is constructed from participants and used as a reference for the deduction process (Rauh, Hagen, Schlieder, Strube, & Knauff, 2000). Further findings showed that during the validation phase alternative models are constructed by small modifications to the initially constructed model. This was the reason why the mental model theory for spatial reasoning was extended within the framework of preferred mental models (Rauh et al., 2000).

A new research line (Oaksford & Chater, 2007) focuses on Bayesian explanations for preferred solutions, e.g. for syllogistic reason. The authors use here the notion of informativeness to explain why a certain quantifier is used. The question is still open, if the Bayesian approach is sufficient to model spatial reasoning.

This paper is structured as follows: First, we will present an empirical investigation analyzing the question about preferences in cardinal direction. Our empirical findings are then analyzed w.r.t. the main theories in the field (Theory of Mental Models, Theory of Mental Logic) with heuristics and we present a Bayesian model representing subjective belief of the reasoner. Finally, we discuss the different findings.

Preferences in Cardinal Direction

The language of cardinal direction consists of points in the euclidean plane $\mathbb{R}^2$. Based on the point algebra it is possible to distinguish 9 base relations $a, b \in \mathbb{R}^2$:

CD	EQ	N	NE	E	SE	S	SW	W	NW
PA	(=,=)	(=,>)	(>,>)	(>,=)	(>,<)	(=,<)	(>,<)	(<,=)	(>,>)

In other words $a\ N\ b := a_x = b_x \wedge a_y > b_y$, b is a northly. An *assignment* of a set of CD constraints C over the vocabulary $\mathcal{B} = \{\mathrm{N}, \mathrm{NE}, \mathrm{E}, \mathrm{SE}, \mathrm{S}, \mathrm{SW}, \mathrm{W}, \mathrm{NW}, \mathrm{EQ}\}$ is a function $\alpha : V(C) \rightarrow \mathbb{R}^2$, mapping each variable x, occuring in C to coordinates in the real plane.

Over the euclidean plane these *jointly exhaustive* and *pairwise disjoint*-base relations (cp. Fig. 1) with the composition table (cp. Figure 2) form a relation algebra. In the first experiment discussed here we used relations from the set $\mathcal{B}' := \mathcal{B} \backslash \{\mathrm{EQ}\}$ to construct a type of relational reasoning task that is referred to as three-term-series-problems (3ts-problems) in cognitive research (e.g. (Hunter, 1957)). In these tasks always two statements are used as premises and the task of the participants is to generate a statement that is consistent with the premises – the conclusion. E.g.,

A is northeast of B.
B is west of C.
Which relation holds between A and C?

The 3ts-problems can be formally described by the composition of two base relations and the question for a satisfiable relation. The set of all possible relations with premises $a\ R_1\ b$, $b\ R_2\ c$ are denoted by the composition $R_1 \circ R_2$. Normally, it is presented as a composition table (cf. Figure 2).

For the above example $\mathrm{NE} \circ \mathrm{W}$ contains the following three relations: NE, N, NW. Since CD consists of nine base relations, there are without EQ 64 possible compositions of two base relations. In other words, exactly 64 different three-term-series problems exist. If we omit all one-relation cases (cells with one entry in Figure 2), it results in 40 multiple relation cases out of the 64 possible compositions. The participants of our studies were confronted with all 64 problems and had to infer a conclusion.

Empirical Data

The first central question we are interested in is: How do people reason about cardinal directions? Do they construct preferred mental models, and if so, what are the principles? An answer to this question might give hints of how preferences differ between large-scale spaces and small-scale spaces. For the latter scale of space, preferences have already been identified (Ragni, Fangmeier, Webber, & Knauff, 2007).

Participants. 24 students of the University of Freiburg took part in this web experiment (14m/10f, $M = 23.5/22.1, SD = 2.3/2.1$). They were paid for their participation.

