



Spatial relations in reasoning : a computational model

François Olivier

► To cite this version:

François Olivier. Spatial relations in reasoning : a computational model. Philosophy. Université Paris sciences et lettres, 2022. English. NNT : 2022UPSLE003 . tel-03984759

HAL Id: tel-03984759

<https://theses.hal.science/tel-03984759>

Submitted on 13 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à l'École Normale Supérieure, Paris

**Spatial Relations in Reasoning:
A Computational Model**

Soutenue par

François Olivier

Le 24 mars 2022

Ecole doctorale n° 540

**Lettres, arts, sciences
humaines et sociales**

Spécialité

Philosophie



Composition du jury :

Laure VIEU Directrice de recherche, Université de Toulouse / IRIT	<i>Présidente</i>
Paul ÉGRÉ Directeur de recherche, ENS Ulm / IJN	<i>Directeur de thèse</i>
Valeria GIARDINO Chargée de recherche, ENS Ulm / IJN	<i>Examinatrice</i>
Salvador MASCARENHAS Professeur junior, ENS Ulm / IJN	<i>Examineur</i>
Carl Peter Leslie SCHULTZ Professeur, Université d'Aarhus / ECE	<i>Rapporteur</i>
Igor DOUVEN Directeur de recherche, Université Paris 1 / IHPST	<i>Rapporteur</i>

Spatial Relations in Reasoning: A Computational Model

Dissertation

Submitted in partial fulfillment of the requirements for the degree

of

Doctor in Philosophy

by

François Olivier

Under the supervision

of

Professor Paul Égré

Examination Committee

Supervisor:	Paul Égré	École Normale Supérieure - IJN
Committee:	Igor Douven	Université Paris 1 - IHPST
	Valeria Giardino	École Normale Supérieure - IJN
	Salvador Mascarenhas	École Normale Supérieure - IJN
	Carl Peter Leslie Schultz	Aarhus University - ECE
	Laure Vieu	Université de Toulouse - IRIT



DEC
DÉPARTEMENT
D'ÉTUDES
COGNITIVES



This work was supported by a doctoral scholarship from Paris Sciences et Lettres ANR-10-IDEX-0001-02 PSL. The writing of the dissertation has been helped by the Département d'Études Cognitives of the École Normale Supérieure de Paris, by the Institut Jean Nicod, and received support from grants ANR-17-EURE-0017 FrontCog, and ANR-19-CE28-0019-01 AMBISENSE.

Abstract

Studies in the psychology of spatial reasoning often focus on positional problems, such as "A is to the left of B, C is to the right of B, where is A compared to C?". The computational model PRISM developed within the model-based approach, simulates in detail the psychological processes that enable us to solve this type of problem, especially by including relevant cognitive elements such as attentional limits. However, no computational model seems to exist for our reasoning about other types of relations, such as circle inclusion, or those used in geometry. This dissertation aims to fill this gap by building a computational model generalized to all types of spatial relational problems.

To achieve this goal, a clarification of the nature of the mental representations used in spatial reasoning is necessary in the first place. By means of arguments concerning the inferential possibilities of different kinds of representations, it is concluded that the mental representations underlying our spatial reasoning capacities must be diagrammatic in nature. This clarification enables us to propose a formalization of these mental representations, as well as a representation system where they can be linked to the natural language sentences used to create them. It is then possible to benefit from the computational tool *Answer Set Programming Modulo Theories for the Qualitative Spatial domain* ASPMT(QS), recently created in artificial intelligence, to take into account all the geometric constraints that apply in this representation system.

Indeed, ASPMT(QS) allows us to define relations from natural language mathematically, to reason about these relations formally, but above all for our purpose, to produce a diagrammatic representation compatible with a description using these relations. Since the representation produced by ASPMT(QS) is also diagrammatic in nature, it can be considered as the mental representation underlying our reasoning. Based on these possibilities and by including the psychological principles already present in PRISM, the computational model EPRISM ('E' for Extended) is built to simulate our psychological reasoning for all types of spatial relational problems.

Finally, some theories of cognitive linguistics show that many expressions and notions, whose meaning is nevertheless abstract, seem to be understood based on our spatial reasoning capacities. This opens up the possibility of using EPRISM to shed new light on topics discussed in cognitive science, such as the modulation process of the theory of mental models or the study of conditional reasoning.

Résumé

Les études en psychologie du raisonnement spatial se concentrent souvent sur des problèmes de positionnement d'objets, comme par exemple "A est à gauche de B, C est à droite de B, où est A par rapport à C ?". Le modèle computationnel PRISM développé dans la lignée de la théorie des modèles mentaux, simule en détail les processus psychologiques qui nous permettent de résoudre ce genre de problème, notamment en intégrant des éléments cognitifs pertinents comme les limites de notre attention. Aucun modèle computationnel ne semble toutefois exister pour modéliser nos raisonnements sur d'autres types de relations, comme par exemple l'inclusion de cercles, ou encore celles utilisées en géométrie. Cette thèse tente de combler cette lacune en proposant un modèle computationnel généralisé à tout type de problème relationnel spatial.

Pour ce faire, une clarification de la nature des représentations mentales qui permettent le raisonnement est d'abord nécessaire. Au moyen d'arguments concernant les possibilités inférentielles de différents types de représentations, il est conclu que les représentations mentales qui sous-tendent notre capacité de raisonner spatialement doivent être de nature diagrammatique. Cette clarification permet de proposer une formalisation de ces représentations mentales, ainsi qu'un système de représentation où elles peuvent être mises en lien avec les phrases du langage naturel utilisées pour les créer. Il est ensuite possible de bénéficier de l'outil computationnel

Answer Set Programming Modulo Theories for the Qualitative Spatial domain ASPMT(QS), récemment proposé en intelligence artificielle, pour rendre compte de toutes les contraintes géométriques qui s'appliquent dans ce système de représentation. En effet, ASPMT(QS) permet de définir mathématiquement des relations du langage naturel, de raisonner formellement sur ces relations, mais surtout, de produire une représentation diagrammatique compatible avec une description utilisant ces relations. Vu que la représentation produite par ASPMT(QS) est aussi de nature diagrammatique, elle peut être considérée comme la représentation mentale sous-tendant notre raisonnement. En exploitant ces possibilités tout en incluant les principes psychologiques déjà présents dans PRISM, le modèle computationnel EPRISM ('E' pour Extended) est proposé afin de simuler notre raisonnement psychologique pour tout type de problème relationnel spatial.

Enfin, des théories de linguistique cognitive montrent que de nombreuses expressions et concepts, dont le sens est pourtant abstrait, semblent être compris grâce à nos capacités de raisonnement spatial. Cela permet d'utiliser EPRISM pour apporter un éclairage nouveau sur différents problèmes discutés dans les sciences cognitives, comme le processus de modulation de la théorie des modèles mentaux ou l'étude du raisonnement conditionnel.

Acknowledgments

Ever since my first year as a philosophy student, the idea of formalizing mental representations has fascinated me and has been the topic of many of my essays. When studying cognitive science in Paris, I became aware of the essential role of space in cognition during a first internship with Valeria Giardino, and thanks to a second internship with Paul Égré, I was able to sharpen my formal skills which would be useful to me later on.

Paul suggested working on conditionals and negations, a topic so rich and captivating that it ended up being a stepping stone to my PhD. We managed to obtain successful results for an experiment, which led to an article at the student session of the European Summer School in Logic, Language and Information (ESSLLI 2019). In parallel, I was exploring formalisms for spatial reasoning, convinced that it could benefit some conditional sentences. The exploration of these spatial systems led to a first result published in the international cycle of Spatial Cognition conferences in 2020, which aimed at the extension of the computational model PRISM. This dissertation can be seen as the in-depth elaboration of this result. Eventually, a possible application of spatial analysis to conditionals became more and more apparent, and Chapter 10 describes this possibility. Now, besides the opportunity

that these few words offer to understand the reason for the (possibly too wide) range of topics addressed in this dissertation, I hope they will fill the acknowledgments that follow with meaning.

My deepest thanks go to my supervisor, Paul Égré, who always offered me all his support, encouragement and insights, whatever the topic studied. Paul also managed to create a great working environment that made these years exciting and a genuine source of emancipation in my academic life. To the members of the Examination Committee, Laure Vieu, Valeria Giardino, Igor Douven, Carl Schultz and Salvador Mascarenhas, I would like to express my gratitude for accepting to be *rapporteurs* and *examineurs-rices* of this dissertation. Special thanks go to Carl, Valeria and Salvador for their help and insightful ideas during my PhD. Finally, I would like to thank Valeria Giardino and Hans Rott for being in my Supervision Committee during these years.

I am thankful to my colleagues and to researchers, from the Institut Jean Nicod and beyond, who dedicated some time and thoughts to my research, Steven Franconeri, Markus Knauff, Beidi Li, Jakob Suchan, Benjamin Spector, Hitoshi Omori, Heinrich Wansing, Bruno Leclercq, Roberto Casati, Barbara Tversky, Denis Buehler, Jeremy Zehr, Jackson Graves, Sébastien Vacher, Janet Foot, Nathalie Evin, Marie Lods and Vincent Gaudefroy. To my colleagues and friends who also shaped this dissertation through numerous coffee breaks and discussions, Antoine van Muylder, Mathias Sablé-Meyer, Dawei Bei, Axel Baptista, Romain Bourdoncle, Janek Guerrini, Benjamin Icard, Géraldine Carranante, Anastasia Tsilia, Can Konuk, Jakob Süskind, Céline Boisserie-Lacroix, Constant Bonard, Chloé Tahar and Sofiane Mbarki. To my friends in Paris and Belgium, thank you for all the fun and memories.

Finally, I want to thank with all my heart my parents, for their constant support

throughout my studies and PhD, their help in so many ways, and most importantly, their presence. Last but certainly not least, I would like to thank Ellen Boes, who has been ineluctably on the front line with me during this PhD. Not only has Ellen been a sharp-witted interlocutor to discuss my ideas, she has also been, with an unsurpassable constancy, a source of affection, support and complicity during these important years.

Contents

Abstract	v
Acknowledgments	ix
1 Introduction	1
1.1 Spatial Relations in Reasoning	1
1.2 How Advanced is the Psychology of Spatial Relational Reasoning? . .	3
1.3 Why is an Extension to All Kinds of Spatial Reasoning Called For? . .	6
1.4 Structure of the Dissertation	7
1.5 Preliminary Useful Remarks	8
1.5.1 <i>Formal</i> versus <i>Psychological</i> Resolution	8
1.5.2 <i>Homomorphic, Scientific</i> and <i>Logical</i> Models	9
I Motivations	11
2 The Theory of Mental Models	13
2.1 Facing Mental Logic	13
2.2 General Principles of the Theory of Mental Models	14
2.2.1 Models are Conjunctions of Possibilities	15
2.2.2 Dual-system Processing and the Search for Counterexamples .	16

2.3	How to Characterize Mental Models?	18
3	The Space to Reason Theory	21
3.1	Spatial Information in the Brain	22
3.1.1	The Two-Streams Hypothesis	22
3.1.2	Imagery: Recreating the World in Our Mind	24
3.1.3	Cognitive Evidence for the Two-Streams	25
3.2	Spatial Models Support Reasoning	26
3.3	The Nature of Spatial Models	27
3.3.1	We Reason Based on Qualitative Spatial Layout Models	27
3.3.2	Spatial Models Reduced to the Structure of an Array	29
4	The Preferred Model Theory and PRISM	31
4.1	The Main Principles of the Preferred Model Theory	33
4.1.1	What Are The Alternative Models?	33
4.1.2	The Preferred Model and the Principle of Cognitive Economy	34
4.1.3	No Spontaneous Search for Alternative Models	34
4.2	PRISM: The Computational Model	35
4.2.1	Resolution Through Reasoning Phases	36
4.2.2	The Attentional Focus	37
4.3	Motivations for a Computational Model for All Kinds of Relations	39
4.3.1	Limits Imposed by PRISM	39
4.3.2	Why an Extension?	40
II	Spatial Models as Diagrammatic Representations	43
5	The Diagrammatic Hypothesis	45

5.1	The <i>Inferential Argument</i>	47
5.1.1	The Representation (token) and its Information (types)	48
5.1.2	<i>Constraints</i> Between Types	49
5.1.3	The Potential for FREE RIDE (P1)	51
5.1.4	The Potential for CONSISTENCY CHECK (P1 Continued)	52
5.1.5	Model-Based Approaches Presuppose Cognitive Potentials (P2)	53
5.1.6	Diagrams: the Most Simplified of the Graphical Class (P3)	55
5.1.7	Conclusion: the <i>Diagrammatic Hypothesis</i> (C)	57
5.2	Further Discussions about the Hypothesis	59
5.2.1	Spatial Models as Images	61
5.2.2	Attentional Capacities, Retinotopy and the Hypothesis	62
5.2.3	The Information, Not the Representation, is Qualitative	64
6	OVER-SPECIFICITY: its Formal and Psychological Consequences	67
6.1	A Representation System for Psychology	68
6.1.1	Shimojima's Presentation of OVER-SPECIFICITY	68
6.1.2	The Relevant Spatial Aspects for Psychology	69
6.1.3	A Spatial Model in the Representation System	70
6.1.4	Clarifying Knauff's Characterization	72
6.2	Default Specifications	73
6.2.1	Naturalness of the Spatial Model Constructed	73
6.2.2	Default Specifications and Constraints	75
6.2.3	Harmless in Reasoning, Explanatory for Psychology	77
6.3	Variation Phase and Deductive Reasoning	78
6.3.1	Putting the Notions of <i>Variation</i> in Order	79
6.3.2	Variation of the Default Model	80
6.3.3	Certainty and Generality: the Trade-Off	83

6.3.4	Psychological Deductive Strategies	85
6.4	Concluding Remarks	86
III	EPRISM: A Computational Model for Spatial Reasoning	89
7	Architecture	91
7.1	ASPMT(QS): Implementing the <i>Representation System for Psychology</i> . .	93
7.1.1	Represented Objects: Qualitative Relations as Facts	93
7.1.2	Representation: Values for a Spatial Model	98
7.1.3	Governed by Geometric Constraints	99
7.2	The Roles of the Reasoning Components	102
7.2.1	The Controller	103
7.2.2	ASPMT(QS), its <i>Semantic, Formal</i> and <i>Psychological</i> Roles	104
7.2.3	Representations in the Spatial Working Memory Module	105
7.2.4	The Roles of the Reasoner's Attention	107
8	Reasoning Phases: a First Approach	109
8.1	Reasoning Process and its Resolution	109
8.1.1	Reasoning Steps	110
8.1.2	The Construction Phase	110
8.1.3	Reasoning on the Constructed Model	114
8.2	Two Examples in EPRISM	116
8.2.1	A Typical PRISM Problem in EPRISM	116
8.2.2	Claire's Reasoning in EPRISM	120
8.3	Concluding Remarks	122

IV	Natural Language, World Knowledge and Connectives	123
9	Formalizing Natural Language	125
9.1	Mapping to Formal Objects and Relations	126
9.1.1	Mapping to Formal Objects	126
9.1.2	Formal Definitions of Relations from Natural Language	127
9.1.3	Constraining the Set of Possible Relations	129
9.2	Structures and Non-Spatial Sentences	130
9.2.1	The Image Schema Theory	130
9.2.2	Three Different Components	133
9.2.3	Capturing the Spatial Dimension of Simple Image Schemas . .	134
9.2.4	Existing Formal Approaches	135
9.2.5	Defining CONTAINMENT for EPRISM	137
9.2.6	Defining SOURCE-PATH-GOAL for EPRISM	139
9.2.7	Difficulties with the Image Schema Theory and Promising Trends	140
10	Logical Connectives	143
10.1	Constraints and Sets of Spatial Models	144
10.1.1	Sets of Spatial Models	144
10.1.2	Constraints Apply To Each Spatial Model of a Set	145
10.2	Connectives in the Construction Phase	148
10.2.1	The <i>Modulation Process</i> in the Theory of Mental Models	148
10.2.2	Replacing the Modulation Process by Geometric Constraints . .	149
10.3	Reasoning with Connectives	152
10.3.1	Some Valid Arguments In Classical Logic	152
10.3.2	The Satisfiability Requisite	154
10.3.3	Accounting For Problematic Arguments	155

10.3.4 A Case of Acceptability	158
10.3.5 Concluding Remarks	159
11 General Conclusion	161
11.1 To Sum Up	161
11.2 Deductive Reasoning as Spatial Reasoning	163
Appendices	166
A Installation of ASPMT(QS)	166
B Basic Principles and Logical Foundation of ASPMT(QS)	167

Chapter 1

Introduction

1.1 Spatial Relations in Reasoning

This dissertation is about our cognitive capacity to reason about spatial relations described through language. Examples of sentences that produce such reasoning are the following:

- (1) John is in the yard, and the yard is separated from the lawn. So John is not on the lawn.
- (2) Point A is on Segment B, and Segment B crosses Circle C. So Point A might be in Circle C.
- (3) Peter is to the left of Cathy, so Cathy is to the right of Peter.
- (4) The ball is in the box, so the ball is smaller than the box.
- (5) Mary lives in London, but London is not in France. So Mary does not live in France.

The main hypothesis of this dissertation is that we understand such sentences by creating a diagrammatic representation in our mind, and then evaluate the inference

by inspecting this constructed representation. For a sentence, its diagrammatic representation, either mental or drawn on a graphical support, can be defined as follows:

Diagrammatic Representation: the *diagrammatic representation* of a sentence is a configuration of objects that displays the spatial information contained in the sentence. The objects of the representation are reduced to the simplest geometrical shapes that yet enable us to understand and reason about this spatial information.

Based on this definition, two sentences that seem to convey different information might turn out to have a common diagrammatic representation underlying them, as is the case for Sentences 1 and 5. The content of the upper left frame in Figure 1.1 corresponds to their possible diagrammatic representation.

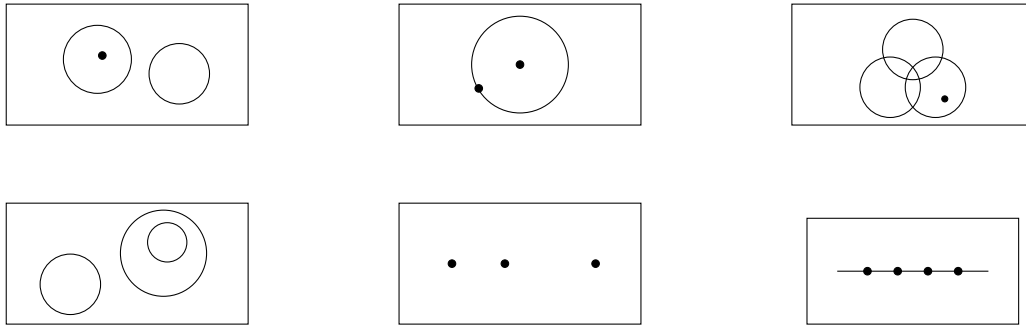


Fig. 1.1. Examples of diagrams discussed in the rest of the dissertation.