	NW	N	NE	W	E	SW	S	SE
NW	1 [NW] NW 1.0 0.21	1 [NW] NW 1.0 0.21	3 [NW,N,NE] N 0.91 0.21	1 [NW] NW 1.0 0.29	3 [NW,N,NE] N 0.89 0.36	3 [SW,W,NW] W 0.83 0.14	3 [SW,W,NW] W 0.78 0.36	8(9) [B] W 0.29 0.0
N	1 [NW] NW 1.0 0.36	1 [N] N 1.0 0.14	1 [NE] NE 1.0 0.43	1 [NW] NW 1.0 0.29	1 [NE] NE 1.0 0.36	3 [SW,W,NW] W 0.67 0.36	2(3) [N,S] N 0.5 0.14	3 [SE,E,NE] E 0.64 0.21
NE	3 [NW,N,NE] N 1.0 0.29	1 [NE] NE 1.0 0.5	1 [NE] NE 1.0 0.36	3 [NW,N,NE] N 0.89 0.36	1 [NE] NE 1.0 0.29	8(9) [B] N 0.21 0.0	3 [SE,E,NE] E 0.73 0.21	3 [SE,E,NE] E 1.0 0.29
W	1 [NW] NW 1.0 0.29	1 [NW] NW 1.0 0.39	3 [NW,N,NE] N 0.6 0.29	1 [W] W 1.0 0.57	2(3) [W,E] E 0.58 0.14	1 [SW] SW 1.0 0.29	1 [SW] SW 1.0 0.21	3 [SW,S,SE] S 0.5 0.29
E	3 [NW,N,NE] N 0.88 0.43	1 [NE] NE 1.0 0.5	1 [NE] NE 1.0 0.29	2(3) [W,E] W 0.77 0.07	1 [E] E 1.0 0.29	3 [SW,S,SE] S 0.78 0.36	1 [SE] SE 1.0 0.43	1 [SE] SE 1.0 0.21
SW	3 [SW,W,NW] W 0.9 0.29	3 [SW,W,NW] W 0.78 0.36	8(9) [B] W 0.29 0.0	1 [SW] SW 1.0 0.57	3 [SW,S,SE] S 0.44 0.36	1 [SW] SW 1.0 0.43	1 [SW] SW 1.0 0.29	3 [SW,S,SE] S 0.91 0.21
S	3 [SW,W,NW] W 0.7 0.29	2(3) [N,S] S 0.82 0.21	3 [SE,E,NE] E 0.78 0.36	1 [SW] SW 1.0 0.43	1 [SE] SE 1.0 0.64	1 [SW] SW 1.0 0.5	1 [S] S 1.0 0.29	1 [SE] SE 1.0 0.29
SE	8(9) [B] SE 0.43 0.0	3 [SE,E,NE] E 0.73 0.21	3 [SE,E,NE] E 0.9 0.29	3 [SW,S,SE] S 1.0 0.36	1 [SE] SE 1.0 0.36	3 [SW,S,SE] S 1.0 0.21	1 [SE] SE 1.0 0.29	1 [SE] SE 1.0 0.21

Figure 2: The preferred relations in reasoning with cardinal direction. In each cell, the first number gives the number of correct relations and the relations. In the second row we have the preferred relation, then in the indeterminate case the relative frequency of this relation, i.e. how often it was chosen by the participants and then the error rates.

Materials. The experiment used the whole set of Cardinal Direction relations presented in Fig. 1. In the main part of the experiment all participants had to solve the same set of 64 3ts-problems. Here is an example-problem:

A is northwest of B.
B is southeast of C.
Which relation holds between A and C?

In half of the trials we asked for the relation between A and C and in half of the trials between C and A.

Procedure and Design. The experiment was conducted as a web experiment (partially conducted at our site for control) using webexp2. Tasks were presented in a randomized order. The premises were presented sequentially, i.e. the first premise disappeared when the second premise appeared. In other words, the participants were forced to hold the premise information in the working memory. All premises were presented in a self-paced procedure. Finally, the participants had to give a relation as an answer.

Overall, 87% of the problems were correctly solved. The results regarding the preference effects can be found in Figure 2.

As shown in Figure 2 out of the given 64 problems exactly 24 are determinate problems and 40 are indeterminate problems. Most of the indeterminate problems exactly 90% (only 4 relations were not significantly preferred: $N \circ S$, $W \circ SE$, $W \circ E$, $SW \circ E$) were solved with a clear preference for one relation. However, it is remarkable that several relations could have been chosen as a possible conclusion, but, in fact, the participants chose just one of them and their preferences also often corresponded.

Discussion. There are differences between preferred relations in small-scale spaces and in large-scale spaces. Contrary to the small-scale spaces (Ragni, Fangmeier, et al., 2007) where the first-free fit strategy has been identified in relational reasoning in large-scale spaces participants used a first-fit strategy. In other words they inserted the third object C in-between A and B (cp. the relations $S \circ N$ and $E \circ W$ where in the first case S and in the second W has been reported). The inverse composition $N \circ S$ and $W \circ E$ are not statistically significant.

By a formal analysis it was possible to explain the preferred mental model in indeterminate cases by the following distinction

- Principle 1 (*In-between Insertion Principle*): If the two relations of the composition are inverse (e.g. S and N, W and E) then the third object C is inserted in-between A and B,(e.g. A is S of C and B is north of C, and so on).
- Principle 2 (*Cut Principle*): Choose always the relation in the geometrical cut of the two relations, i.e. if $NE \circ NW$ is composed and the relations NW, N, NE are possible than the relation N is chosen.

The participants preferred the cut between relations, e.g. in the composition of $NE \circ NW$ and $NW \circ NE$ they preferred the relation N. The same pattern holds as well for $SW \circ NW$ and so on. This gives an indication that without additional information they use (independently of projection based or cone based representation of Cardinal Direction) similar distances.

Theories of Deduction

In this section we ground the intuitively used theories formally (and mathematically) and analyze them with respect to their reasoning power.

A *relational structure* is a tuple $(D,(R_i)_{(i \in I)})$ consisting of a domain D (sometimes called discourse universe) and a set of (usually binary) relations R_i (Russell & Norvig, 2003). For example, geographic knowledge like *New York is north-east of Washington* can be expressed by cardinal direction relations $N, NE, E, SE, \dots$ over the domain of cities. More complex expressions can be formed by using connectives like conjunctions (New York is north-east of Washington *and* New York is in the U.S.) and disjunctions ($\dots$ *or* $\dots$). By allowing negations, we have a propositional relational language $\mathcal{L}$ over cardinal direction relations. Such relational structures can be used to describe *knowledge representation*. But how can new information be derived?

Theory of mental logic

The theory of mental logic (Rips, 1994) assumes that we use (transitivity) rules to draw conclusions, whereas the classical model theory argues that we use models for this inference process. The classical mental model theory (Byrne & Johnson-Laird, 1989) claims that in multiple model cases (i.e. more than one model is consistent with the premises) other models are inspected.

1. $West(x,y)$ & $North(z,x) \rightarrow West(z,y)$
2. $West(x,y)$ & $North(z,y) \rightarrow West(x,z)$
3. $West(x,y)$ & $West(y,z) \rightarrow West(x,z)$
4. $West(x,y) \leftrightarrow East(y,x)$
5. $(West(y,x)$ & $West(z,x)) \rightarrow (West(y,z)$ *or* $West(z,y))$
6. $(West(y,z)$ *or* $West(z,y))$ & $North(w,z) \rightarrow$ $(West(y,w)$ *or* $West(w,y))$

Figure 3: Set of (incomplete) inference rules specified for spatial reasoning adapted from Van der Henst (2002).