Only sentences for which the graphical representation corresponds to diagrams no more complex than those presented in Figure 1.1 (i.e. static two-dimensional representations that only contain *points*, *segments* and *circles*) will be used to study our capacity to reason spatially. The diagrams that can be constructed with these objects are sufficient to explain all the psychological principles of our approach, and furthermore, no additional principle is needed to extend it to more complex objects.

Finally, we use the term *imagery* for the capacity to create a mental representation of a situation that is not physically perceived (Kosslyn et al., 2006). For example, imagery enables us to think and talk about an apple without actually seeing any.¹ Imagery is often discussed in regard with perception, and we will see that, to a great extent, the two capacities share the same neurological systems in the brain (Pearson, 2019).

1.2 How Advanced is the Psychology of Spatial Relational Reasoning?

Although spatial relational inferences such as the ones triggered by Sentences 1-5 above are constantly made in cognition, they have rarely been the primary focus of psychological studies. Instead, these inferences have often been integrated indiscriminately into studies about reasoning *in general*. One of these theories in the middle of the twentieth century is *mental logic* (Braine and O'Brien, 1998). Its main assumption is that people reason by means of rules similar to those of logic, which clearly makes it a symbolic approach of the mind. However, several types of inferences quickly turned out to be difficult to explain simply by means of rules, and the need for a richer psychological approach soon became apparent.

The main alternative is the model-based approach, which postulates that people understand and reason about a situation by constructing a *model* of it mentally. Such a model is neither symbolic, nor a rich image, but rather a simplification of the situation that only conserves the aspects of it that are relevant for the inference -

¹ Even if the present work is mainly about visual imagery, a *mental image* is not necessarily visual, and imagery can also exist for other modalities (e.g. the taste of an apple, or the sound that a bite makes in it, are also defined as mental images).

such as a map does with the world. For this reason, mental models are considered *homomorphic*, or *structure-like*, to the situation they stand for. These ideas of models in the mind were already present in the work of the psychologist Craik ([Craik, 1943](#)), or even the philosopher Peirce ([Peirce, 1938, 4.6](#))², but it is only a few decades ago that Johnson-Laird developed them in the theory of Mental Models ([Johnson-Laird, 1983](#)). Presently, it is still in reference to this theory and its developments that studying our cognition from a model-based perspective is made.

The principles of the theory permitted to explain many interesting aspects of reasoning that could not be accounted by any other approach until then. However, the formal tools that it uses do not really differ from those of propositional logic: a model is represented by means of words on a line, each word corresponding to a piece of information. This formal choice shows some limits when it comes to capturing the semantics of sentences, especially considering the various range of inferences that the theory intends to cover: sentential, spatial, deontic, temporal, epistemic, causal, kinematic and many others ([Khemlani and Johnson-Laird, 2013, p.4](#)). This limitation is particularly noticeable in spatial reasoning, for which the claim that mental models are homomorphic clearly requires a different mode of presentation. Although more spatial techniques are adopted by the theory in these cases, they remain very limited, and the general principles of the approach do not apply easily. In consequence, the need for a specific development of the theory for spatial reasoning was evident.

This specific development has mainly been achieved through two complementary theories. The first one is the *space to reason theory* developed by Knauff ([Knauff, 2013](#)), which mainly investigates the nature of mental models in spatial reasoning. Based on neurological evidence that shows separated systems in the brain for the visual and the spatial process of information, the theory says that mental models are qualitative -

² We follow here the conventional notation for citing papers in Peirce's *Collected Works*: 4.6 refers to Volume 4, Paragraph 6.

without any exact measure - representations that only contain the spatial information relevant for the task. These mental models are called *spatial (layout) models* in the theory and are even postulated supporting comparative relations that are not spatial (e.g. "Mary is richer than John"). Despite the generality of these characterizations, only spatial models that have the structure of an array are used by Knauff, as he clearly announces it: "In the space to reason theory, I suggest defining a spatial layout model as a pattern of filled and unfilled cells in symbolic spatial array" (Knauff, 2013, p.37)

The second theory is the *preferred model theory* developed by Ragni and Knauff (Ragni and Knauff, 2013). Its main goal is to account for preferences in positional problems (e.g. "A is to the right of B", "B is below C" and so on), but it can also be seen as a practical application of some ideas of the space to reason theory. Indeed, an array is used to represent the spatial models and no visual detail is given to the objects contained in it.³

The theory has been implemented into a computational model called PRISM (i.e. Preferred Inferences in Reasoning with Spatial Mental Models), which accurately models our difficulties, errors and preferences in spatial reasoning for these positional problems.⁴ What enables PRISM to make more accurate predictions than other computational models (Ragni et al., 2021) is that it integrates some interesting principles highlighted in psychology, such as the limited capacity of our attention.⁵

³ By 'position', we understand relations such as 'left', 'right', but also 'South', 'East' and so on. This excludes topological relations such as 'inside', or size relations like 'larger'. Relations of distance, such as 'close to', 'far from', can partly be explained through the position of the objects in the array, as Ragni and Knauff show at the end of their article (Ragni and Knauff, 2013, 579).

⁴ The computational model in action can be seen at <https://www.modeltheory.org/portfolio/spatial-reasoning/>

⁵ The article is in fact a comparison of the "few theories [about spatial relational reasoning that] have been implemented as computational models" (Ragni et al., 2021, abstract). PRISM turns out having the highest performances for all the tests, which dispenses us from making an extensive literature review here.

1.3 Why is an Extension to All Kinds of Spatial Reasoning Called For?

Despite the general claims contained in the space to reason theory, only spatial models that have the structure of an array are studied in depth by Ragni and Knauff. Whether this limitation has a conceptual, or a more formal and computational origin remains open. Whatever the answer, the psychology of reasoning turns out to be generally reduced to these simple problems of positions, even in articles that intend to give an overview of the field ([Johnson-Laird et al., 2015](#); [Ragni et al., 2021](#)).⁶

The computational model PRISM is sufficient for illustrating the preferences in reasoning about positions, but all the principles it incorporates are limited to explaining this kind of inference only. Many other spatial inferences exist for plenty of objects more complex than cells in an array, such as circle inclusion, size or distance comparison, segment crossing and so on. In fact, it is most of our geometric capacities of reasoning that are excluded from this computational model. For example with the sentences presented at the beginning of this introduction, only our reasoning on Sentence 3 can be analyzed in PRISM. These last claims give a considerable support to the idea of extending the computational model to all kinds of spatial inferences.

Furthermore, it turns out that theories of other fields of cognitive science explain some fundamental capacities of cognition through our capacity to reason spatially. For example, the spatial relation of *inclusion* is postulated by the image schema theory as a fundamental relation in our comprehension of abstract linguistic expressions and mathematical concepts ([Lakoff and Johnson, 1980](#); [Geeraerts et al., 2005](#); [Lakoff and Núñez, 2000](#)). Therefore, a computational model that supports all kinds of spatial

⁶ It is of course possible to find other computational models for more than only positional relations in the literature ([Schultheis and Barkowsky, 2011](#); [Sima, 2014](#)), but in these cases, they do not allow to account for psychological effects in detail. These approaches are not even considered in ([Ragni et al., 2021](#)).

models would not only enable us to study spatial reasoning in a general way, but also to make all the other fields that attribute this fundamental role to spatial reasoning benefit from this contribution. The conceptual and computational possibilities of such a model are the topic of this dissertation.

1.4 Structure of the Dissertation

This dissertation is organized in four parts. Part 1 details the main principles of the theory of mental models (Chapter 2), before turning to the specifications of the space to reason theory (Chapter 3) and the preferred model theory (Chapter 4). The computational model PRISM is presented in this last chapter, and the motivations to extend it to spatial reasoning of all kinds are detailed at the end of it.

Part 2 investigates what the nature of spatial models might be when they are not restricted to an array (Chapter 5). Some well-known considerations in the literature about the inferential possibilities of representations are applied on mental representations in order to show that spatial models must, in fact, be of a diagrammatic nature. We call this argument, the *inferential argument*, and its conclusion, the *diagrammatic hypothesis*. A formalization of spatial models is proposed and elements from cognitive science are given to prevent the interpretation of the hypothesis as a claim for their imagistic nature. The hypothesis also enables us (Chapter 6) to develop a *representation system adapted for psychology*, which links the spatial models to the natural language sentences used to create them. Finally, some consequences of the hypothesis for the psychology of reasoning in general, such as the possibility to make deductive inferences, are discussed at the end of the chapter.

In order to create a computational model of spatial reasoning of all kinds, the representation system for psychology constructed in Part 2 must capture all the geo-

metric constraints that govern spatial models. It is demonstrated (Chapter 7) that the recent computational tool developed in the field of artificial intelligence, called *Answer Set Programming Modulo Theories for the Qualitative Spatial domain* ASPMT(QS), is perfectly adequate to implement all these geometric constraints. Combined with the psychological principles already present in PRISM, we build a novel computational model called EPRISM (i.e. 'E' for extended). The concrete functioning of EPRISM is sketched out (Chapter 8), and two examples are analyzed in detail: the first example is a typical problem used in PRISM, the second is a problem that involves topological relations between circles.

Finally, Part 4 highlights (Chapter 9) the fruitfulness that a computational model such as EPRISM represents for theories of cognitive science that attribute a fundamental role to our capacity to reason spatially. This is implemented in EPRISM by means of an additional linguistic module that translates sentences from natural language into formal objects and relations understandable by ASPMT(QS). The image schema theory receives a particular attention in this study. In the rest of this part (Chapter 10), the theory of mental models and the formal study of connectives are shown to benefit from these developments.

1.5 Preliminary Useful Remarks

1.5.1 *Formal versus Psychological Resolution*

The distinction between the *formal* resolution of a problem and the study of its *psychological* resolution is important. The formal resolution of a spatial problem aims to assess a conclusion by means of any logical systems in which the spatial information can be encoded and manipulated. In such a formal approach, the system is not

expected to process the information in accordance with the psychological principles known for humans (e.g. some tables containing all the possible combinations of spatial relations can be used to assess an inference, whereas it is clear that humans do not have such tables in mind).

The psychological resolution is made by the reasoner, who is subject to errors, cognitive limitations, and biases. The main purpose of the psychologist, based on experimental results, is to propose a theory that accounts for the psychological processes that lead the participant to their answer, regardless of its correctness. In this explanation, the formal resolution described above is only used to get the correct answer of the task that the reasoner is trying to solve. Consequently, this specific role of formal systems in psychological models of reasoning should never be taken as a psychological argument whatsoever. This last remark will turn out to be even more relevant for the case of ASPMT(QS) in EPRISM. Indeed, ASPMT(QS) is used in both types of resolution since it enables us to evaluate a problem formally, but also, it helps simulating the psychological resolution by generating the most probable spatial model constructed in the reasoner's mind.

1.5.2 *Homomorphic, Scientific and Logical Models*

The word *model* constitutes a key notion in this work. Several meanings exist for this word ([Hodges, 2009](#)), but all these meanings should be clearly separated from each other.

The most common meaning of the word 'model' in this dissertation is for 'mental model'. In this acceptation, a model corresponds to a mental representation that is *homomorphic* to a situation. This means that the configuration of elements in the mental representation is the same as the configuration of the elements in the situation,

such as a map is a model of the physical world.

This meaning should not be confused with the *scientific* use of the term, where a model represents a physical system or a natural phenomenon that is mathematized. In this acceptance, there is no need for the model to be homomorphic to the object it represents. For example, an equation models the behavior of waves against the coast, although it only consists of symbols. The computational models PRISM and EPRISM are models in this scientific sense since they model the human cognitive operations that enable us to reason spatially.

Finally, a more *logical* sense of the term model comes from Model Theory developed by Tarski. In this logical sense, a model is considered as a structure that satisfies all the formulas of a theory. Typically, the theory is in first-order logic and its models correspond to structures in which the non-logical symbols of the theory are interpreted and their relations represented set-theoretically.

Whether mental models should be considered as logical models or as homomorphic models from the first category is an interesting question that has not often been addressed in the literature. Moreover, in the few passages where the topic is addressed, Johnson-Laird remains ambiguous about the answer ([Johnson-Laird and Byrne, 1993](#), p.370). However, the lack of a clear formalism for his approach, as well as the characterization of models as *homomorphic* representations, are two facts that contributed to consider mental models rather as models of the first definition. The present work intends to bring some clarity regarding this question, although it will not resolve it completely.

Part I

Motivations

Overview

This part presents the main principles of the psychology of reasoning about spatial relations, and motivates the extension of the field to more than positional problems.

Chapter 2 starts by presenting the general principles of the model-based approach through the classical *theory of mental models* proposed by Johnson-Laird. The inadequacy of its formal tool is shown for the case of spatial reasoning.

Chapter 3 presents the *space to reason theory*, which develops the model-based approach for spatial reasoning and characterizes the nature of spatial models. Whenever it is possible, these characteristics are enhanced by more recent facts about cognition.

Finally, Chapter 4 presents the *preferred model theory* proposed by Ragni and Knauff, which mainly intends to account for our preferences in positional problems with multiple solutions. A computational model, called PRISM, implements the principles of the space to reason theory and the preferred model theory. After an illustration of PRISM by means of an example, several of its limitations are emphasized in order to motivate an extension to spatial relational problems of all kinds. Additional theories that would also benefit from such an extension are presented at the end of that chapter.

Chapter 2

The Theory of Mental Models

2.1 Facing Mental Logic

Early in the history of cognitive science, some authors postulated that people possess a kind of formal language in the mind. This hypothesis was mainly supported by the conceptual argument that only formal languages had the properties that were found in cognition ([Fodor, 1975](#)). For this reason, a similar language had to be present in the mind.

Within the psychology of reasoning, this hypothesis took mainly shape in a theory called Mental Logic ([Braine and O'Brien, 1998](#)). Its main assumption is that people reason based on formal rules similar to those of classical logic. But clearly, these mental rules cannot simply be those of propositional logic, or any other propositional approach. Propositional logic would fail to capture the spatial information of sentences as simple as "Circle A is inside Circle B, Circle B is disconnected from Circle C, so Circle A is disconnected from Circle C". However, it is precisely on this spatial information that people reason when they evaluate the inference. In order to save this logicist approach, one would need a more

expressive logic such as first-order logic. The rules thus would become similar to $\forall x, \forall y, \forall z ((\text{inside}(x, y) \wedge \text{disconnected}(y, z)) \rightarrow \text{disconnected}(x, z))$, and equivalences like $\text{disconnected}(x, y) \leftrightarrow \text{disconnected}(y, x)$ would be needed. As can be easily imagined, the number of such rules postulated in the mind would increase rapidly, and so would the formal derivations that are required to evaluate any simple sentences such as the one above. More complex geometric problems would undeniably make the hypothesis difficult to hold.

Opposed to this symbolic approach, the model-based approach consists in postulating *models* in people's mind that have the same structure as the situation described. These models differ from logical rules by conserving some aspects of the situation as a map does with the world. However, these mental models are not images either since they only conserve the information that is relevant for the inference. The defense of such an idea was already present in the work of the philosopher Peirce (Peirce, 1938, 4.6), or the psychologist Craik (Craik, 1943), but it is mainly through the *theory of mental models* proposed by Johnson-Laird that it is discussed nowadays (Johnson-Laird, 1983). The most general principles of this theory are presented in the next section.

2.2 General Principles of the Theory of Mental Models

The mental representations postulated in people's mind received the name of *mental models* in the theory (Johnson-Laird, 1983). Although the approach intends to account for many different kinds of reasoning (e.g. sentential, spatial, temporal, epistemic, kinematic and so on) (Khemlani and Johnson-Laird, 2013), it is often through linguistic problems that involve logical connectives such as *if*, *and* and *or*, that the main principles of the theory are generally presented.

2.2.1 Models are Conjunctions of Possibilities

A typical example of problems discussed by the theory is the following: suppose your lamp is not working anymore. You know that the fault might come "from the plug, the light bulb, or both". This situation allows three possible states of affairs, or psychologically speaking, three mental models, each of which being represented by a line below. We will refer to this format of presentation as the *one-model-per-line* format, and contrast it to the *homomorphic* format presented later in this section.

plug	
	bulb
plug	bulb

Now, suppose that someone checks the plug by connecting another device to it, and that device turns out to work. From this, you get the information that "the plug works correctly" and then you update your mental models with this new information. This corresponds to removing all the models (i.e. all the lines) that contain the plug. This results into only one model with the bulb, as presented below:

bulb

Since the bulb is part of all the resulting models (i.e. here only one), it can be concluded that it is the light bulb that causes the problem.

Not all the constructions of mental models are as easy as this one, and some problems might require more search from the reasoner to infer a valid conclusion. Several types of reasoning are then postulated in these cases, as presented in what follows.

2.2.2 Dual-system Processing and the Search for Counterexamples

For each logical connective, the theory predicts the number of models that will be constructed in people's mind. The theory also endorses a dual process of reasoning (Khemlani et al., 2018), which means that two types of inferential system are used to solve the problem. By means of a fast and intuitive type of reasoning (i.e. called System 1), a first set of models is produced by the reasoner. These models are called *mental models*, Table 2.1 (central column) presents them for each connective. The dots represent the possibility to add information or alternative models, which is done by means of a slower and more deliberate type of reasoning (i.e. called System 2). The additional models are called *alternative models* and the whole set formed is the set of *fully explicit models*, as visible in Table 2.1 (right column).

Table 2.1. The mental models constructed by Systems 1 and 2 for the different connectives (Khemlani et al., 2018, p.9).

Assertion	Mental models produced by System 1		Fully explicit models obtained by System 2	
If P then Q .	P	Q	P	Q
	...		$\neg P$	$\neg Q$
			$\neg P$	Q
If and only if P then Q .	P	Q	P	Q
	...		$\neg P$	$\neg Q$
P or Q or both.	P		P	$\neg Q$
		Q	$\neg P$	Q
	P	Q	P	Q
P or else Q but not both.	P		P	$\neg Q$
		Q	$\neg P$	Q

Based on this approach of dual reasoning, the theory is able to account for errors and difficulties in reasoning. For example, a sentence such as "if it rains, then the road will be wet", produces the first mental model based on System 1:

rain wet
...

However, by means of a more deliberate type of reasoning (i.e. System 2), a reasoner can construct alternative models that are also possible with the meaning of the 'if-then' connective (see Table 2.1). The fully explicit models are then:

rain	wet
\neg rain	\neg wet
\neg rain	wet

In the case where the piece of information "it is not raining" is obtained, can we conclude to the fact that "the road is not wet"? The search for the validity of a conclusion is an important psychological principle that the theory explains by means of the search of *counterexamples*. Authors claim that "The major principle of human reasoning is accordingly that inferences are good only if they have no counterexamples, that is, possibilities in which the premises hold, but the conclusion does not" (Johnson-Laird, 2020, p.219). In other words, the reasoner scans all the remaining models in order to check if the conclusion holds in all of them.

Following this process, it is easy to see why the problem with the lamp leads to a valid conclusion. Regarding the conditional example above, the two models that would remain with the information "it is not raining" (i.e. the second and third lines) do not allow to conclude that the "road is not wet". Indeed, the road is still wet in one of the models, which constitutes a counterexample to the conclusion. Reasoners that do not use the more deliberate reasoning of System 2 only construct the first model, and consequently, wrongly conclude to the validity of the argument.