The central idea of this approach can be characterized as follows: "Reasoning consists in the application of mental inference rules to the premises and conclusion of an argument. The sequence of applied rules forms a mental proof or derivation of the conclusion from the premises, where these implicit proofs are analogous to the explicit proofs of elementary logic" (Rips, 1994, p. 40). Hagert (1984) defined a first set of spatial inference rules (cf. Fig. 2). This set of rules has been extended by two additional rules (cf. the rules 5 and 6 in Fig. 2) to deal with indeterminacy by Van der Henst (2002). The rules in Fig. 2 are successively applied to the premises of a problem description.

There is, however, no recent theory in explaining the construction of the preferred relations (Figure 2).

Theory of mental models

The mental model theory assumes that the human reasoning process consists of three distinct phases: The *model generation phase*, in which a first model is constructed out of the premises, an *inspection phase*, in which the model is inspected to check if a putative conclusion is consistent with the current model. In the *validation phase*, finally, alternative models are generated from the premises that refute this putative conclusion. The indeterminacy effect is mainly responsible for human difficulty in reasoning (Johnson-Laird, 2001).

Recent findings indicate a phenomenon encountered in multiple-model cases, namely that humans generally tend to construct a *preferred mental model* (PMM). This model is easier to construct, less complex, and easier to maintain in working memory compared to all other possible models (Knauff et al., 1998). The principle of economicity is the determining factor in explaining human preferences (Manktelow, 1999). This principle also explains that a model is constructed incrementally from its premises. Such a model construction process saves working memory capacities because each bit of information is immediately processed and integrated into the model (Johnson-Laird & Byrne, 1991). In the model variation phase, this PMM is varied to find alternative interpretations of the premises (Rauh et al., 2000). From a formal point of view, however, this theory has not been formalized yet and is therefore not fully specified in terms of necessary operations to process such simple problems as were described above.

A model $\mathcal{A}$ is called *consistent* with a set of premises Φ over a relational language $\mathcal{L}$ (mathematically $\mathcal{A} \models \Phi$) if all expressions of Φ are true in $\mathcal{A}$. Then a conclusion Ψ can be derived from the premise set Φ (mathematically $\Phi \models \Psi$, whereby $\models$ is called the *consequence relation*) if

$$\begin{aligned}\Phi \models \Psi \quad &\Leftrightarrow \quad \text{All models of } \Phi \text{ are models of } \Psi. \\ &\Leftrightarrow \quad \text{There is no model } \mathcal{A} \text{ with} \\ &\qquad \mathcal{A} \models \Phi \text{ and } \mathcal{A} \models \neg\Psi.\end{aligned}$$

A model $\mathcal{A}$ with the property $\mathcal{A} \models \Phi$ and $\mathcal{A} \models \neg\Psi$ is called *counter-example*. It follows if there is a counter-example to Φ and Ψ then $\Phi \models \Psi$ cannot hold.

This classical (mathematical) consequence relation, however, does not explain how initial mental models are constructed and varied (Rauh et al., 2000). Since there is a huge empirical evidence supporting the preferred mental model theory for different calculi (Rauh et al., 2000; Ragni, Fangmeier, et al., 2007; Ragni, Tseden, & Knauff, 2007) it seems worth to ground this theory mathematically.

A Probabilistic Approach

As already stated, a new approach are probabilistic models (Oaksford & Chater, 2007) to explain preferred relations. Those are based on the consideration to use probabilities instead of truth values as the representation of semantics. This is a valid consideration as a probability might be interpreted in a *subjective* manner describing a subjective degree of belief rather than a relative frequency of an event. Following this subjective interpretation probability theory can be utilized for belief updating and inference. The probabilistic approach to inference is based on:

$$P(\text{“If } p \text{ then } q\text{”}) = P(q|p). \tag{1}$$

Thus, the probability of a conditional proposition is identified with the conditional probability of the proposition. The a-posteriori belief in the fact q in face of certainty about the fact p is given by the a-priori conditional probability: $P_1(q) = P_0(q|p)$, if $P_1(p) = 1$. This is called “conditionalization”. It constitutes the basis of probabilistic inference.

The probabilistic representation of conditionals as given in equation 1 enables the application of Bayes' theorem:

$$P(q|p) = \frac{P(p|q)\, P(q)}{P(p)} \tag{2}$$

This has two advantages: First, whereas $P(q|p)$ is a rather abstract value, the probabilities of its right hand side can often be derived from the agent's experience. Second, it implies basic patterns of performance while reasoning with conditional propositions which appear as “errors and biases” from a logicistic standpoint.

Bayesian Rationality arises from a rational analysis of the problem, the environment, and the constraints of an agent while conducting deductive tasks. As such, it is not a theory of the actually psychological processes in use, but a description of general regularities. It is further independent of cognition *about* probabilities. It shows that cognition often obeys the laws of probabilistic theory.

The following models[1] are to reproduce the frequency distribution of the 3ts-task on cardinal directions this way.

Spatial Bayesian Models The spatial reasoning task of the previous section uses the set of cardinal relations $\mathcal{B}' = \{\mathrm{N}, \mathrm{E}, \mathrm{S}, \mathrm{W}, \mathrm{NE}, \mathrm{SE}, \mathrm{SW}, \mathrm{NW}\}$. The statement of an item in the 3ts-task is given by a pair of relations $R_1, R_2 \in \mathcal{B}'$ with aR_1b and bR_2c for three locations a, b, c. The subject's guess for the relation between a and c is another relation $R_3 \in \mathcal{B}'$. The relative frequency of R_3 for an item R_1, R_2 will be referred to as $f_{R_1,R_2}(R_3)$.

The objective of a Bayesian model for the 3ts-task is to implement a probability distribution of R_3 parametrized by the task item R_1, R_2, i.e. $P_{R_1,R_2}(R_3)$. This probability distribution is assumed to be a prediction of the experiment's relative frequencies f_{R_1,R_2}. Thus, model M's preferred relation given the task's relations R_1 and R_2 is

$$M(R_1, R_2) := \underset{R_3 \in \mathcal{B}'}{\arg\max}\, P_{R_1,R_2}(R_3).$$

The per-item probability distribution of R_3 can be identified with the probability of R_3 conditioned by the item's relations. Therefore, it further allows the application of Bayes' theorem (equation 2):

$$P_{R_1,R_2}(R_3) := P(R_3|R_1, R_2) = \frac{P(R_1, R_2|R_3)\, P(R_3)}{P(R_1, R_2)} \tag{3}$$

Consequently, it is sufficient for a Bayesian model of the 3ts-task to specify merely the reversed conditional probability $P(R_1, R_2|R_3)$ as well as the marginal probabilities $P(R_3)$ and $P(R_1, R_2)$.

The following sections will describe such implementations. The quality of each model M will be compared to the empirical data by three factors: a) the mean correlation C^M between P_{R_1,R_2} and the empirical data f_{R_1,R_2}, b) the sum E^M of the squared differences between P_{R_1,R_2} and f_{R_1,R_2} and c) the number N^M of correctly predicted preferred relations.

The Unit Layout (Model M_1) The computation of $P^{M_1}(R_1, R_2|R_3)$ is based on a heuristic for detours when moving by R_3 in the so called *unit layout*. R_1 and R_2 describe the detour. The farther the detour the smaller is the conditional probability of R_1, R_2.

The unit layout is a rectangular subset of $\mathbb{Z}^2$ and separately defined for each direction R_3. The brackets $[\cdot]^{R_3}$ map the locations a and c each to a field in $\mathbb{Z}^2$ such that $[a]^{R_3} R_3 [c]^{R_3}$. Each pair of relations R_1, R_2 with $R_3 \subset R_1 \circ R_2$ is likewise mapped to a field in $\mathbb{Z}^2$ by $[\cdot]^{R_3}$ such that

$$[a]^{R_3} R_1 [R_1R_2]^{R_3} \text{ and } [R_1R_2]^{R_3} R_2 [c]^{R_3}.$$

The fields $[a]^{R_3}$ and $[c]^{R_3}$ must be chosen in such a way that each $[R_1R_2]^{R_3}$ is definite. That way, the unit layout is definite in $\mathbb{Z}^2$ up to translations. Figure 4 shows the unit layout for $R_3 = \mathrm{NW}$.