As in general for logical reasoning, note the asymmetry between the requirement of finding *only one* counterexample to a conclusion in order to show the invalidity of the inference, compared to the requirement of checking *all* the possible models to prove its validity. This asymmetry also holds for reasoning with mental models, as recalled in (Johnson-Laird and Byrne, 1993, p.338): "an argument is invalid if there is

at least one counterexample. By contrast, an argument is valid only if every model of the premises is a model of the conclusion".

2.3 How to Characterize Mental Models?

One of the most important principles of the theory is the *integration* principle, which claims that people do not keep the information of the premises separately in reasoning, but combine them into a single mental model as much as possible (Johnson-Laird, 2008, p.124). Therefore, each line of the previous examples has to be understood as a single mental representation in people's mind. Johnson-Laird clearly confirms this fact by saying that "Of course, people don't use words and phrases to represent possibilities, but actual models of the world akin to those that the perceptual system constructs" (Johnson-Laird, 2020, p.219).

This alternative to symbolic approaches is appealing, but it consequently opens an important question: what is the nature of mental models exactly? Many discussions arose in the literature to answer this question. Over the years, the theory has brought some elements of answers by means of several principles. Some of the most important are the following (Johnson-Laird et al., 2015):¹

- Mental models are *homomorphic*, which means that they have a structure similar to the situation they represent, such as a map does with the world.
- Mental models are *possibilities*, which means that no mental model of an incoherent situation can be constructed. A 'square-circle' is for example an impossible model.

¹ The list of all the principles can be found on the official website of the theory, at <https://www.modeltheory.org/about/what-are-mental-models/>

- Mental models are neither propositions, nor concrete images, but rather similar to *diagrams* in the sense that they only preserve the relevant aspects of the situation to make the reasoning possible.

These characteristics of mental models are supposed to make the theory more semantic than a logicist approach. However, it is sometimes hard to know how they should be applied to the mental models constructed for some sentences. For the examples of the lamp or the rain given above, these principles do not really make sense, especially the homomorphic principle. Furthermore, since mental models are represented in this one-model-per-line format that does not differ from classical logic, some authors came to the conclusion that "Johnson-Laird's theory is *semantic* in the same sense as truth-tables are semantic" ([Osta Vélez, 2020](#), p.28).

There exists, however, a type of reasoning for which the principles make more sense, namely, spatial reasoning. For such inferences, it is even common to see the authors putting aside the one-model-per-line format as presented above, in favor of a representation that is indeed homomorphic to the spatial information conveyed. For example, a spatial problem considered by the theory is the following ([Johnson-Laird, 2020](#), p.224):

The cup is to the right of the plate.
The spoon is to the left of the plate.
The knife is in front of the spoon.
The saucer is in front of the cup.

What is the relation between the knife and the saucer?

The mental model constructed for such a problem is displayed by the two-dimensional representation below:

spoon	plate	cup
knife		saucer

As can be seen, the structure of the model represented is homomorphic to the spatial information conveyed. This enables a reasoner to actually *see* the conclusion in the model constructed whereas such an inference would not have been possible in a one-model-per-line format.

Despite the fruitfulness of such a format of representation, no formalism is proposed by the authors of the theory to generalize their studies. Some systems have been developed, but always specifically for the kind of spatial inference under consideration in the task and without holding for spatial reasoning in general (Johnson-Laird, 2006).² The authors even conclude that "reasoning about the domain of two-dimensional spatial relations, such as *The cup is on the left of the saucer*, is intractable for all but the simplest deductions" (Johnson-Laird, 2020, p.218). Regarding the psychological processes used in these spatial inferences, they also require a specific study different from the one for connectives, but this is not really undertaken by the theory.

In conclusion, a tension occurs between the aim of the theory to take into account the semantics of spatial problems in accordance with the principle put forward, and its inability to process this spatial information as well as the psychological inferences made with it. All these difficulties converge towards the need for a specific version of the theory of mental models for spatial reasoning. This specific study was mainly carried out by Knauff, and resulted in the *space to reason theory* a few years ago (Knauff, 2013). This theory is the object of the next chapter.

² A method that has often been used by the authors is to declare a binary relation by a constraint such as ((100) *spoon plate*), where the digits correspond to a placement in a coordinate system. Deductions on the locations of other objects are then made by means of calculations on these digits (Johnson-Laird, 2006, p.125).

Chapter 3

The Space to Reason Theory

Several goals can be seen in the space to reason theory. A first one is of course to lay out the psychological principles that govern the way people reason spatially. A second and more general goal is to propose a new characterization regarding the nature of mental models. Knauff sees this characterization as an element of resolution in a debate that lasted for decades about the role of images in reasoning. On the one hand, the pictorialists argued that reasoning is helped whenever the problem easily suggests an image in the mind ([Shaver et al., 1974](#)). On the other hand, the authors of the theory of mental models themselves postulated that people reason by creating mental models that are homomorphic to the situation considered, without being images at the same time, as already explained above ([Johnson-Laird, 1998](#)). Within the view endorsed by the authors of the theory of mental models, images in the mind are only epiphenomenal. It turns out that both sides found results corroborating their own hypothesis in a large amount of experiments.

By means of three assumptions, the space to reason theory has brought some clarity into the debate:

- (i) the *visual* and the *spatial* contents of premises have to be distinguished in the brain, and consequently in experiments.
- (ii) the visual content (when it is not necessary for the task) impedes reasoning, whereas the spatial content plays an important role for reasoning in general.
- (iii) mental representations that support reasoning are *spatial layout models* of a qualitative nature.

These assumptions are supposed to better characterize mental models, and consequently, to help accounting for the way we reason with them. Each of the three assumptions is presented in one of the following sections. Whenever they exist, additional arguments from recent studies enhance the cognitive facts already put forward by Knauff.

3.1 Spatial Information in the Brain

3.1.1 The Two-Streams Hypothesis

For the first assumption that visual and spatial information should be distinguished in mental representations, Knauff relied on the two-streams hypothesis primarily highlighted in the field of neurology ([Mishkin et al., 1983](#)). To understand this hypothesis, it is useful to know that the brain is, roughly speaking, divided into four different lobes as represented in Figure 3.1 (a).¹

¹ As most commonly done to represent brain images, the forehead is towards the left, and the back of the head towards the right. The two brain images of Figure 3.1 are thus two left hemispheres. Moreover, representing only one hemisphere generally suffices to explain a neurological process since cognitive functions are, most of the time, symmetrically distributed between the two hemispheres. This is considered to be the case for the functions discussed here, at least for the level of detail taken for the present explanation.

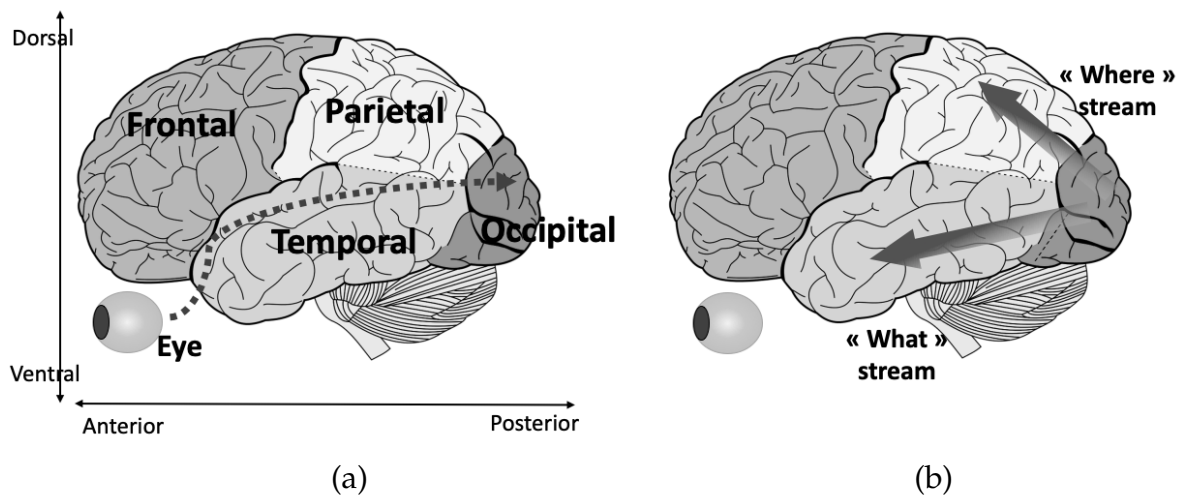


Fig. 3.1. Pictures of two left hemispheres. In (a), the visual stimulus goes from the retina into the occipital lobe at the back of the brain. In (b), the visual information in the occipital lobe exits it by taking a dorsal path (i.e. the "where" stream) that processes the spatial information, and a ventral path (i.e. the "what" stream) that processes the information of the objects. (Figure created with a public domain image.)

Studies about perception tell us that the visual information that enters the retina is then brought into the occipital lobe, more specifically into a region called V1, at the back of the brain. This transport from the retina is represented by the dotted lines in Figure 3.1 (a). The visual information in the occipital lobe yet requires some further treatments, which is done by other parts of the brain. For more than three decades now, the *two-streams hypothesis* has claimed that the treatment of this visual information is mainly divided into two processes ([Mishkin et al., 1983](#)). A ventral path, called the "what" path, goes into the temporal lobe and processes the objectual information, whereas a dorsal path, called the "where" stream, goes into the parietal lobe and processes the spatial information (Figure 3.1(b)).

The objectual information processed by the "what" stream concerns the aspect, details, color and shape of the objects of the scene. It is mainly involved in tasks such as object recognition and form representation. On the other hand, the spatial information processed by the "where" stream concerns the location, orientation and

size of the objects. The movements of dynamic scenes are also processed by this stream.

Before giving some pieces of evidence for the reality of these two streams in the brain, one can argue that the hypothesis only concerns perception, and might not hold for a situation that is only described linguistically. The following facts intend to prevent this argument from being made.

3.1.2 Imagery: Recreating the World in Our Mind

Imagery is often defined comparatively to *perception*. Perception denotes the cognitive systems that process and give meaning to the information coming from the world and brought into our brain by means of our senses (i.e. vision, audition, and so on). Imagery is the pendant of perception since it corresponds to our mental life that is not (directly) created by external stimuli.

Interestingly, it is often shown that imagery is made possible by our mental ability to create a content *as if it were physically perceived* (Kosslyn et al., 2006; Koenig-Robert and Pearson, 2021). This means that thinking about, for example an apple, will be made possible by creating a mental object similar to an apple, and then innerly looking at this object as if it were an actual apple in front of our eyes. Language is naturally considered as the principal medium by which mental imagery is activated.

This important fact about our capacity to recreate the world internally permits a great economy on the neurological level, that is, the neurological processes used in perception can be similarly used in imagery. This fact is now confirmed by a large amount of experiments and brain scans (Kosslyn et al., 2006; Pearson, 2019). Consequently, neurological principles that are true for perception are also generally true for imagery, as has been proved for the two-streams hypothesis (Pearson, 2019).²

² Interestingly, this study also shows a stronger overlap of the neurological systems in case of

3.1.3 Cognitive Evidence for the Two-Streams

The two-streams hypothesis has received support from many fields of cognitive science and is largely accepted nowadays. Results in neuropsychology, for example, show that people who present damage to one of the two streams do not necessarily have deficits in the functions associated with the other stream ([Levine et al., 1985](#); [Farah et al., 1988](#)).

From a psychological perspective, the experiments presented in the space to reason theory also constitute support for the hypothesis. By means of a dual-task paradigm, Knauff and his colleagues showed that a spatial task conflicts with a secondary spatial task but not with a secondary visual task, and vice versa ([Knauff, 2013](#), p.67). The same result was also obtained by Klauer and Zhao who showed a double dissociation of the two cognitive processes ([Klauer and Zhao, 2004](#)).

Finally, a more recent source of evidence for the two streams hypothesis can be found in the study of people with aphantasia ([Pearson, 2019](#)). This cognitive characteristic prevents the construction of visual mental images in the mind. For example, people with aphantasia cannot picture the image of an apple, or themselves lying on a beach. However, this inability to form specific images in their mind does not prevent them for reasoning spatially. Up to now, it has remained unclear whether aphantasia is a deficiency of the "what" path, or if the visual information is processed but cannot access consciousness.

voluntary imagery.

3.2 Spatial Models Support Reasoning

Based on the first assumption about the two-streams hypothesis, Knauff makes the second assumption that only the spatial system used in the "where" path plays a fundamental role in reasoning, whereas the content of the visual system used in the "what" path can even impede the reasoning if it does not participate in the resolution of the task. This claim is supported by experimental results for relational problems with different degrees of imageability (Knauff, 2013, p.47). In these studies, participants receive a relational description (e.g. "Ann is taller than Bill, Bill is taller than Carla") and are asked to find the implicit relation (e.g. "Ann is taller than Carla"). Interestingly, the relations in the problems vary in their degree of spatiality and imageability (e.g. 'above-below' is highly spatial, 'cleaner-dirtier' is highly visual). Results show that problems with a high degree of spatiality are solved faster than highly visual ones. It is assumed that additional time in visual problems is spent for the creation of images in the mind, although they do not help to solve the problem. A similar experiment with blind people (Knauff and May, 2006) showed the results expected, that is, blind people do not present the impedance effect in the visual problems since they do not spend time creating any mental images.

The important other part of Knauff's assumption is that the spatial system supports reasoning *in general*. This fact is clearly expressed in the following quotation:

[...] visual brain areas are only involved if the problem information is easy to visualize. Reasoning in general, however, does not involve the construction of visual images. Instead it involves more abstract spatial representations held in parietal cortex. Only these spatial layout models are crucial for the genuine reasoning process. (Knauff, 2013, p.xiii)

To justify this assumption, Knauff relies on experiments where the authors showed that the parietal areas dedicated to spatial reasoning were activated in all comparative inferences, regardless of the kind of relation (i.e. 'taller', 'bigger', 'richer', 'smarter'

all create activation in the parietal lobe), (Ruff et al., 2003). This fact gives support to the idea that mental representations of a spatial nature are used even for comparison of abstract relations. Recently, this idea has been enhanced by additional neurological evidence (Alfred et al., 2020).

3.3 The Nature of Spatial Models

The first assumption of the space to reason theory separates the visual from the spatial information, and the second assumption postulates that mental representations purely spatial in the parietal lobe support reasoning in general. In the third assumption, Knauff intends to define the nature of these spatial mental representations located in the parietal lobe.

3.3.1 We Reason Based on Qualitative Spatial Layout Models

According to Knauff, the mental representations used in reasoning are neither completely imagistic, nor completely abstract, but rather of a *qualitative* nature, which means that they contain no exact value.³ In line with the model-based approach, Knauff calls these mental representations *Spatial Layout Models*, or more simply *Spatial Models* (Knauff, 2019). He describes them as follows:

I argue against the hypothesis that human reasoning is completely embedded in language and relies on formal rules of inference akin to those of formal logic. Instead I argue that reasoning requires going beyond language and is based on the construction, inspection, and variation of spatially organized mental representations. These representations are inherently spatial, more concrete than words, but more abstract than visual images. I refer to these spatial representations as *Spatial Layout Models*. (Knauff, 2013, p.xi)

³ Knauff even considers these mental representations to be *amodal*, or *supramodal*. This claim is not directly discussed in this dissertation, although the arguments of the second part might be connected with it.

Some structures very similar to the spatial models described by Knauff have already been highlighted in the literature under the same name (Tversky, 1993), or different ones.⁴ As can be seen in the next quotation, Knauff also characterizes these spatial models as mental representations that only conserve the parts of the situation that are relevant for reasoning:

Following the principle of representational parsimony, my conjecture is that on the algorithmic level, spatial layout models are purely qualitative representations whose parts contain only the relations sufficient to permit particular calculations or operations. [...] the term qualitative representation originated in computer science and refers to configurations among distinct spatial entities that preserve location in space without incorporating information such as shape, size, texture, or color of objects. (Knauff, 2013, p.37)

Finally, another description has also been given by Knauff more recently, and permits to understand how a concrete image differs from a spatial model:

Visual mental images are concrete. They represent, color, shape, texture, size, etc. They also do not allow to represent indeterminacies, ambiguities, or different interpretations of the given information. You cannot imagine that the flower is red or yellow. If you want to do that you have to imagine two flowers. Spatial mental models are more abstract. They represent just the information pertinent to the inference. You can imagine that one flower is to the left of another flower without accounting for the exact visual features of the flowers or the exact distance between the flowers. (Knauff, 2019, p.8)

Knauff considers a subway map as a good example of spatial model (Knauff, 2013,

⁴ For instance, Kosslyn calls such mental representations *object maps* and describes them as follows: "If a spatial image is all that is required by the task, stored information is unpacked to create an *object map* in the spatial-properties-processing subsystem (in the posterior parietal lobes). The object map specifies the relative locations of objects, or parts of a single object, but does not indicate depictive properties of shape" (Kosslyn et al., 2006, p.143). A direct comparison of *object maps* and *spatial layout models* is made in (Sima et al., 2013). Amorapanth and his colleagues also postulate such spatial structures that they call a *schema*: "The term schema is used in this paper as any kind of representation (external or cognitive) where perceptual detail has been abstracted away from a complex scene or event while preserving critical aspects of its analogous qualities. Schemas, as such, occupy a representational middle-ground: more abstract than very concrete representations of objects, but unlike truly symbolic representations (e.g., words) a schema preserves some of the spatial-relational aspects of the thing it stands in for. The most critical aspect of schemas, as the term will be employed in the present paper, is that they occupy an intermediate position between abstract words and concrete percepts in a graded model of representation". (Amorapanth et al., 2012, p.143).

p.192). The stations do not have their real size and the distances are not conserved between the stations, but the topological relations are preserved.

3.3.2 Spatial Models Reduced to the Structure of an Array

Despite the characterizations provided by the third assumption, Knauff suggests to represent spatial models by means of an array: "In the space to reason theory, I suggest defining a spatial layout model as a pattern of filled and unfilled cells in symbolic spatial array" (Knauff, 2013, p.37). This formal characterization suffices to describe the spatial models and reasoning in positional problems such as "A is to the left of B", "C is above B" and so on.

In this respect, Knauff and his colleagues studied in depth the psychology of reasoning for such problems, and especially the preferences of participants in case of several possible models. The theory that results from this research is called the *preferred model theory* and is the object of the following chapter. A computational model, called PRISM (i.e. Preferred Inferences in Reasoning with Spatial Mental Models), is also proposed for the theory. It is presented at the end of the next chapter.

Chapter 4

The Preferred Model Theory and PRISM

The preferred model theory and its computational model PRISM result from Ragni and Knauff's work. The main goals of the preferred model theory are to explain how people reason about positional problems (e.g. "A is to the left of B, C is below A" and so on), and to highlight their preferences in case of problems with multiple solutions. The computational model PRISM simulates our reasoning for these problems based on the results from experimental research. This computational model can also be seen as a practical realization of the principles formulated in the space to reason theory since spatial models only have the structure of an array, and objects are not represented with their aspectual features in it.

One advantage of not representing detailed object information is that models in PRISM are not confused with visual images. We have argued for a careful distinction between spatial models and pictorial mental images in many other publications [i.e. the space to reason theory]. ([Ragni and Knauff, 2013](#), p.565)

Almost all the examples discussed in the preferred model theory are similar to Problem 6 presented on the left of Figure 4.1. In formal terms, it means that all the

problems have the logical form $\phi \wedge \dots \wedge \psi \models \chi$. On the right of Figure 4.1 is the array used by the authors to represent the spatial model postulated in people's mind. The cars used in the problem's description are simply represented by their names within the cells of the array.

- (6) The Porsche is to the right of the Ferrari.
 The Beetle is to the left of the Ferrari.
 The Dodge is in front of the Beetle.
 The Volvo is in front of the Ferrari.

Where is the Volvo compared to the Dodge?

	Beetle	Ferrari	Porsche	
	Dodge	Volvo		

Fig. 4.1. On the left, a typical problem discussed in the preferred model theory. On the right, the representation of the spatial model constructed for this problem as proposed by Ragni and Knauff ([Ragni and Knauff, 2013](#), p.154).

Many general principles from the theory of mental models are included in the preferred model theory, but often with modifications specific to spatial reasoning. However, one principle that remains exactly the same is the *integration* principle that claims that the information of the premises is combined into a single model ([Johnson-Laird, 2008](#), p.124). It is the reason why all the premises are combined into a single array, as can be seen in Figure 4.1.

Regarding the other principles from the theory of mental models, their specific adaptations for spatial reasoning are presented in the next sections.

4.1 The Main Principles of the Preferred Model Theory

4.1.1 What Are The Alternative Models?

A major difference with the theory of mental models is the definition of *alternative* models. In the theory of mental models, alternative models are obtained by negating some clauses of the problem according to the semantics of logical connectives in the problem (i.e. *and*, *or*, *if*, etc). Since only problems in the format of conjunctive information are discussed in the preferred model theory (i.e. ϕ *and* ..., *and* ψ), no alternative model resulting from the negation of clauses is created by participants. Indeed, using just the conjunctive connective only creates a single model, and in a classical logic view, this model is always the first line of the truth-table.

Instead, alternative models in the preferred model theory occur in problems where the information is not sufficient to construct only one configuration of objects in the array. It follows that alternative models in the preferred model theory are all *contained* within the same valuation of the problem, which is this first line of the truth-table where all the pieces of information are considered true.

A problem where only one spatial configuration can be constructed is called a *determinate* problem, whereas it is called an *indeterminate* problem if more than one configuration is possible. For example, the problem "the circle is to the left of the square, and the square is to the left of the triangle" is determinate since only one model can be constructed for it. On the contrary, the problem "the circle is to the left of the square, and the circle is to the left of the triangle" is an indeterminate problem. Indeed in this last case, the triangle can be placed to the left of the square, but also to the right. Finally, it is worth mentioning that PRISM does not allow two objects to be at the same place, which precludes the possibility of a third model.

4.1.2 The Preferred Model and the Principle of Cognitive Economy

A fundamental contribution of the theory is to predict preferences in case of indeterminate problems, in other words, what is the first model constructed among the possible ones. This model is called the *preferred model* in the theory.

Importantly, this preferred model is supposed to be the same for every participant, and even through different cultures (Ragni and Knauff, 2011). The rationale of these preferences is based on a fundamental principle of psychology: every cognitive operation is performed at the lowest cognitive cost as possible, that is, the one that requires the least amount of cognitive resources (Ragni and Knauff, 2013, p.563). Although the authors have not named this principle, we will refer to it as the *cognitive economy* principle.

Based on several experimental results, the authors show that people prefer to place a new object of the premises without moving any other objects previously placed. This strategy is called the *fff-strategy* (first free fit), and is opposed to the *ff-strategy* (first fit), where new objects are inserted between old ones.¹ According to this claim, the sentence "The circle is to the left of the square, and the circle is to the left of the triangle" creates the model where the square is to the left of the triangle.

4.1.3 No Spontaneous Search for Alternative Models

Consequently to the discrepancies in the definition of alternative models, the preferred model theory differs from the theory of mental models regarding the search for alternative models. Whereas the theory of mental models endorses a dual-system approach of reasoning (i.e. System 1 and System 2), the preferred model theory does

¹ Regarding the principle that preferences are calculated according to the cognitive cost of construction, the *fff-strategy* is cognitively less demanding since it does not require the additional operation of moving an already placed object.

not predict a spontaneous search for alternatives by means of System 2. This is clearly announced by Knauff:

My preference theory offers a different explanation [than the theory of mental models]: there is a single system in which a preferred model is constructed and inspected. Usually people end at this model, because this is the model that in many cases suffices to draw a valid conclusion. The single, preferred model does not overload working memory and is feasible in most situations. Also, people do not usually search for counterexamples to validate a conclusion. (Knauff, 2013, p.221)

Although indeterminate problems are one of the main focuses of the preferred model theory - especially for studying what is the first model constructed by people - they will not be discussed at a psychological level here. Therefore, the presentation of PRISM below and the examples analyzed in EPRISM afterwards are always determinate problems, that is, only one model is possible. Nonetheless, alternative models will play an important role on a theoretical level, which is the reason why they could not be ignored in the presentation of the theory.

4.2 PRISM: The Computational Model

The computational model PRISM (i.e. Preferred Inferences in Reasoning with Spatial Mental Models) is used for modeling spatial reasoning in accordance with the assumptions of the preferred model theory (Ragni and Knauff, 2013). A screenshot of the model during the resolution of a problem is visible in Figure 4.2.²

² An animation of PRISM can be watched at <https://www.modeltheory.org/portfolio/spatial-reasoning/> and its official website is <http://spatialmentalmodels.appspot.com/>.

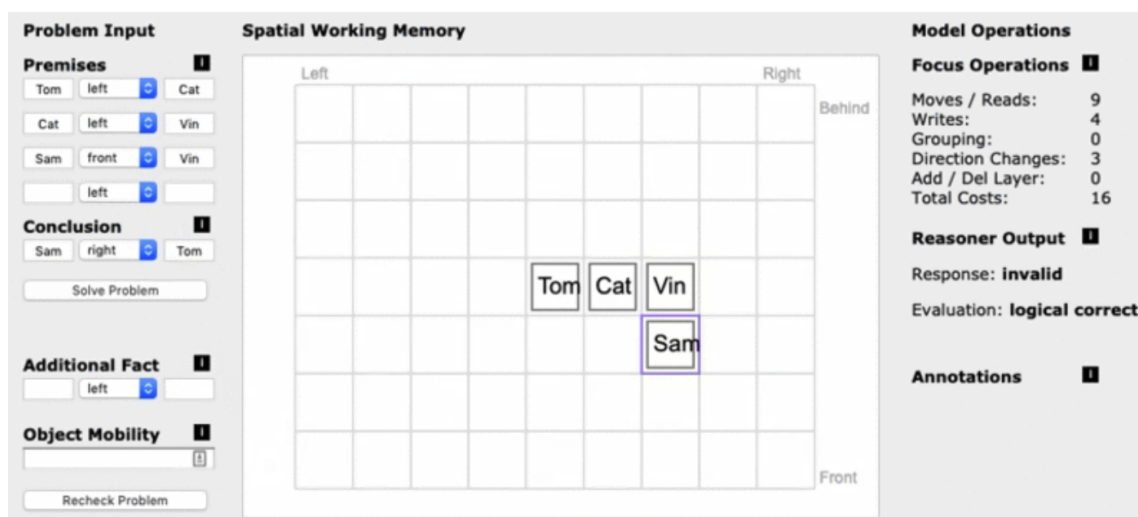


Fig. 4.2. Screenshot of the computational model PRISM while solving the spatial problem described on the left side. On the right side, the cognitive resources used for the resolution of the problem and the psychological and formal predictions for its evaluation. [source: <https://www.modeltheory.org/portfolio/spatial-reasoning/>]

The array in the center of Figure 4.2 represents the spatial working memory of the reasoner in which the spatial model is constructed.³ The premises are presented to the left of the array. In Figure 4.2, these premises correspond to the assertions "Tom is to the left of Cat", "Cat is to the left of Vin", and "Sam is in front of Vin". The conclusion to assess, which is "Is Sam to the right of Tom?", can be seen below the premises.

4.2.1 Resolution Through Reasoning Phases

Similarly to the theory of mental models, the psychological resolution of problems is supposed to be done by means of three successive phases of reasoning: (i) the model *construction* phase that creates the preferred model, (ii) the model *inspection* phase in which this model is inspected, and (iii) the model *variation* phase where alternative

³ Ragni and Knauff based their theory on Baddeley's model of working memory (Repovš and Baddeley, 2006).

models are searched. Since only determinate problems are discussed in detail here, only the construction and inspection phase need to be described.

During the model *construction* phase, the participant understands the information conveyed by the premises and constructs the spatial model according to it. Because premises are read sequentially, the construction of the model is realized by inserting the objects of one premise after another. As will be detailed below, attentional limitations force us to process each premise through several steps. When all the premises have been processed, a spatial model such as the one in the spatial working memory of Figure 4.2 is constructed.

The model *inspection* phase starts by reading the conclusion that has to be checked in the model, which is, in our example: "Is Sam to the right of Tom?". This checking operation mainly consists in scanning the spatial model in order to verify whether the two objects of the conclusion are indeed in the relation stated by it. As can be seen for the specific example of Figure 4.2 (right side), the conclusion is predicted invalid because 'Tom' is not wholly to the left of 'Sam' in the array.⁴ Since the problem is a determinate one, no alternative model is searched during a variation phase and the reasoning stops there.

The psychological operations in these reasoning phases (i.e. the insertion of objects and the search for objects in the model) are carried out by an essential component of PRISM that remains to be presented, namely, the *attentional focus*.

4.2.2 The Attentional Focus

The attentional focus is an important component that enables PRISM to make better predictions than other computational systems in the literature (Ragni et al., 2021). It

⁴ More will be said about this evaluation while discussing PRISM's limitations.

can be considered as the reasoner's attention towards the objects and places of the spatial model. In the snapshot of Figure 4.2, the focus corresponds to the blue square that frames Sam's cell.

During the construction phase, attentional shifts among the different cells are made in order to place the objects in accordance with the relations stated in the premises. Note that from a cell of the array, a shift can only reach an adjacent cell. Since one object of each premise will already be placed in the array when the premise is processed, an important role of attention is to find this already placed object. When it is found, the second object of the premise is inserted according to the relation stated by it. The inspection phase for checking the validity of a conclusion is quite similar, except that no object is inserted.⁵

Each operation of the focus is counted on the right side of Figure 4.2. The total of the operations corresponds to the cognitive cost required by the problem. Using an attentional focus enables the authors to account naturally for the different effects that are caused by a problem, such as the continuity effect (Ragni et al., 2021).⁶

Restricting attention to focusing only on one object at a time in the spatial model is a judicious choice for which a large amount of evidence has brought support in the literature, both in the study of perception and imagery (Franconeri et al., 2012; Laeng et al., 2014). For example, this limited capacity of attention is clearly visible in an experiment where people are reading a text while a complex device changes all the words that are not currently read by the participant into a series of 'X' symbols. Beside the fact that the participants are not bothered in their reading, they remain blind to the modification (Slattery et al., 2011).

An important consequence of this limitation is that a relation cannot be perceived

⁵ All these operations can be seen on the online animation at <https://www.modeltheory.org/portfolio/spatial-reasoning/>.

⁶ A *continuous* problem (e.g. "A is to the left of B, and B is to the left of C") is more easily solved than a non-transitive problem (e.g. "A is to the left of B, and C is to the right of B").

per se, but is only obtained by means of attentional shifts between its elements. In other words, it is the attentional shift itself that enables us to extract a piece of information from the spatial model.⁷ It follows that, for instance, the information "Object A is to the left of Object B" can be obtained by a shift from A to B, but that an additional shift from B to A will be needed to get the converse information "B is to the right of A". This important principle is correctly instantiated in PRISM.

4.3 Motivations for a Computational Model for All Kinds of Relations

4.3.1 Limits Imposed by PRISM

On the one hand, PRISM is one of the computational models of spatial reasoning that best instantiates some general psychological principles, such as the homomorphism of spatial models, or the consequences of attentional limits. On the other hand, it constraints in a very drastic way the infinite possibilities of spatial models, and consequently of spatial inferences that can be analyzed.

The major limit imposed by PRISM is that spatial models can only have the structure of points placed in an array. It means that any spatial model involving other objects such as segments or circles, and other relations such as 'larger' or 'inside', cannot be accounted by it.

This limit goes hand in hand with the other limits that PRISM "has no link to a mental lexicon in which, for instance, the semantics of spatial prepositions are represented" (Ragni and Knauff, 2013, p.583). Therefore, only objects that are perfectly positioned horizontally or vertically compared to other objects are considered by

⁷ This *sequential* extraction of information has been confirmed experimentally (Franconeri et al., 2012).

the system, and any situation with a slight deviation is not supported. To a greater extent, it also means that PRISM cannot take into account diagonal relations either, as the authors acknowledge it, along with other relations: "PRISM struggles with inferences in the diagonal of the spatial array (reasoning with *diagonally behind* is not possible) and also cannot handle the spatial relation *in the same place*" (Ragni and Knauff, 2013, p.584). Knauff and Ragni accept these limits by saying that "In the real world, solid objects cannot usually occupy the same place at the same time. Liquids or container-like objects might be exceptions, but to dealing with such entities would make PRISM unjustifiably complex and would have too far-reaching consequences for our model construction process" (Ragni and Knauff, 2013, p.584).

4.3.2 Why an Extension?

These limits can be ignored for the main goals of the preferred model theory. However, it seems also promising to extend such a model to any kinds of spatial relational inferences. Two main motivations support this extension.

The first is to account for a broader part of our spatial reasoning capacities, such as those used in geometry. Indeed, few approaches in the literature seem able to propose a computational model general enough to simulate geometric reasoning, although it is clearly considered as an important capacity of cognition. Some very commonly used relations, such as circle inclusion or size comparison, also motivate the creation of such a computational model. Authors in the literature clearly show their support for extending the research area of the psychology of spatial reasoning.

Previous studies have almost exclusively focused on discrete positional relations (e.g., to the left of, to the right of, above, below, behind, in front of, etc.) or on one-dimensional topological relations, but they have not typically studied topological relations involving regions and their boundaries (such as being inside, outside, or intersecting) that are yet fundamental to our spatial understanding. (Hamami et al., 2021, p.2).

Therefore, this extension would enable the psychology of spatial reasoning to broaden its activities to more than positional problems, and furthermore, to focus on these relational inferences that are known to play an important role in our spatial capacities.

The second motivation for this extension comes from the sum of other fields from cognitive science that are not directly linked with the psychology of reasoning, but which turn out to postulate spatial structures very similar to spatial models in order to account for cognitive abilities. Some of these theories coming from the field of neurology have already been mentioned with the presentation of Knauff's definition of spatial layout models ([Kosslyn et al., 2006](#); [Amorapanth et al., 2012](#)), but it is mainly when we turn to theories of cognitive linguistics that these structures postulated in the mind become numerous. In particular, the image schema theory ([Gibbs and Colston, 1995](#)) postulates such structures in order to account for cognitive capacities as various as linguistic comprehension of concrete and abstract expressions, and concept acquisition in mathematical knowledge. The inferences about the topology of circles constitute a good example for supporting this claim, considering the number of inferences we supposed permitted by them: in geometry and topology, but also in syllogistic reasoning with Euler diagrams, mathematical logic with the set theory and so on ([Lakoff and Núñez, 2000](#)). Several signs for the need of a formal framework adapted to these structures can be found in the literature ([Suchan et al., 2015](#); [Hedblom, 2020](#)). In conclusion, this second goal may be less direct than the one for the psychology of spatial reasoning, but is clearly more general for the field of cognitive science.

Considering the large scope of applications that such an extension seems to offer to cognitive science, it might be surprising that no framework has been proposed to fulfill these functions so far. We believe that Knauff's characterization of spatial models does not really permit this extension to happen. It is clear that the space

to reason theory brings many interesting elements to describe the nature of spatial models, but it does not enable us to understand what spatial models are precisely, and how we can formalize them when they have a structure more complex than an array.

The next part mainly investigates what spatial models are in spatial reasoning of all kinds. What we call the *inferential argument*, uses concepts from representation systems' studies to argue that spatial models must have a diagrammatic nature in order to be in line with some essential principles of the theories presented in this first part. In particular, two characteristics given by Knauff to spatial models will be shown incoherent. These clarifications will also open the possibility to consider spatial models from a formal perspective.

Part II

Spatial Models as Diagrammatic Representations

Overview

This part mainly intends to clarify the nature of spatial models when they are considered in all kinds of spatial inferences.

Chapter 5 starts by presenting what we call the *inferential argument*, which basically says that the inferences generally considered to be permitted by spatial models correspond to the *cognitive potentials* attributed to diagrammatic representations in the literature about representation systems. This inferential argument enables us to state the *diagrammatic hypothesis* that considers spatial models as diagrammatic representations. A formalization of spatial models is proposed accordingly. Finally, several facts are given to narrow down the interpretation of the hypothesis, especially regarding the idea that a spatial model cannot be considered as an image in the mind.

Chapter 6 discusses the formal and psychological consequences of an important implication of the diagrammatic hypothesis: spatial models over-specify the information they stand for. This property of OVER-SPECIFICITY is defined through the construction of a *representation system adapted for psychology*. Finally, this chapter ends by discussing the consequences of this property for deductive reasoning in general.

Two characteristics given by Knauff, namely that spatial models are "qualitative representations", and that they "contain only the relations sufficient to permit particular calculations" (Knauff, 2013, p.37), are clarified along these two chapters.

Chapter 5

The Diagrammatic Hypothesis

The computational model PRISM, and for its main part, the preferred model theory, only deal with positional problems. The array used as a framework is sufficient for inferences restricted to the positional aspect of relations. However, it is clear that many other spatial aspects can be used to study spatial inferences. For example, Sentence 7 presents a problem that also involves spatial relations, but addresses the topological aspect between the objects.

(7) Circle A is inside Circle B, and Circle C is disconnected from Circle B.

Is Circle A disconnected from Circle C?

Few psychological studies directly tackle our reasoning capacities for such a spatial aspect, and consequently, few theories describe the spatial models that underlie this kind of inference. For example with Sentence 7, it is easy to infer that Circle A is indeed disconnected from Circle B, but no approach clearly states the spatial model or the psychological processes that enable us to make this conclusion.

Based on Knauff's claims that spatial reasoning "requires going beyond language and is based on the construction, inspection, and variation of spatially organized mental representations"(Knauff, 2013, p.xi), the most probable representation of the

spatial model constructed for Sentence 7 is the diagrammatic one presented in Figure 5.1:¹

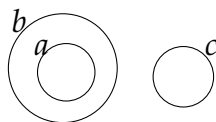


Fig. 5.1. Spatial model postulated for the sentence "Circle A is inside Circle B and Circle C is disconnected from Circle B".