[1] The source code is available at http://tiny.cc/hmi3f.

SE-NW	S-NW	SW-NW	SW-N	SW-NE
E-NW	*a*	W-NW	W-N	W-NE
NE-NW	N-NW	NW-NW	NW-N	NW-NE
NE-W	N-W	NW-W	*c*	NW-E
NE-SE	N-SE	NW-SE	NW-S	NW-SE

Figure 4: The *unit layout* for $R_3 = \text{NW}$. Field a is to the northwest of c. All other field are uniquely labeled with relations R_1-R_2. It holds for each of them that field a is R_1-wards of it and it is R_2-wards of c.

For $R_3 \subset R_1 \circ R_2$, the unit layout entails the costs of a "detour" moving from field $[a]^{R_3}$ via $[R_1 R_2]^{R_3}$ to $[c]^{R_3}$ utilizing a metric d on $\mathbb{Z}^2$.

$$c^{R_3}_{R_1,R_2} := \frac{d([a]^{R_3}, [R_1 R_2]^{R_3}) + d([R_1 R_2]^{R_3}, [c]^{R_3})}{d([a]^{R_3}, [c]^{R_3})}$$

The costs $c^{R_3}_{R_1,R_2}$ for $R_3 \not\subset R_1 \circ R_2$ are defined by the model parameter distimposs. This cost measure entails the wanted conditional probability:

$$P^{M_1}(R_1, R_2 | R_3) := \frac{{c^{R_3}_{R_1,R_2}}^{-1}}{\sum_{R_1', R_2' \in \mathcal{B}'} {c^{R_3}_{R_1',R_2'}}^{-1}}.$$

This points out the influence of the model parameter distimposs: For infinity, the model performs accurate and it simulates errors for positive numbers.

This is how model M_1 computes the conditional probability of the right-hand side of equation 3. The marginal probability of $P^{M_1}(R_3)$ is a unit distribution which can be furnished with a probability gain for the main cardinal directions by cardinalgain and an additional gain towards the west by westgain. The probability of $P^{M_1}(R_1, R_2)$ is assumed to be a unit distribution.

Parameter Variations Varying the metric d between Euclidian, Manhattan, and maximum had no noteworthy effect on the quality estimation factors (C^M, E^M, N^M). So we chose the euclidian metric, as it matches the intuitive concept of distance best. The model parameter distimposs was varied systematically between 20 and 200, the parameters cardinalgain and westgain were varied between 0.1 and 0.9.

We found a maximal convergence against the empirical data with model parameters distimposs $= 150$, cardinalgain $= 0.2$, and westgain $= 0.2$.

It has a mean correlation $C^{M_1} = 0.91$, a summed error of $E^{M_1} = 2.82$ and predicts the preferred relation correctly in $N^{M_1} = 59$ cases. This instance of model M_1 has a mean correlation of 0.96 in 60 items of the task. Nevertheless, the mean correlation for the task items with opposing intermediate directions is as little as 0.17. This suggest the appearance of another strategy in these cases.

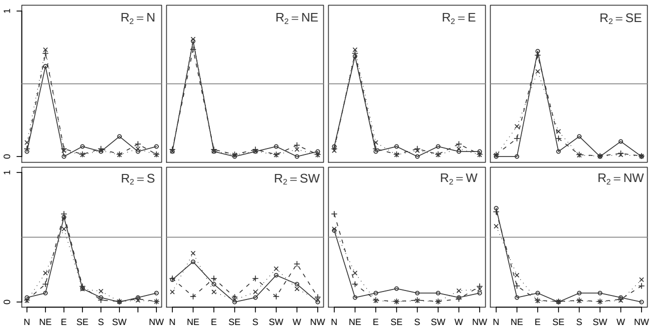

Figure 5: Relative frequencies of R_3 for task items with $R_1 = \text{NE}$ from the experiment (circles: $\circ$) as well as M_1's (pluses: $+$), and M_2's (crosses: $\times$) probabilities .