On the one hand, postulating this kind of representation seems to fit perfectly with the idea developed in PRISM according to which one must inspect the model to find out if the conclusion is true or not. This is also in line with the more general claims of the classical theory that considers mental models similar to diagrams in which the information is encoded in an homomorphic way. On the other hand, it is unclear to know whether the representation in Figure 5.1 is in line with some parts of Knauff's characterization, especially when the author says that spatial models are "representations whose parts contain only the relations sufficient to permit particular calculations or operations" (Knauff, 2013, p.37).

In what follows, several assertions about representation systems in general are used to form what we call the *inferential argument*. Its conclusion, which we call the *diagrammatic hypothesis*, enables us to define the nature of spatial models and thus clarify how Knauff's characterization should be understood.

¹ Although objects are labeled with capital letters in problems' formulations, lower case letters are preferred in diagrams in order to be aligned with the notation of the forthcoming computational model where objects are always in lower case.

5.1 The *Inferential Argument*

The inferential argument can be presented as follows:²

- P1. *Graphical* and *linguistic* representations constitute two different classes of representations, where only the former enables inferences (i.e. FREE RIDE and CONSISTENCY CHECK) without the need of rules. These inferences constitute the *cognitive potentials* of graphical representations.
- P2. Knauff and Ragni's approach supposes that model-based reasoning relies on the properties of our spatial models to enable cognitive potentials.
- P3. Diagrammatic representations are the most abstract representations among the class of graphical ones.
- C. Spatial models must be diagrammatic in nature. It is the *diagrammatic hypothesis*.

The concept of *cognitive potentials* enabled by diagrammatic representations has been proposed by Shimojima. Interestingly, Shimojima's work can perfectly be ascribed in a framework for informational flow developed by Barwise and Seligman, as the latter have shown from the start ([Barwise and Seligman, 1997](#), Chap.20). Shimojima himself also used this framework to summarize his ideas in a recent book called *Semantic Properties of Diagrams and their Cognitive Potentials* ([Shimojima, 2015](#)).

To sum up, the inferential possibilities of concrete representations is a well-studied topic in the literature. What we see as the novelty of the present section, in contrast, is to consider *mental* representations as representations of these systems.

² The letters P1, P2, P3 and C will be used in the headings of this section in order to keep track of the argument progression.

5.1.1 The Representation (token) and its Information (types)

Barwise and Seligman name any representation that supports information a *token*.³ For example, the written Sentence 7 above and the concrete diagram constructed for it in Figure 5.1 are both tokens. For this example, the diagram is a *representation* (noted s) of the sentence, and the sentence is the *represented object* (noted t). These two elements compose a *representation system* because we can say that the diagram *represents* the sentence, which is also noted $s \rightsquigarrow t$, as visible in Figure 5.2.⁴

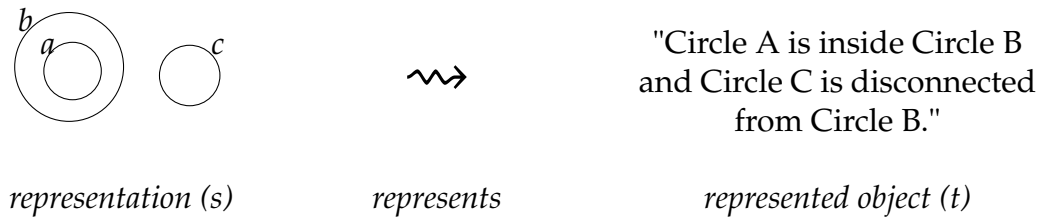


Fig. 5.2. A diagrammatic representation s on the left and a represented object t on the right. The first token *represents* the second token.

The pieces of information supported by a token are called *types* in Barwise and Seligman's approach.⁵ A type is noted σ and is called a *source type* when it is supported by a representation s , whereas it is noted θ and called a *target type* when it is supported by a represented object t .⁶ In the example above, when they are supported by sentences, the pieces of information that Circle A is inside Circle B and that Circle C is disconnected from Circle B correspond to θ^1 and θ^2 . The two same

³ As clearly mentioned by the authors, the forthcoming distinction between *type* and *token* differs from the one usually used in philosophy (Barwise and Seligman, 1997, p.27).

⁴ Defining a representation system completely also requires to specify the aspects of the tokens that are taken into account, which will be done in the second chapter of this part. Despite this, the term *representation system* can already be used to help understanding the content of this chapter.

⁵ A difficult aspect of types should be emphasized: a type is neither a sentence, nor a diagram, but the abstract piece of information itself, which can be supported by any token. This abstract level of types should never be confused with the concrete level of token, especially since we will inevitably be obliged to use concrete sentences to describe types at some points.

⁶ In our approach, σ will always be used for types supported by diagrammatic representations, and θ for linguistic representations.

pieces of information are also supported by the diagrammatic representation, and they correspond to σ^1 and σ^2 in this case. It can be said that the source types σ^1 and σ^2 indicate the target types θ^1 and θ^2 , which are noted $\sigma^1 \Rightarrow \theta^1$ and $\sigma^2 \Rightarrow \theta^2$.

When the tokens and types are combined in a general framework, a two-tier analysis of representations appears. Figure 5.3 presents the overall framework formed by tokens and types.

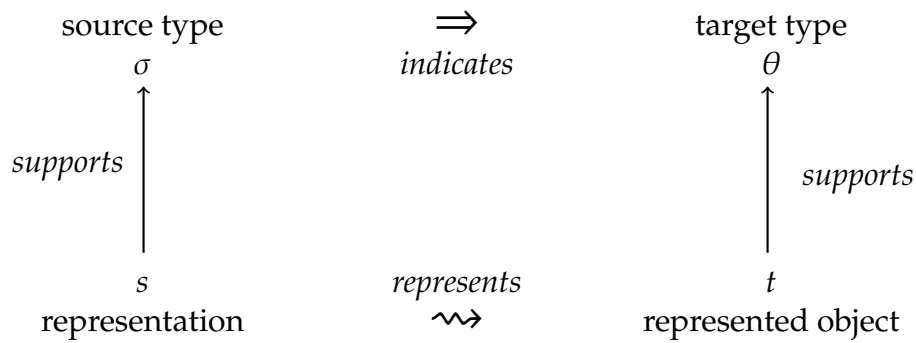


Fig. 5.3. General framework for the analysis of representation systems.

5.1.2 Constraints Between Types

A constraint can be paraphrased as "a regularity, of whatever origin or origins, that governs the distributions of types in a particular class of tokens" (Shimojima, 2015, p.30). Constraints apply among sets of types, either the source or the target ones. Let us consider Γ and Δ , two sets of types. It is possible to say that "there is a constraint $\Gamma \vdash \Delta$ from antecedent set Γ to consequent set Δ if *some* type in Δ must hold of a token if *all* types in Γ hold of that token. Thus $\{\gamma^1, \gamma^2\} \vdash \{\delta^1, \delta^2\}$ means that if γ^1 and γ^2 hold of a token (conjunctively), then δ^1 or δ^2 must hold of that token (disjunctively)." (Smessaert et al., 2020, p.2).

An example of a particularly interesting constraint is the one where a singleton consequent set follows from a set of types, which corresponds to $\Gamma \vdash \{\delta\}$. In such a case,

we talk about a *consequential constraint* (Shimojima, 2015, p.31). Such a consequential constraint can be seen in Figure 5.4, where the set on the left contains two types and the set on the right contains a single type that results from the former. Since the constraint is of the form $\Gamma \vdash \{\delta\}$, it can be called a consequential constraint. Note that here, we are at the level of types, without using any specific token to support them. It should also be clearly kept in mind that constraints are not the types themselves, but as said above, the laws that govern the distribution of types.

$$\left\{ \begin{array}{cc} (\gamma^1) & (\gamma^2) \\ \text{circle 'A' is} & \text{circle 'C'} \\ \text{inside circle 'B'} & \text{contains circle 'B'} \end{array} \right\} \vdash \left\{ \begin{array}{c} (\delta^1) \\ \text{circle 'A' is} \\ \text{inside circle 'C'} \end{array} \right\}$$

Fig. 5.4. An example of *consequential constraint*.

One of the most fundamental contributions made by Shimojima is to highlight two very different classes of tokens: the *graphical class* of representations, and the *linguistic class* of representations (Shimojima, 1999). An essential characteristic that differentiates graphical representations from linguistic ones is the *naturalness* of the constraints that govern their types. That is, diagrammatic representations follow geometric constraints that inherently apply to them and their types. This is not the case for linguistic representations since no constraints inherently apply to a set of sentences that convey some information. This remark also applies to formal sentences where constraints exist, but which lack the naturalness of the geometric ones.⁷ The best way to understand how constraints apply differently to graphical and linguistic tokens is by studying how graphical representations enable FREE RIDE and CONSISTENCY CHECK, and linguistic representations do not.

⁷ The natural constraints of geometry and the conventional constraints of formal systems have sometimes respectively been called *nomic* constraints and *stipulative* constraints by Shimojima. It seems that the author does no longer use this terminology, although some authors continue to find it relevant (Carter, 2021).

5.1.3 The Potential for FREE RIDE (P1)

The notion of FREE RIDE is psychological since it corresponds to obtaining a new piece of information in reasoning without actually making any real deduction. Let us consider the two formulas $A \subset B$ and $C \supset B$. Each of them conveys a piece of information that can be noted θ^1 and θ^2 . These target types are represented in the upper part of Figure 5.5. If a diagram with circles is used as a representation of the two sentences, the two target types are indicated by a circle labeled 'A' inside a circle labeled 'B' (σ^1), and a circle labeled 'C' that contains the circle labeled 'B' (σ^2). The two types supported by the representation are visible in the lower part of Figure 5.5.⁸ By virtue of geometric constraints, the source type that the circle labeled 'A' is inside the circle labeled 'C' (σ^3) must also be supported by the representation.⁹ Since σ^3 also has a meaning in the represented object, that is θ^3 , we can take this piece of information for granted in the represented object (i.e. $A \subset C$). Deducing this new piece of information is called a FREE RIDE since it appears automatically in the diagrammatic representation by means of the geometric constraints that govern its types. On the contrary, deducing $A \subset C$ without this representation system would have required several formal steps in the formal system.¹⁰

⁸ Remember that these two types (i.e. pieces of information) are supported by *all* the graphical representations where a circle labeled 'A' is inside a circle labeled 'B', and a circle labeled 'C' contains a circle labeled 'B'. It is the reason why no diagram is drawn. Indeed, the particular diagram is at the level of tokens, and not at the level of the information conveyed.

⁹ To clearly experience the naturalness of this constraint, the reader should draw a diagram that satisfies σ^1 and σ^2 to see that σ^3 appears inherently within the drawing of σ^1 and σ^2 .

¹⁰ Of course, this inference has to be *valid* as the authors call it, which means that a meaningful match must exist between the represented object and its diagrammatic representation (Shimojima, 2015, p.33).

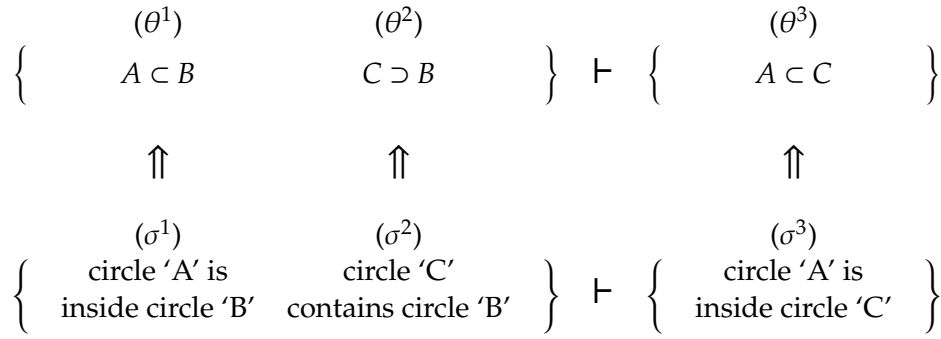


Fig. 5.5. An example of FREE RIDE.

5.1.4 The Potential for CONSISTENCY CHECK (P1 Continued)

The second cognitive potential of a representation system is CONSISTENCY CHECK. It occurs when the impossibility to construct a diagrammatic representation indicates an inconsistent formula or description in the represented object.

More formally, let us consider Δ a set of target types and Γ a set of source types.¹¹ We can say that if $\Gamma \Rightarrow \Delta$, and that $\Gamma \vdash \emptyset$, then the representation system tracks an inconsistency in Δ (Shimojima, 2015, p.87).¹² Shimojima calls the constraint of the form $\Gamma \vdash \emptyset$, an *inconsistency*, and considers it as another kind of constraint with the consequential constraint. Moreover, similarly to the fact that consequential constraints enable the cognitive potential of FREE RIDE, inconsistency enables the cognitive potential of CONSISTENCY CHECK. As an example of CONSISTENCY CHECK, let us consider the sets of types in Figure 5.6. Although they look very similar, the source and target types should not be confused since the source types are supported by a

¹¹ When the sort of the types is known (i.e. source or target) as is the case here, we believe that the symbols Σ and Θ might be a clearer terminology to use. However, for convenience, we conserve Shimojima's terminology that uses Γ and Δ to represent sets of types, even for source and target types. The use of Γ and Δ is justified by the general formal system of Barwise and Seligman, where sets of types correspond to *sequents*.

¹² A few additional conditions are formally specified by Shimojima to deal with abnormal cases. We do not include them here and refer the reader to Shimojima's book for further details (Shimojima, 2015, p.87).

diagrammatic representation, whereas the target types are supported by sentences, such as Euclidean propositions.

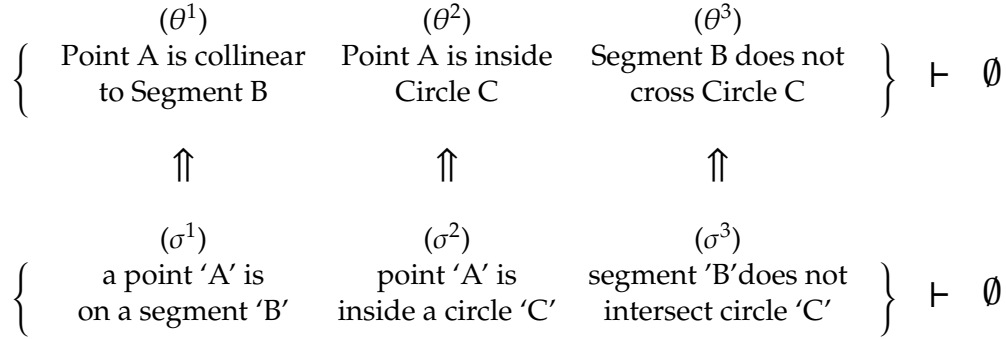


Fig. 5.6. An example of CONSISTENCY CHECK.

After a quick reasoning, it is clear that no diagrammatic representation supporting all the source types can be constructed, which means that the set of types conveyed by the sentences is also inconsistent. Once again, it is only in the diagrammatic representation where geometric rules apply that the CONSISTENCY CHECK has been made. Indeed, no graphical representation supporting σ^1 , σ^2 and σ^3 all together can be constructed.

In what follows, it is argued that model-based approaches precisely suppose that spatial models enable the cognitive potentials of FREE RIDE and CONSISTENCY CHECK.

5.1.5 Model-Based Approaches Presuppose Cognitive Potentials

(P2)

In the case of FREE RIDE where a consequential constraint applies to source types (e.g. $\{\sigma^1, \sigma^2\} \vdash \{\sigma^3\}$), no knowledge is presupposed in the reasoner's mind since the new information σ^3 simply appears in the diagram constructed for σ^1 and σ^2 . For

example in PRISM, when an object labeled 'A' is inserted to the left of an object labeled 'B' in the array (σ^1), and that an object labeled 'C' is inserted to the right of the object labeled 'B' (σ^2), the information that the object labeled 'A' is to the left of the object labeled 'C' (σ^3) becomes apparent. This new piece of information can then simply be noticed through an attentional shift, and therefore constitutes a FREE RIDE (i.e. the information is available without the need of any rule).

This potential to make possible inferences without the need of rules is precisely the claim defended by the model-based approach, and we can almost say, what differentiates the theory of mental models from mental logic in the first place. Only the semantic comprehension of the premises combined with the capacity to form a spatial model suffices to account for our inferential capacities.

Before moving to the next step of the argumentation, another resemblance between graphical representations and mental models should be emphasized: both are *agglomerative* representations (Stenning, 2000). This term means that all the pieces of information are represented together. It is opposed to *discursive* representations, where the information is only presented in a concatenated way. Of course, it is by means of the *integration* principle that mental models can be considered as agglomerative representations (Johnson-Laird, 2008, p.124).

Finally, it is important to notice that being an agglomerative representation is a necessary condition for the application of geometric constraints to the representation. Indeed, no cognitive potentials are permitted in representations that keep the pieces of information separated. In this respect, the authors of the theory of mental models have correctly characterized mental models so that cognitive potentials can occur.

5.1.6 Diagrams: the Most Simplified of the Graphical Class (P3)

So far, we have argued that the class of graphical representations differs fundamentally from linguistic representations by enabling cognitive potentials (i.e. FREE RIDE and CONSISTENCY CHECK), and that the theory of mental models presupposed these cognitive potentials in reasoning. However, it remains unclear what is exactly inside this class of graphical representations and what is not.

It is possible to answer this question by clarifying Knauff's characterizations that consider spatial models "more concrete than words, but more abstract than visual images"(Knauff, 2013, p.xi). This characterization easily invites the reader to construct a scale where *visual* and *symbolic* constitute the two extremes, and spatial models would be lying somewhere in the middle. But if we accept the previous arguments that spatial models have to enable cognitive potentials, and that only graphical representations do so, a spatial model can be moved towards the symbolic extreme of the scale as long as it does not lose the graphical properties that enable these cognitive potentials to occur. Otherwise, it would fall into the class of linguistic representations for which inferences require the use of rules.

To find this crossing point between the two classes, let us start with a clear instance of the graphical class, such as a rich image. This representation can be simplified by gradually removing features from it, such as colors, aspectual properties of the objects, details of the shapes and so on. The diagrammatic representation that results from this simplification still enables cognitive potentials, as we have seen with the examples above. Now, it is even possible to change some spatial aspects of the diagram in order to highlight other ones, as for example reducing the distances between the elements in order to make relative orientation more visible. However, changing these aspects of space does *not* remove them from the representation, and the resulting diagram is still a diagram in the same sense as the unmodified one.

This limit of simplification that necessarily applies to graphical representations can be better apprehended by the distinction between a digital and an analog form of representations proposed by Dretske.¹³ The author distinguishes the representations in *digital* form that only convey the information expected, from the representations in *analog* form that always convey superfluous information.¹⁴ Dretske illustrates the distinction by comparing how a picture and a sentence differ in the way they convey a piece of information:

To illustrate the way this distinction applies, consider the difference between a picture and a statement. Suppose a cup has coffee in it, and we want to communicate this piece of information. If I simply tell you, "The cup has coffee in it," this (acoustic) signal carries the information that the cup has coffee in it in digital form. No more specific information is supplied about the cup (or the coffee) than that there is some coffee in the cup. You are not told how much coffee there is in the cup, how large the cup is, how dark the coffee is, what the shape and orientation of the cup are, and so on. If, on the other hand, I photograph the scene and show you the picture, the information that the cup has coffee in it is conveyed in analog form. The picture tells you that there is some coffee in the cup by telling you, roughly, how much coffee is in the cup, the shape, size, and color of the cup, and so on. (Dretske, 1981, p.136)

Just as the shape, size and color of the cup cannot be removed from the picture, the location, distance or size of the objects cannot be removed in diagrams. Although a diagram can be said 'less' digital than a picture (i.e. it conveys less information than a picture), they both belong to the category of analog representations. As an illustration, consider drawing two circles on a piece of paper in order to illustrate the sentence "Circle A is externally connected to Circle B". In this drawing, the two circles will undeniably be assigned a size, a specific point where they are connected and so on.

¹³ The reader should keep in mind that this forthcoming distinction between *analog* and *digital* differs from the one proposed by Goodman for the representation of numbers (Goodman, 1976).

¹⁴ Although Dretske's work is at the origin of Barwise and Seligman's theory, the distinction does not seem to have been analyzed in their formalism to the best of our knowledge. Such an analysis is nevertheless perfectly possible by means of types and tokens: a representation s supports information σ^1 in *digital* form if and only if no information complementary to σ^1 is contained in representation s . On the contrary, if additional pieces of information to σ^1 are also conveyed by s (i.e. $\sigma^1, \sigma^2, \dots$), the representation carries the information in *analog* form.

Therefore, any diagrams constructed to represent a piece of information always requires to specify the spatial aspects for which no information is given. For this reason, a diagrammatic representation cannot be simplified, and therefore, constitutes the most simplified possible instance of the graphical class.¹⁵

When a source representation instantiates more information than the object it represents, Shimojima talks about OVER-SPECIFICITY. The author considers it as an additional cognitive potential of diagrams with FREE RIDE and CONSISTENCY CHECK. However, unlike the latter, OVER-SPECIFICITY should not be seen as a beneficial cognitive potential since it might convey a piece of information that is not true in the represented object. More will be said about OVER-SPECIFICITY in the next chapter. For now, we have enough elements to wrap up the inferential argument.

5.1.7 Conclusion: the *Diagrammatic Hypothesis* (C)

We have argued that linguistic and graphical representations constitute two very different classes of tokens, and that only graphical representations enable the cognitive potentials (i.e. FREE RIDE and CONSISTENCY CHECK) that permit inferences without rules. However, it is precisely one of the main goals of the theory of mental models, and also the space to reason theory proposed by Knauff, to reject the idea that model-based reasoning requires rules in the mind. For this reason, mental models have to be representations from the graphical class. Additional reflexions have demonstrated that diagrams are logically the most simplified instances of the graphical class that still enable the cognitive potentials. As a general conclusion,

¹⁵ Note that, as suggested above, one diagram may be less complex than another (e.g. two circles can be simplified to points if their interior is not used for the inference). Nonetheless, the resulting representation remains a diagrammatic representation.

it can be said that the nature of spatial models must be diagrammatic in order to make the inferential principles supported by the model-based approach viable. This conclusion is stated in the following hypothesis:¹⁶

The Diagrammatic Hypothesis: spatial models are diagrammatic representations in order to make cognitive potentials (i.e. FREE RIDE and CONSISTENCY-CHECK, but also OVER-SPECIFICITY) possible.

A direct benefit of this hypothesis is to allow a formalization of spatial models. By means of a coordinate system and parameters similar to those used in analytic geometry, each object composing a spatial model can be exhaustively characterized. Since only points, segments and circles are considered in this dissertation, the following definitions suffice to formalize all the forthcoming spatial models.

Spatial Model: a *spatial model* is formally characterized by giving exact values to its objects as follows:

- For all *points*, the two parameters x , y .
- For all *segments*, the double pair of parameters x_s , y_s and x_e , y_e respectively for a starting and an ending point.
- For all *circles*, the two parameters x and y for its center, and r for the radius.

Specifying the values of these parameters enables us to precisely describe a spatial model. For example, Figure 5.7 presents a spatial model (left) and the parametric values that can be used to formalize it exhaustively (right).

¹⁶ Despite the fact that the *inferential argument* supports it, we prefer the name of *hypothesis* for this claim, especially because the rest of the dissertation can also be seen as an assessment of its psychological viability.

Such formalization of spatial models directly follows from the diagrammatic hypothesis. However, it is important to keep in mind the mathematical goal of such a formalization, and to separate it from the psychological goal of describing our experience with spatial models. In this psychological goal, no coordinate system such as the one in Figure 5.7 will be used for the description of spatial model, which is the reason why it is already grayed out in Figure 5.7 .

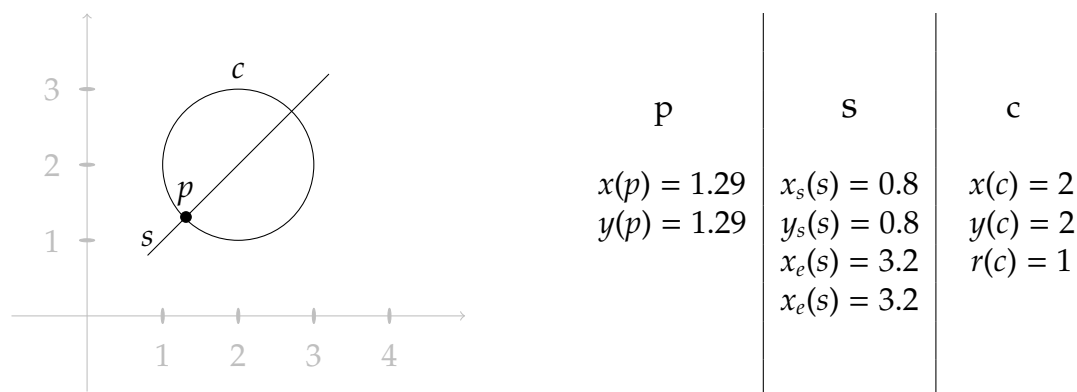


Fig. 5.7. On the left, a possible spatial model constructed from a set of sentences. On the right, the values associated with each of its objects. The axes are grayed out to represent the fact that they are only used to construct the spatial model.

The diagrammatic hypothesis might be considered as a strong claim at first sight, but the next section mainly intends to bring nuances to its comprehension, and most importantly, to prevent some misinterpretation of it. A few historical references are also given to highlight similar ideas in the literature.

5.2 Further Discussions about the Hypothesis

It is possible to find some authors who defended ideas similar to the hypothesis in the literature. Berkeley, through his attack on abstraction, is probably the most famous of them ([Berkeley, 1881](#)). The author argues against the existence of *abstract ideas* in

the mind and claims that our thoughts must always have a certain concreteness. He supports his point through a reflexion about *triangles*:

For example, does it not require some pains and skill to form the general Idea of a triangle (which is yet none of the most abstract, comprehensive and difficult) for it must be neither oblique nor rectangle, neither equilateral, equicrural, nor scalenon, but *all and none* of these at once? In effect, it is something imperfect that cannot exist, an Idea wherein some parts of several different and inconsistent ideas are put together. (Berkeley, 1881, p.5)

More recently, some experiments confirmed Berkeley's claim by showing that people indeed reason about a specific triangle, and moreover, that this triangle has a prototypical structure (Lupyan, 2017).

Of course, many counter-arguments to the ideas of the diagrammatic hypothesis exist as well. Pylyshyn, for example, thinks that it is a typical property of mental representations to be non-specific (Pylyshyn, 1973). Dennett also argues against specific mental representations by taking the example of an imagined tiger, and claiming that no one can clearly see how many stripes this tiger has in imagery (Dennett, 2002).

We believe that the debate is ill-posed for at least two reasons. The first reason is because no distinction is made between the *spatial* structure and the *visual* content of mental representations in the first place. The second reason is because several psychological facts, such as the limitation of attention, seem ignored in the debate. The cognitive facts that will be presented in what follows intend to give multiple pieces of evidence that support these two reasons, and thereby, nuance the diagrammatic hypothesis. In particular, it will be shown that the diagrammatic hypothesis cannot be understood as the simple claim that we have diagrammatic images in the mind.

5.2.1 Spatial Models as Images

Spatial models are not imagistic mental representations, but rather the spatial structures that underlie these mental representations. For example, in a spatial relation such as "Mary is to the left of Jack", some rich images can appear in the mind to represent Mary and Jack, but they are not the spatial model. The spatial model that captures the spatial information of the sentence (e.g. it can simply correspond to two points next to each other) seems more abstract than images.

Moreover, two conflicts with the previous statements of this dissertation might arise if spatial models were treated as images. Firstly, people with aphantasia and blind people would not be able to form spatial models since they are not able to form images. However, we have seen that these people can perfectly reason spatially. A second conflict would be the circularity that it creates with the two streams hypothesis. The spatial model would be processed by the "what" path as any other images, rather than the "where" path. However, it is clear that the parietal lobe, and thus the "what" path, is the location of spatial models in the brain ([Sack, 2009](#)).

Two variations might also be considered. In the first variation, the distinction between visual and spatial mental representations is refused, and instead, a unique type of *visuospatial* mental representations that vary on a continuum is considered ([Schultheis et al., 2006](#)). In this view, the mental representations for spatial models would be so poor in visual details that people with aphantasia would still be able to reason based on them. An interesting neurological fact to notice in this regard is that the representations created in the parietal lobe, such as spatial models, are not easily accessible to conscious experience ([Kosslyn et al., 2006](#)). This might explain why spatial models are more difficult to apprehend than rich images, which are, for their part, known to be easily accessible to conscious experience. A second variation would be to investigate the extent to which spatial models are sensorimotor

in nature. In other words, it would be through our bodily experience and physical comprehension of space that we reason spatially. A related idea is investigated in (Giardino, 2016), where external diagrams are considered the artifacts that permit an *external connection* between our visuo-spatial, conceptual and motor systems. But precisely, if external diagrams have the potential to link all these embodied capacities, there is no reason why an internal diagram could not do it either for enabling spatial inferences mentally.

5.2.2 Attentional Capacities, Retinotopy and the Hypothesis

Important facts about attention should not be forgotten in discussions akin to the one above. The reasoner's attention, as we have seen for PRISM, is a very limited cognitive capacity that can only focus on one thing at a time. In our approach where objects more complex than points are considered (i.e. *segments* and *circles*), it implies that multiple attentional shifts will be needed among the features of these objects in order to apprehend them. This multiplicity of shifts has been confirmed in the literature:

One remarkable finding of several studies of imagery is that while imagining something there appears to be a lot of motor activity, which resembles the exploratory movements typically made during perceptual scrutiny of an object or scene. (Laeng et al., 2014, p. 264)

Such attentional shifts towards the content of our mind can be seen as evidence for the fact that our mental representations are spatially structured, and moreover, in an equivalent way to representations formed through perception (Borst and Kosslyn, 2008). This structural equivalence is even visible at a neurological level since some brain regions have been shown *retinotopic*, that is, their neurons form a pattern of activation that is homomorphic to the main traits of the scene perceived (Tootell et al.,

1982). It is mainly the area V1 that has this property, but interestingly, the parietal lobe where spatial models are supposed to be constructed, also present a form of retinotopy (Saygin and Sereno, 2008).¹⁷

Knowing whether these neurological elements should be taken as a claim for the imagistic nature of spatial models is difficult. It is not surprising that similar inquiries occur in the literature about eye movements:¹⁸

When people imagine a visual scene, their eyes move even though there is nothing to look at. Intriguingly, there is a spatial correspondence between fixation positions during perception and visual imagery of the same content, suggesting that visual imagery has a pictorial format of representation. (Gurtner et al., 2021, p.1)

Whatever the answer to the imagistic nature of mental representations, it is important to take into account all the consequences following the attentional limitations. Indeed, even if this attentional limit does not affect the physical objects in the case of perception, this might be different for imagery, where our brain is at the same time creating *and* inspecting the representation. Questioning whether a spatial model has a pictorial format would then sound like an ill-posed question since we never have the opportunity to see such a representation as a whole anyway. In line with this last remark, some studies are more inclined to consider mental representations as a sequence of oculus motor movements encoded in the brain, instead of a fixed representation existing *per se* (Mast and Kosslyn, 2002).

¹⁷ This particularity of brain regions allows surprising possibilities, as explained in (Pearson and Kosslyn, 2015), p. 2): "Researchers have been able to *read* or *decode* a mental image from patterns of activation in area V1. That is, just based on brain activity, researchers can learn what an individual is visualizing. Critically, overlapping activity patterns occur in retinotopically organized visual areas during imagery and visual perception".

¹⁸ Despite the pieces of evidence given through the quotation of (Gurtner et al., 2021), it is also clear that the *physical* eye movements are not the reasoning process itself, but an external sign of the process carried out innerly. In this respect, some authors managed to block the participants' eye movements during spatial reasoning. The participants were still able to solve the tasks, but interestingly, they were less efficient than when their eyes could move freely (Postle et al., 2006).

5.2.3 The Information, Not the Representation, is Qualitative

It might be opportune to end these questions about the imagistic nature of spatial models by discussing how Knauff's characterization of spatial models as "qualitative representations" (Knauff, 2013, p.37) can be understood. As already explained, *qualitative* means that no exact measure is used, and it is opposed to *quantitative*.

When we consider physical representations, the term qualitative makes sense in the case of uttered sentences. Indeed, language is almost always used to convey information in a qualitative way (e.g. "move to your right", "you are taller than me", "you are late"). In the case of diagrams, it is less clear whether the representation itself can be said qualitative. It is possible to have a *qualitative* approach of a diagram (e.g. noticing the relative size, position, or topological situation between the objects), or it is possible to have a *quantitative* approach (e.g. taking a graduated rule and measuring the size of the objects, the distances between them and so on). The diagram, in itself, is just a concrete graphical representation for which it sounds inadequate to say that it is quantitative or qualitative.

When we turn to mental representations such as spatial models, we have shown that they cannot be linguistic: this is one of the main assumptions of the space to reason theory, and the inferential argument above confirms that linguistic representations do not enable the cognitive potentials expected in the model-based approach. Then, if spatial models are considered diagrammatic representation as we do it, the same line of arguments used for concrete diagrams above can be used: it does not really make sense to ask whether exact measures compose spatial models or not. In fact, supposing that spatial models do not have exact measures might even be taken as an argument for considering them as linguistic representations, since only linguistic representations can represent several values at the same time.

In contrast, what makes more sense is to consider that *only the information* retrieved

from spatial models is of qualitative nature. Our inability to calculate any measure precisely by introspection is evident: the mental distance or exact size of objects in a mental representation cannot be calculated, and nor can time be counted exactly.¹⁹ In conclusion, whereas Knauff's consideration that "spatial models are qualitative" seems to argue for a linguistic nature of models, considering that it is only the information extracted from them that is qualitative makes more sense.

Most of the questions about the imagistic nature of spatial models in this section have remained open-ended, but they have only been considered to help our comprehension of the diagrammatic hypothesis. Moreover, despite their interest for cognitive science in general, we do not believe that they directly impact our approach. Indeed, some effects that follow from the diagrammatic hypothesis can already be seen in the psychology of reasoning, even if the fundamental nature of spatial models in the brain remains an open question. Equipping our approach with a representation system that enables us to account for these effects is the goal of the next chapter. As will become clear, the over-specificity of spatial models is undeniably the most difficult aspect that has to be taken into account in this endeavor.

¹⁹ Following Peacocke's terminology for expressing this idea, authors often say that the content of the mind is *unit free* (Peacocke, 1986).

Chapter 6

OVER-SPECIFICITY: its Formal and Psychological Consequences

The biggest difficulty with the diagrammatic hypothesis is to formally and psychologically deal with OVER-SPECIFICITY (i.e. diagrammatic representations over-specify the information they represent). In this chapter, we propose an approach of spatial models that takes into account OVER-SPECIFICITY, without falling into a mathematical description that loses all psychological plausibility. This will be done in the first section, mainly through the construction of a representation system similar to the ones discussed in the first chapter, but adapted for the psychology of reasoning. The second section introduces the concept of *default specifications* that enables us to construct spatial models as close as possible to what a reasoner would intuitively construct. Finally, the third section discusses the consequences of OVER-SPECIFICITY for deductive reasoning in general.

6.1 A Representation System for Psychology

6.1.1 Shimojima's Presentation of OVER-SPECIFICITY

Shimojima defines OVER-SPECIFICITY by first defining a *disjunctive constraint* as follows: let us consider Γ^1 and Γ^2 as two sets of types. If $\Gamma^1 \vdash \Gamma^2$ and Γ^2 contains more than one member (e.g. $\{\gamma^3, \gamma^4, \gamma^5\}$), it is a *disjunctive constraint*.¹ Recall that elements of Γ^1 are taken conjunctively, whereas those of Γ^2 are taken disjunctively. Now, considering sets of target types Δ^1 and Δ^2 , and sets of source types Γ^1 , we can say that a representation *over-specifies* a represented object if $\Gamma^1 \Rightarrow \Delta^1$ and $\Gamma^2 \Rightarrow \Delta^2$, with $\Gamma^1 \vdash \Gamma^2$ being a disjunctive constraint (Shimojima, 2015, p.61).

The indeterminate problems of the preferred model theory are perfect examples of OVER-SPECIFICITY. The premises of the indeterminate problem presented in the first part are "the circle is to the left of the square" (θ^1), and "the circle is to the left of the triangle" (θ^2). When a spatial model is constructed in PRISM to support the two source types σ^1 and σ^2 , the type that the triangle is to the left of the square (σ^3), *or*, that the triangle is to the right of the square (σ^4), must also be the case. This is due to the disjunctive constraint that $\{\sigma^1, \sigma^2\} \vdash \{\sigma^3, \sigma^4\}$. It follows that the reasoner must decide between one of these two types in order to create a representation, although these pieces of information were not in the initial set of premises. Therefore, the representation is over-specific because it indicates the information θ^3 or θ^4 , but not both, although they were both following from θ^1 and θ^2 .

¹ It is the third kind of constraint with *consequential constraints* and *inconsistencies* presented previously, respectively supporting FREE RIDE and CONSISTENCY CHECK. A *disjunctive constraint* produces the cognitive potential of OVER-SPECIFICITY. Contrary to FREE RIDE and CONSISTENCY CHECK, this cognitive potential is not beneficial for reasoning.