A Secondary Strategy (Model M_2) Model M_2 is an extension of the model presented in the preceding section. It adds a probability gain g_{R_1,R_2} to the value of $P^{M_1}_{R_1,R_2}$. This gain implements priming effects on the relations R_1 and R_2. The amounts of priming towards R_1 and R_2 are controlled by the model parameters firstprim and secondprim, respectively. Values of 0 each void the priming effect.

The extent of this probability gain is in turn controlled per task item by the certainty z_{R_1,R_2} of M_1

$$z_{R_1,R_2} := \max_{R_3 \in \mathcal{B}'} P^{M_1}_{R_1,R_2}(R_3).$$

The (yet to be normalized) probability distribution of M_2 is defined as

$$P^{M_2}_{R_1,R_2}(R_3) := z_{R_1,R_2} \cdot P^{M_1}_{R_1,R_2}(R_3) + (1 - z_{R_1,R_2}) \cdot g_{R_1,R_2}(R_3).$$

It weakens M_1's probability distribution $P^{M_1}_{R_1,R_2}$ and strengthens the priming effect g_{R_1,R_2} as a function of decreasing certainty.

Parameter Variations In a systematic search through the parameters of model M_1 as well as firstprim and secondprim we found an instance of M_2 with mean correlation of $C^{M_2} = 0.94$, summed error $E^{M_2} = 2.67$ and $N^{M_2} = 62$ correctly predicted items. Along with it, this instance has a mean correlation of 0.73 for the task items with opposed intermediate directions. The parameters were distimposs = 180, cardinalgain = 0.1, westgain = 0.2, firstprim = 0.3 and secondprim = 0.2.

Figure 5 shows results both from model M_1 and M_2 for $R_1 = \text{NE}$. M_2's improvement is apparent for $R_2 = \text{SW}$.

Interpretation

The following lines give a clue of how the found model parameters can be read as a hints on the underlying cognitive processes.

Utilizing Experience The first model, M_1, shows that the spatial reasoning task can be modelled by a Bayesian approach. The computation of $P(R_3 | R_1, R_2)$ is based on an "intuition of the benefit" to move towards R_1 first and then towards R_2 to attain towards R_3 overall. This intuition might reflect complying knowledge of the subject arising from basic experience navigating through the world.

It was possible to further increase the convergence of the model towards the empirical data by means of higher marginal probabilities of the cardinal directions, and additionally the west. This might reflect frequency effects for the cardinal directions as well as an effect of the reading direction for the western direction.

Shifting Strategies Whereas model M_1 behaved poorly for tasks with opposed intermediate directions, model M_2's correlation on those could be improved by simulating priming effects on the relations given by the current task item. Those tasks excel in a high uncertainty about the answer. This suggests the subjects shift their strategy to be driven by priming effects under uncertainty.

General Discussion

If incomplete information is available only (i.e. several relations are possible), humans tend to take a relation more into account than others. This finding complements a series of findings for preferred spatial reasoning with intervals (Rauh et al., 2000), with the spatial relations right and left (Jahn, Knauff, & Johnson-Laird, 2007), and with topological relations (Ragni, Tseden, & Knauff, 2007).

Our starting point was the question if it is possible to model preference effects for cardinal directions in both theories (the Mental Model Theory and the Bayesian rationality) based on heuristics. Only by a formalization it is possible to compare human reasoning to approaches in AI. A formal handling of the preferred mental model theory by a consequence relation allows to make precise predictions about which kind of conclusion(s) are drawn (from a given set of premises) and which are neglected. These heuristics can be described by two principles: the in-between insertion principle and the cut principle. Both together can explain the preferences in the composition table (Figure 2) and support the theory of cognitive economicity (Manktelow, 1999).