6.1.2 The Relevant Spatial Aspects for Psychology

Generally, only one aspect of diagrams is used in representation systems, such as the topology of circles in Euler diagrams. Since Shimojima's first goal is to study such diagrammatic systems, OVER-SPECIFICITY is generally considered occurring for only one aspect in his approach. However, this restricted analysis of over-specificity only holds by convention in diagrammatic systems (e.g. it is by common agreement that we ignore the size of circles in Euler diagrams). But a diagram can be seen over-specific in more than one way when we consider several aspects of space. Shimojima also notices that OVER-SPECIFICITY depends on the conventions used in the system: "In general, whether accidental features pose real dangers, or whether a representation system is really over-specific, depends on what semantic conventions a theorist associates with the system" (Shimojima, 2015, p.69).

To correctly account for the psychology of reasoning with spatial models, we consider that at least the four following spatial aspects of relations must be taken into account in our representation system:

- **Position:** the location of the objects compared to each other (e.g. 'to the left of', 'in front of', 'above', ...).
- **Topology:** how the objects connect with each other (e.g. 'partially overlap', 'inside', 'disconnected', 'on' ...).
- **Size:** the surface occupied by the objects (i.e. 'bigger', 'smaller', 'equal').
- **Distance:** the space between two objects (e.g. 'far', 'close', 'medium' ...).

More aspects can always be considered (e.g. the size of the lines in the diagram), but we do not attribute any role to them in reasoning. By stipulating the aspects

of space that are taken into account in our approach, we are also creating a specific *representation system for psychology*, which is presented in Figure 6.1.

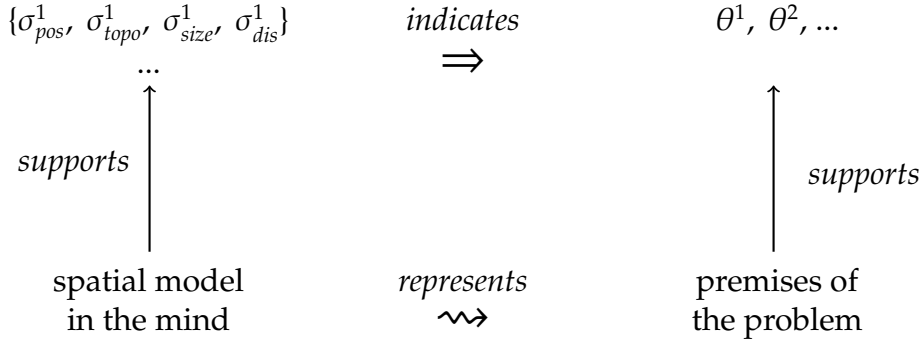


Fig. 6.1. The *representation system for psychology*. Each premise corresponds to a type θ . Each pair of objects in the spatial model corresponds to a set of types $\{\sigma_{pos}, \sigma_{topo}, \sigma_{size}, \sigma_{dis}\}$.

The *represented object* is the problem formulation (read or heard), which supports the target types ($\theta^1, \theta^2, \dots$). The *representation* is the spatial model constructed. All four spatial aspects above can be taken into account in this representation, which means that for each pair of objects, there is now a set of source types $\{\sigma_{pos}, \sigma_{topo}, \sigma_{size}, \sigma_{dis}\}$.

6.1.3 A Spatial Model in the Representation System

When a sentence is considered as a represented object, such as "Circle A is disconnected from Circle B" (θ^1), the spatial model indicating this type can be noted $(a_C, b_C) = \{\sigma_{pos}, disconnected, \sigma_{size}, \sigma_{dis}\}$.² However, as argued in the first chapter with Dretske's distinction and the diagrammatic hypothesis, the spatial model is a diagrammatic representation and thus always specifies the other aspects of space. This means that constructing the spatial model makes the other aspects of space

² Indexes to the name of objects, such as the 'C' here, might be used if it seems necessary: 'C' = Circle, 'S' = Segment and 'P' = Point.

instantiated (e.g. $(a_C, b_C) = \{left, disconnected, sameSize, medium\}$).³ To be clear, these relations are only used to characterize the spatial model, but they are not the spatial model itself, which is always a diagrammatic representation in the mind.

When the number of objects in a spatial model increases, the relations might quickly become difficult to represent. For this reason, a directed graph called a Qualitative Constraint Network (QCN) can be used to facilitate the visualization. The nodes of such a graph represent the objects of the model and the edges are labeled with the relations that hold between these objects. For example, the relation $(a_C, b_C) = \{left, disconnected, sameSize, medium\}$, combined with the relation $(c_C, b_C) = \{right, disconnected, smaller, medium\}$ and the relation $(a_C, c_C) = \{left, disconnected, smaller, far\}$ produce the QCN presented in Figure 6.2 (b). This QCN corresponds to the characterization of the spatial model presented in (a).

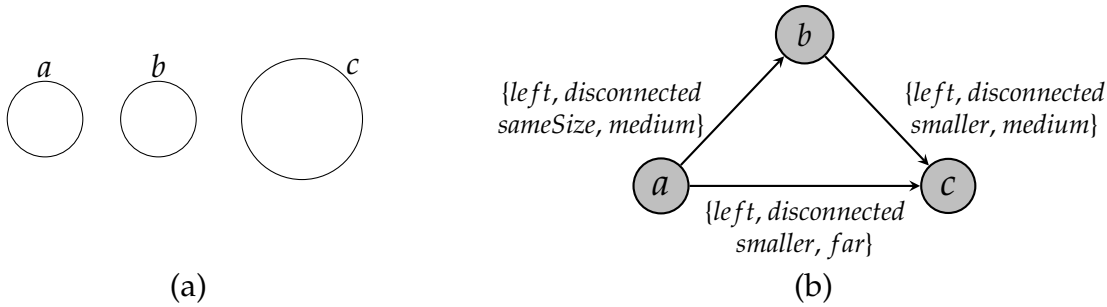


Fig. 6.2. In (a) a possible spatial model. In (b) the Qualitative Constraint Network of this spatial model when the position, topology, size and distance between the objects are taken into account.

Some aspects of space can of course be ignored in the realization of the QCN if they turn out not to play a role in the psychological phenomenon studied.⁴ Despite this

³ A formalization of this fact can be obtained by considering a *complex disjunctive constraint* between the aspects of space taken conjunctively for a relation. For example, with a topological premise, the complex disjunctive constraint is $\{\sigma_{topo}\} \vdash \{\{\sigma_{pos}^1 \vee \dots \vee \sigma_{pos}^i\} \wedge \{\sigma_{size}^1 \vee \dots \vee \sigma_{size}^j\} \wedge \{\sigma_{dis}^1 \vee \dots \vee \sigma_{dis}^k\}\}$. It means that a choice always has to be made among the relations of the other aspects of space.

⁴ In this regard, it is worth noticing that the QCN represented in Figure 6.2 (b) is already incomplete

flexibility and the relations that are omitted in the QCN, the spatial model described by it is always fully instantiated in people's mind.

6.1.4 Clarifying Knauff's Characterization

As a conclusion to this section, we now have enough elements to comment on Knauff's characterization that a spatial model "contains only the relations sufficient to permit particular calculations or operations" (Knauff, 2013, p.37), or that "they represent just the information pertinent to the inference" (Knauff, 2019, p.8). We have seen that a relation given in the premises requires a spatial model to be understood, and that spatial models are diagrammatic representations always containing superfluous information due to their analog form. Once again, Knauff's claim suggests a linguistic comprehension of models since a spatial model that contains *only* one piece of information (e.g. $(a_C, b_C) = \{\sigma_{pos}, disconnected, \sigma_{size}, \sigma_{dis}\}$) must correspond to a class of spatial models, which can only be denoted by means of symbols.

This problematic interpretation of spatial models at the level of classes is clearly endorsed by Knauff when he says that "a spatial layout model can represent a whole class of situations that differ in various visual aspects irrelevant to the inference" (Knauff, 2013, p.35). For example, the different representations in Figure 6.3 for the sentence "A is longer than B and B is longer than C" are considered corresponding to a *single* spatial model according to him.

since the edges only have one direction and the reflexive relation are not represented. These implicit relations are generally not displayed since they can always be inferred from the explicit ones.

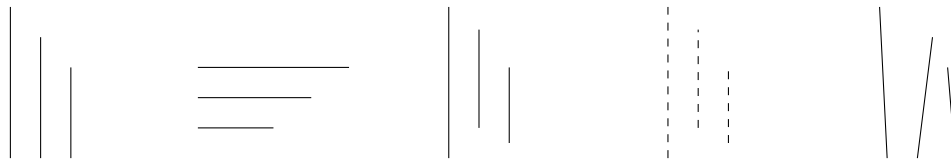


Fig. 6.3. Examples of representation of "A is longer than B and B is longer than C". According to Knauff, they can all be represented by a single spatial model (Knauff, 2013, p.35).

This claim directly conflicts with the inferential argument given above and it is difficult to see how a spatial model could stand for a class of spatial models without being symbolic. This *trade-off* between the expressiveness of linguistic classes and the inferential efficiency of diagrams is well-known in the literature (Stenning and Oberlander, 1995). There is no reason for spatial models to avoid this trade-off, even if they are mental representations: Either a spatial model is a class of diagrammatic representations, but it does not allow cognitive potentials in this case. Or, it is a diagrammatic representation that allows cognitive potential, but it must be a particular diagrammatic representation. Since it is reasonable to consider that Knauff rejects the need for rules, as typically done in a model-based approach, his characterization of spatial models must be modified. A revised version that would be coherent with the inferential argument would be:

"A spatial model is a diagrammatic representation that contains, among all the spatial information it supports, the relations that allow reasoning about the problem."

6.2 Default Specifications

6.2.1 Naturalness of the Spatial Model Constructed

As we have seen in the example above, the aspects of space unspecified by a sentence must be specified in the construction of the spatial model, and it is reasonable to claim

that regularities exist in the way people specify these aspects.⁵ It is the reason why the concept of *default specifications* is introduced in what follows. Default specifications capture what is instinctive in the construction of spatial models, or to put it differently, they deal with OVER-SPECIFICITY in the most natural way.

It is very likely that default specifications are not even consciously used in reasoning. When we understand a relation, the spatial model created for it simply comes up to our mind. However, when unnatural default specifications are used, such as in the case of a computer program that constructs a diagram without additional constraints for a sentence, the spatial model created can easily go against our intuition. For example, a computer program might construct, for the sentence "Point A is on Segment B", a diagram where Point A is at the edge of the segment, whereas we would intuitively place it somewhere in the middle. As will be seen, the aspects of our spatial models relating to default specifications can quickly be changed in our mind. However, we believe that they will also enable us to explain interesting psychological facts.

Note that Ragni and Knauff do not require the concept of default specifications since all spatial aspects except the positional one is specified in PRISM. This mainly results from the use of the array, which imposes adjacent cells to be in the relations of 'equidistance', 'equisize' and 'externally' connected (i.e. $(cell, cell_{adj}) = \{\sigma_{pos}, extConnected, sameSize, close\}$). Consequently, one spatial model differs from another only by the positional aspect.⁶

This is totally different in a framework supporting the diagrammatic hypothesis since all the aspects of a relation have to be instantiated. The way we consider default specifications to be instantiated intuitively is detailed in what follows.

⁵ The prototypical structure found for the triangles in the experimental investigation of Berkeley's thesis supports this idea (Lupyan, 2017).

⁶ In fact, even the positional aspect is not entirely free since two objects cannot be at the same place.

6.2.2 Default Specifications and Constraints

The default specifications presented here are not decided without any rationale, but satisfy the same principle of *cognitive economy* already invoked in the presentation of the preferred model theory. In other words, specifications are made in such a way that they generate the least cognitive cost. The following list presents some possible rules for the default specifications of the spatial aspects considered.

Distance Default Specifications

σ_{dist} The distance between neighboring objects is the same.

Size Default Specifications

σ_{size} Two circles that are not in a relation of inclusion have the same size.

Topological Default Specifications

σ_{topo} The objects of a spatial model are distinct (i.e. 'disconnected') from each other.

σ_{topo} A circle inside another circle is in a non-tangential inclusion (i.e. the boundaries of the two circles do not touch) ([Knauff et al., 1997](#)).

Positional Default Specifications

σ_{pos} When no positional information is provided, the objects of a problem are inserted next to each other into the spatial model according to the reasoner's reading direction.

This last positional default specification also satisfies the principle of cognitive economy, but interestingly, only because of a *cultural* specification at its origin. In other words, such a default specification instantiates an unspecified aspect of a spatial model according to a recurrent pattern encountered in the education or the everyday

life of the reasoner. For example, a westerner will preferably insert a new object to the right of the already placed objects if no positional information is given.⁷

The main origin of this effect is of course the reading direction, but additionally, this direction also shaped the construction of structures that have been learned through education. For example, it is common to postulate that an abstract mental number line is used to ground our capacity of ordinal counting (Restle, 1970).⁸ Interestingly, this mental line has been shown to present this same directional feature as reading (Dehaene et al., 1993). As a result, even the activity of counting itself enhances this directional preference in cognition, and therefore, the default specification in the construction of spatial models in general.

Many other rules can of course be added to this list and those above are only given to introduce the concept of default specifications. Eventually, a default specification function will be used to specify all the unspecified aspects when a spatial model is constructed. For example, if the sentence "Circle A is disconnected from Circle B" is given, the corresponding input of the function is $(a_C, b_C) = \{\sigma_{pos}, disconnected, \sigma_{size}, \sigma_{dis}\}$. The output might be a spatial model specified by $(a_C, b_C) = \{left, disconnected, sameSize, medium\}$, and which corresponds to the representation in Figure 6.4. We characterize the spatial model constructed as the *default model* when its default specification has not been changed yet.

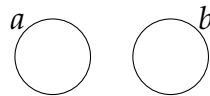


Fig. 6.4. Example of the *default spatial model* constructed for the information "Circle A is disconnected from Circle B".

⁷ Note that for the first premise of a problem, since no object is already placed in the model, this requires to know which object is inserted first. According to Ragni and Knauff, it is the second object of the relation that is inserted first in this case (Ragni and Knauff, 2013).

⁸ This structure will be described in detail in the fourth part.

Finally, an important role that will be carried out by the default specification function is to take into account the geometric constraints while instantiating the other aspects of space. This role is essential, and we believe it is one of the advantages of our approach to make it possible. More concretely, such constraints apply when the semantic of a relation given in the premise *imposes* to specify the unspecified aspects of the relation in a certain way. For example, the premise "Circle A is inside Circle B" instantiates the topological aspect between the two objects, but furthermore, imposes the relation 'smaller' to the size aspect due to geometric constraints. Consequently, the default function must specify $(a_C, b_C) = \{\sigma_{pos}, inside, smaller, \sigma_{dis}\}$ before choosing any other relation for the positional and distance aspects. As presented in the third part, all these constraints between the spatial aspects of a single pair of objects, but also between all the objects of a spatial model, will be automatically handled by a computational tool.

6.2.3 Harmless in Reasoning, Explanatory for Psychology

To conclude this section, it should be clear that using default specifications simply amounts to choosing one model among those that satisfy the premises. When the conclusion of a problem is evaluated, the default spatial model is first used to assess this conclusion. If this default model is already a counterexample to the conclusion⁹, the reasoning stops there and the problem is considered invalid. On the contrary, if the default model satisfies the conclusion, we believe that the reasoner starts a variation phase that consists in varying the default specifications in order to find whether a possible counterexample exists or not.

This last claim completes the description of default specifications: they are pre-

⁹ As explained in the first part, evaluating a problem consists in the search of a counterexample to its conclusion.

cisely intended to be changed if necessary, and therefore, they cannot constitute deducible constraints at the same level as constraints resulting from the premises. In conclusion, default specifications are harmless from a formal reasoning point of view. From a psychological point of view, on the other hand, they enable us to predict the most probable spatial model constructed in people's mind, and consequently, to account for difficulties in reasoning tasks (e.g. when a task is not congruent with the default spatial model constructed for it).

Changing the default specifications is done during a variation phase. In this respect, our approach differs from Ragni and Knauff's theory since we consider that this phase starts spontaneously. The way default specifications are supposed to be changed psychologically in the variation phase is described in the next section. The conceptual consequences of the diagrammatic hypothesis for reasoning in general are also discussed.

6.3 Variation Phase and Deductive Reasoning

Consequently to the fact that no spatial aspect of the relations remains unspecified in PRISM's framework, no variation of default specifications can occur. This might explain why Ragni and Knauff do not postulate a spontaneous variation phase in their approach. In our approach, on the contrary, all the default specifications might be varied in order to find an alternative model that is a counterexample to the conclusion. We believe default specifications are modified quite spontaneously by participants when they evaluate a problem.

Before describing this variation phase psychologically, it might be helping to clarify the different meanings of the word 'alternative' in the theory of mental models, the preferred model theory and our approach.

6.3.1 Putting the Notions of *Variation* in Order

The easiest way to distinguish the different notions of variations is first by explaining how *under-specificity* and *indeterminacy* differ.

Under-specificity comes from the fact that a piece of information only concerns one aspect of a relation without saying anything about the other aspects. Specifying these aspects is done by means of the default specifications, which enables us to select a model (i.e. the *default* spatial model) among all the alternatives. Except when it is in a mathematical context, the description of spatial models is always under-specified.

Indeterminacy, on the other hand, occurs in problems where the premises do not suffice to decide the relation between two objects *for the spatial aspect addressed by the problem*. Within Ragni and Knauff's approach where spatial models differ only by means of their positional aspect, the class of alternative models is small and finite (i.e. it contains only the spatial configurations that are possible in the array, which is generally only two). The *preferred* model is the one constructed by the participant among this class of alternatives, and it is precisely one of the goals of the preferred model theory to predict which among these models is the preferred one.

This description of indeterminacy can be extended to our framework easily by keeping in mind that indeterminacy always depends on the conclusion of the problem. For example, let us take a problem such as "Circle A is inside Circle B and Circle B is disconnected from Circle C. Is Circle A inside Circle C?". It is under-specified since the size or position of the objects are not given. However, it remains a determinate problem since the conclusion addresses the topological relation between A and C, and that precisely, this relation can only be 'disconnected'. Now, if the conclusion is changed and turns out to address the size of the circles, the problem becomes indeterminate. Indeed, several models with circles having a different relative size can be constructed, which makes the evaluation of the conclusion indeterminate.

To sum up, we propose the three following definitions to distinguish the different occurrences of ‘alternative model’ :

1. A *negation alternative* is a model in the theory of Johnson-Laird and is obtained by negating some clauses of a premise. (e.g. with the premise "I live in Paris or I live in Brussels", the mental model "I live in Paris" is a *negation alternative* of "I live in Brussels".)
2. A *preference alternative* relates to the notion used by Ragni and Knauff, and concerns the models that can change the evaluation of a conclusion.
3. Finally, a *default alternative* is specific to our approach, and concerns a spatial model obtained by variation of default specifications.

Of course, more can be said about the connections between these definitions, for example in terms of class.¹⁰ Only the notion of variation of default specifications will be needed for what follows since only determinate problems will be analyzed in depth. For this reason, the term *alternative* will implicitly refer to a default alternative.