The primer raised question, if the Bayesian approach is expressible enough to model preference effects in spatial reasoning (with cardinal directions) can be positively answered. Moreover, it reproduces the full frequency distribution quiet well: The first model is based on a heuristic for detours which explains the preferences (Figure 2). It has a mean correlation of 0.91 and predicts the preferred relation correctly in 59 from 64 cases. The second model which adds a priming effect leads to an increase from 0.17 to 0.73 in the correlation in the four cases of opposed intermediate directions.

A possible limitation of the Bayesian model is connected to the certainty of the conclusion. While each statement is given with absolute certainty (Berlin is north-east of Paris) a conclusion has only a degree of certainty. Taken together, the results clearly indicate that the preference effect can be explained by heuristics in both mental models and bayesian approach. Further research necessarily requires an investigation for a general heuristic explaining preference relations for the diverse spatial calculi.

One point, however, is certain: the role of heuristics has been vastly underestimated in explaining the preferences in spatial reasoning.

References

Byrne, R. M., & Johnson-Laird, P. N. (1989). Spatial reasoning. *Journal of Memory & Language*, *28*(5).

Hagert, G. (1984). Modeling mental models: Experiments in cognitive modeling spatial reasoning. In T. O'Shea (Ed.), *Advances in artificial intelligence.* Elsevier.

Hunter, I. M. (1957). The solving of three-term series problems. *Br J Psychol*, *48*(4), 286–98.

Jahn, G., Knauff, M., & Johnson-Laird, P. N. (2007). Preferred mental models in reasoning about spatial relations. *Memory and Cognition*.

Johnson-Laird, P. N. (2001). Mental models and deduction. *Trends in Cognitive Sciences*, *5*(10).

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Hove (UK): Erlbaum.

Knauff, M., Rauh, R., Schlieder, R., & Strube, G. (1998). Continuity effect and figural bias in spatial relational inference. In *Proceedings of the 20th cognitive science conference* (pp. 573–578). Mahwah, NJ: Lawrence Erlbaum.

Ligozat, G. (1998). Reasoning about cardinal directions. *Journal of Visual Language Computing*, *9*(1), 23–44.

Manktelow, K. (1999). *Reasoning and Thinking*. Hove: Psychology Press.

Oaksford, M., & Chater, N. (2007). *Bayesian rationality the probabilistic approach to human reasoning*. Oxford University Press.

Ragni, M., Fangmeier, T., Webber, L., & Knauff, M. (2007). Preferred mental models: How and why they are so important in human spatial reasoning. In C. Freksa, M. Knauff, B. Krieg-Brückner, B. Nebel, & T. Barkowsky (Eds.), *Spatial cognition v.* Berlin: Springer.

Ragni, M., Tseden, B., & Knauff, M. (2007). Cross cultural similarities in topological reasoning. In S. Winter, M. Duckham, L. Kulik, & B. Kuipers (Eds.), *Proceedings of 8th International Conference on Spatial Theory, COSIT 2007* (Vol. 4736). Springer.

Rauh, R., Hagen, C., Schlieder, C., Strube, G., & Knauff, M. (2000). Searching for Alternatives in Spatial Reasoning: Local Transformations and Beyond. In *Proceedings of the 22nd Annual Conference of the Cognitive Science Society.* Mahwah, NJ: Lawrence Erlbaum Associates.

Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking*. Cambridge, MA: MIT Press.

Russell, S., & Norvig, P. (2003). *Artificial intelligence: A modern approach* (2nd ed.). Prentice-Hall.

Van der Henst, J. (2002). Mental model theory versus the inference rule approach in relational reasoning. *Thinking and Reasoning*, *8*, 193-205.

Van der Henst, J., & Schaeken, W. (2007). The wording of conclusions in relational reasoning. *Cognition*, *97*, 1–22.