6.3.2 Variation of the Default Model

Few articles study how the search for counterexamples is psychologically performed. An article that nevertheless brings important concepts and results to start filling this gap is proposed by Hamami, Mumma and Almaric ([Hamami et al., 2021](#)). One of the

¹⁰ If the constructed model of a determinate problem is trivially considered as the preferred one, this model can be seen as a class that contains all the default alternatives that produce the same evaluation of the conclusion. Since the alternative models defined by Ragni and Knauff are contained in the same mental model of Johnson-Laird’s theory (i.e. a same line in the truth table), this enables us to make the following general claim: the concept of model used by Johnson-Laird subsumes the concept of model used by Ragni and Knauff, which in turn subsumes the concept of default model described here.

fundamental principles highlighted by the authors is that people start searching for a counterexample by modifying the model they have in mind instead of constructing a new one.¹¹ This fact seems in line with the idea of cognitive economy that has prevailed in other psychological processes discussed so far, since constructing a new model might be cognitively more demanding than just modifying the constructed one.

Furthermore, the authors show that people do not modify all the objects of the spatial model at once, but, in accordance with the principles given about the limitations of attention, only change one object at a time. An important notion about the variation of a diagrammatic representation is the *conceptual distance*. This notion has existed for decades in the field of qualitative spatial reasoning ([Freksa, 1992](#)), and can be defined as follows:¹²

The *counterexample distance* is the distance between the diagram and the “closest” counterexample. Distance is measured here in terms of the traditional geometric transformations of translation, rotation, and uniform scaling as applied to the object of the considered scanning problem. ([Hamami et al., 2021](#), p. 16)

To illustrate this notion, the authors consider the description "Point P is outside Circle C, and Point P is on Segment S" and show the two diagrams of Figure 6.5 to represent this sentence. Whenever the conclusion "Segment S intersects Circle C" is given, these two diagrams can be varied in order to find a counterexample. It can be

¹¹ The general results of their article are obtained by means of experiments where participants first see a graphical configuration of objects, and then reason about this configuration by modifying it mentally. Since there is no reason to consider that such mental reasoning would have been different if the mental representation were constructed from a linguistic description, the results of this study can be taken into account in our approach. The graphical representation from which people start to reason in the experiment can perfectly be considered as the spatial model produced after the construction phase for linguistic premises.

¹² Note that the authors also define the notion of *counterexample density*, which corresponds to "the proportion of counterexamples obtained when considering all possible ways of placing the object to which the reasoning question applies in the diagram according to the constraints" ([Hamami et al., 2021](#), p.12).

seen that (b) is *close* to the closest counterexample, whereas (a) is *far* from the closest counterexample.



Fig. 6.5. Three models for the sentence "Point P is outside Circle C, Point P is on Segment S". When a reasoner starts searching for counterexamples to the sentence "Segment S intersects Circle C", (b) is *close* to the closest counterexample whereas (a) is *far* from it (Hamami et al., 2021, p.16).

Interestingly, it is shown that increasing the counterexample distance impairs the reasoning performances of the participants (Hamami et al., 2021, p.18). This result suggests that people modify the objects of the model in a continuous way, rather than by discrete changes that would significantly modify a model in an instantaneous way. Moreover, since the conceptual distance of the counterexample is correlated to a metric distance (i.e. increasing the conceptual distance corresponds to increasing the distance over which an object must be moved to obtain a counterexample), authors took these results as "further evidence that scanning operations [variation of the model] are sensitive to metric aspects of the diagram" (Hamami et al., 2021, p.18). Seeing such a property of a spatial aspect in our mental representations, once again, corroborates the general idea of the diagrammatic hypothesis.

A lot of research is still necessary to shed light on the way a variation phase is actually carried out. Whatever the principles one might find, a conceptual issue about the search for counterexamples must be discussed: valid arguments do not have any counterexample, but since an infinity of models can be constructed for premises in our approach, how can deductive reasoning be psychologically possible? This

conceptual problem is illustrated more in detail in what follows. Of course, it is not meant to be solved since it corresponds to centuries-old philosophical issues still discussed today.

6.3.3 Certainty and Generality: the Trade-Off

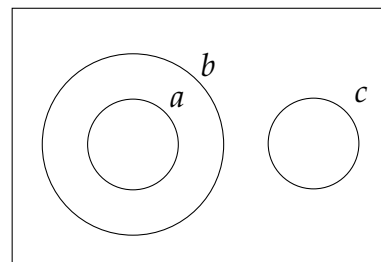
The problem results from the *trade-off* between linguistic and graphical representations already mentioned above: On the one hand, a diagrammatic representation gives us the certainty about an inference since the validity of the conclusion is directly visible in the representation. On the other hand, a diagrammatic representation only represents one model of the premises, and therefore, does not allow us to satisfy the requirement of a valid inference: the conclusion must be true in *all* possible models. For diagrammatic representations to be used in deduction, it would be necessary to construct the infinity of diagrammatic representations, which is of course problematic (Kozak, 2020).

Barwise illustrates this issue by supposing that Claire has to solve a problem similar to the one presented below (Johnson-Laird and Byrne, 1991, 1993).¹³ The diagram on the right is considered to be the spatial model constructed in Claire's head after reading the premises (i.e. after the construction phase):

P: Circle A is inside Circle B,
and Circle C is disconnected from Circle B.

Q₁: Circle A is disconnected from Circle C.

Q₂: Circle C is to the right of Circle B.



¹³ It is in fact a comparative size problem discussed by Barwise (Johnson-Laird and Byrne, 1993, p.337), but which is logically similar to the one discussed here.

As Barwise says, it is clear that Q_1 follows from P but this is not the case for Q_2 . In the case of Q_2 , Claire first inspects her model to see that Circle C is indeed to the right of Circle B. In order to make this inference valid, she starts modifying the model (i.e. a variation phase) for finding a counterexample to the problem. After a certain time, she is perfectly able to think of a model where Circle C is to the left of the other circles, while keeping the other constraints of the premises true. The resulting model is a counterexample to Q_2 , and Claire concludes that the problem is invalid.

Similarly to Q_2 , the inspection phase that Claire carries out for Q_1 terminates successfully. However, the variation phase that follows cannot result in a counterexample as is the case for Q_2 , since Q_1 does not have any counterexample. For this reason, every model that can be constructed by Claire will always satisfy Q_1 . Therefore, Claire's variation phase is necessarily doomed to failure in an endless search, as Barwise explains it :

If a counterexample is found, as in the case of Q_2 , then Claire knows that the conclusion does not follow. If none are found, as must be the case with Q_1 , then eventually Claire must give up searching and take it to follow. But since there is no way to search through all possibilities, she has to draw the line someplace. But at just that point Claire loses her claim on logical certainty. She may have reasons to suppose Q_1 follows from P , but she doesn't know it does. ([Johnson-Laird and Byrne, 1993](#), p.338)

Despite these remarks, it seems also evident, at least for a problem as simple as the one that Claire is trying to solve, that we are able to assert with certainty that the conclusion Q_1 holds. Therefore, there must be some inferential capacities that enable humans to make judgments at the model class level. Some ideas about such a reasoning capacity are briefly discussed in what follows.

6.3.4 Psychological Deductive Strategies

How the infinity of possible models can be encompassed in order to make a genuine deductive inference is a fundamental question that has prompted great philosophical investigations over the centuries.

In consequence to the fact that Berkeley can be considered as a proponent of the diagrammatic hypothesis, it is not surprising that he also stumbled on the same conceptual problem. Interestingly, he suggested that we are able to make universal judgments based on our capacity to understand how the aspects of space constraint each other. More specifically, it is because we can understand that a spatial aspect has no causal role in the spatial aspect addressed by the problem that we can ignore it in our search for counterexamples. The author describes this capacity by continuing his example about triangles:

It seems therefore that, to be certain this proposition is universally true, we must either make a particular demonstration for every particular triangle, which is impossible, or once for all demonstrate it of the abstract idea of a triangle, in which all the particulars do indifferently partake, and by which they are all equally represented. To which I answer, that though the idea I have in view whilst I make the demonstration, be, for instance, that of an isosceles rectangular triangle, whose sides are of a determinate length, I may nevertheless be certain it extends to all other rectilinear triangles, of what sort or bigness soever. And that, because neither the right angle, nor the equality, nor determinate length of the sides, are at all concerned in the demonstration. ([Berkeley, 1881](#), p.7)

Despite the interesting line of research that Berkeley's idea constitutes, it is difficult to see how to implement this capacity in our approach currently. Moreover, the literature remains quite sparse so far compared to the great amount of research that would be necessary to fully describe this capacity.

Nonetheless, the representation system for psychology proposed previously seems on the right track to implement this idea, since it gets rid of the infinite quantitative values in the same way as human do when they inspect their spatial

model. Within our formalism, only models that differ for a qualitative relation are actually considered different. It results that the possible models that have to be inspected is no longer infinite, but correspond to the finite combinations of all the qualitative relations that are considered.

6.4 Concluding Remarks

The main contributions of this part have been: (i) The clarification of the nature of spatial models by showing essential characteristics that they must have in order to allow inferential possibilities generally attributed to them. This characterization resulted in the diagrammatic hypothesis, which allows (ii) a formalization of spatial models by means of parameters. (iii) The construction of a representation system for psychology that helps characterizing spatial models in a psychological way. (iv) The proposition of default specifications to cope with OVER-SPECIFICITY in psychology. Finally (v), a stepping stone to possible research about the way people overcome the difficulty of deductive reasoning.

The question of the imagistic nature of spatial models has also been opened, but only for bringing additional elements to our comprehension of the diagrammatic hypothesis. Whether spatial models are imagistic, more abstract or even embodied will not change the fact that some effects related to the diagrammatic hypothesis can be observed in the psychology of reasoning. In this respect, the results of Hamami, Mumma and Almaric's article ([Hamami et al., 2021](#)) can already be seen in adequacy with our hypothesis, and therefore, can find a relevant framework in the one proposed here.

Now, despite all the conceptual and formal arguments given in this part, the *representation system for psychology* that has been proposed remains hollow due to

an important missing element: a formal system that accounts for all the geometric constraints and that formally defines the qualitative relations. Such a formal system has challenged the field of computer science for decades. The field of Qualitative Spatial Reasoning (Dylla et al., 2017) has produced a series of qualitative calculi that each formalizes one spatial aspect of relations for one type of object. For example, the Region Connection Calculus that accounts for the topology of regions is one the most used calculi (Randell et al., 1992). The adequacy of this field for modeling our psychology has often been noticed in the literature (Dechter and al., 2003; Mani and Pustejovsky, 2012; Forbus, 2019; Hedblom, 2020), and a specific application to Ragni and Knauff's approach has even been investigated (Olivier, 2020).

However, these calculi also suffer from important limits when it comes to account for reasoning with different types of objects, or between several aspects of space. These limitations mainly come from their logical foundations on algebraic structures - generally a relation algebra - that do not permit any modification of the calculus once it is defined (Ligozat and Renz, 2004; Inants and Euzenat, 2020). Even if some specific combinations have been shown successful (Gerevini and Renz, 2002; Kurata, 2009), the possibility of mixing all aspects of space and all types of objects at the same time, as humans naturally do, remains inconceivable with this logical foundation.

Fortunately, a recent approach (Bhatt et al., 2011; Wałęga et al., 2017; Schultz et al., 2018) solves all these problems by using polynomial constraints over the reals to define the relations, instead of algebraic methods. It is then possible to check for a possible combination of all these polynomial constraints by means of efficient solvers for first-order theories over the reals. Several computational tools that all capture the geometric constraints as we need for the representation system for psychology have been constructed.¹⁴ Among them, the tool Answer Set Programming Modulo

¹⁴ It is in fact a suite of spatial reasoning tools produced by the same group of computer scientists: <http://hcc.uni-bremen.de/spatial-reasoning/>. Although this is not necessary for the rest of

Theories for the Qualitative Spatial domain ASPMT(QS) (Walega et al., 2017) implements the diagrammatic hypothesis perfectly since it automatically constructs a diagrammatic representation as output.¹⁵ By using these possibilities to implement all the geometric constraints in our representation system, and with the addition of the psychological processes highlighted so far, it is possible to create a computational model for all kinds of spatial inferences. This computational model is called EPRISM (i.e. 'E' for Extended) and is presented in the next part.

this dissertation, the tool that will be used can be downloaded from their website, and installed by following the instructions given in Appendix A at the end of this dissertation.

¹⁵ There is also the computational argument that ASPMT(QS) enables non-monotonic and abductive reasoning, which eventually, will turn out useful for our psychological approach. More about the logical foundation of ASPMT(QS) can be found in Appendix B.

Part III

EPRISM: A Computational Model for Spatial Reasoning

Overview

This part presents the computational model EPRISM ('E' for Extended), which is a computational model similar to PRISM, but extended to spatial models of all kinds. An important goal of this part is to highlight the relevance of the computational framework Answer Set Programming Modulo Theories for the Qualitative Spatial domain ASPMT(QS) ([Wałęga et al., 2017](#)) to make EPRISM possible.

Chapter 7 describes EPRISM from a theoretical point of view. How ASPMT(QS) helps implementing the *representation system for psychology* is first presented. Then, the roles of EPRISM's components are detailed in order to allow a comprehension of the whole functioning of the computational model.

Chapter 8 is more practical since it explains how the psychological resolution of a problem is carried out. Two concrete examples are analyzed at the end of the chapter: the first example is a typical PRISM problem already presented in Part 1, and the second is the topological problem solved by Claire at the end of Part 2.

Since the literature does not provide substantial evidence for the way spatial inferences of all kinds are psychologically made, it is clear that modeling these inferences cannot be fully achieved. For this reason, the computational model described in what follows is straddling two levels of existence: on the one hand, it already integrates some psychological principles from PRISM and other psychological studies. On the other hand, it cannot be fully developed as long as no more results are obtained regarding some other aspects of reasoning (e.g. the way complex relations are understood by attentional shifts remains to be studied in depth). Therefore, even if EPRISM might not contain all the psychological principles required for being an exhaustive model of psychology, at least its use as a framework that *enables* the systematic study of these missing psychological principles is perfectly possible.

Chapter 7

Architecture

EPRISM is composed of four modules: the *natural language processing* - or more simply called the *linguistic module* -, the *controller*, the *computational tool ASPMT(QS)* and the *spatial working memory*.¹

From the problem to its final resolution, two processes follow one another in EPRISM: the *linguistic process* and the *reasoning process* (Figure 7.1).

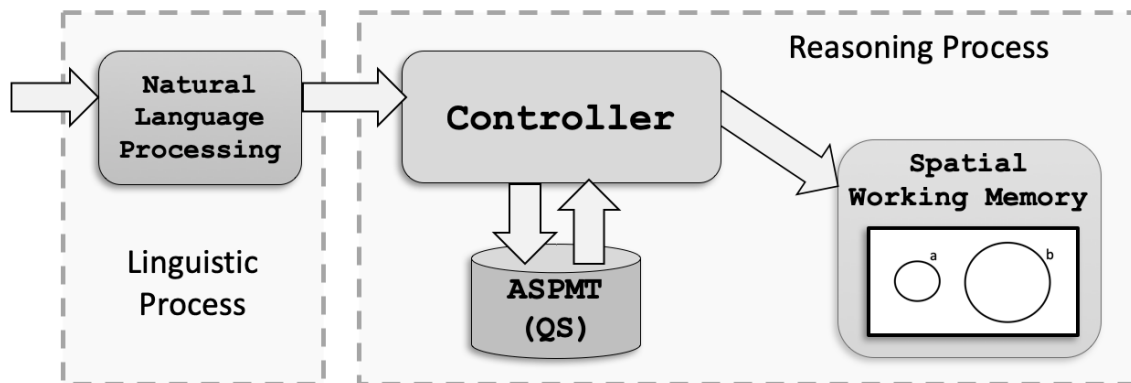


Fig. 7.1. Architecture of EPRISM showing the connections between the four modules, and the distinction between the *linguistic process* and the *reasoning process*.

¹ It is the tool ASPMT(QS) as constructed by (Walega et al., 2017) which is used in the presentation of EPRISM's architecture in this chapter. Eventually, ASPMT(QS) will be more integrated within the other components of EPRISM.

The *linguistic process* mainly translates the problem from natural language into its formal version understandable by the other components, which corresponds to a *list of triplets* (i.e. it only contains the relations in the format *relation(object, object)*). Due to its nature, the presentation of this linguistic process is saved for the next part, which is entirely dedicated to linguistic matters.

Regarding the *reasoning process*, it consists in all the operations realized on the list of triplets for producing an evaluation of the problem. Therefore, it is important to keep in mind that the only input with which the reasoning process works is this list of triplets. The controller, which can be seen as the reasoner's intention during the resolution of the problem, guides the whole reasoning process. It interacts with ASPMT(QS) and the spatial working memory when it is needed. One of the main roles of the controller is to process the different elements of the premises similarly to the way people understand the information of the problem.

Since ASPMT(QS) is the core reasoning mechanism of the reasoning process, its formal characteristics condition the format of all the information that flows between the modules. For this reason, ASPMT(QS) is presented before any other module in the next section. Moreover, we have constructed a representation system adapted for the study of spatial models in the previous part. In this system, sentences of the problem are the *represented objects*, and spatial models are the *representations* created for these sentences. In order to demonstrate the adequacy of ASPMT(QS) in implementing this representation system, each of its formal possibilities is presented by showing what element of this system it implements.

7.1 ASPMT(QS): Implementing the *Representation System for Psychology*

Basically, ASPMT(QS) implements all the elements of the representation system for psychology because: (i) Similarly to the premises that are understood through sentences of natural language (i.e. the represented object), ASPMT(QS) takes qualitative relations as input. (ii) Similarly to the spatial model that is created in the reasoner's mind (i.e. the representation), ASPMT(QS) creates a diagrammatic representation as output. Finally, (iii) similarly to the geometric constraints that govern our mental creation and use of spatial models, ASPMT(QS) cannot produce a model that goes against the laws of geometry. Each point of adequacy is detailed in what follows.

7.1.1 Represented Objects: Qualitative Relations as Facts

ASPMT(QS) is developed in a logic programming approach, which means that each problem is mainly defined by means of *facts* and *rules*. A problem given as input is called a *program* in this computational approach.

The *facts* are the true statements declared in a program, and they must be satisfied in the solution. In ASPMT(QS), a fact is a spatial relation declared in a qualitative way (e.g. *inside(a,b)*). For this reason, a fact in a program can perfectly be seen as a premise of a problem since both are qualitative.

However, declaring the premises of a problem by means of qualitative relations in ASPMT(QS) requires the previous definition of these relations. Indeed, ASPMT(QS) has to know how to interpret these relations, exactly as humans have learned how to interpret the relations in language. Defining qualitative relations is done through *polynomial constraints* in ASPMT(QS), which is the topic of what follows.

Defining Qualitative Relations as Polynomial Constraints

Relations in ASPMT(QS) are defined in an analytic manner (Bhatt et al., 2011), that is, by using object parameters similar to those used for the formalization of spatial models in the previous part. To recall, *points*, *segments* and *circles* have the following parameters:

- A *point*, noted P, has the two parameters x , y .
- A *segment*, noted S, is the double pair of parameters x_s , y_s for a starting point, and x_e , y_e for an ending point.
- A *circle*, noted C, has the two parameters x and y for the center, and the parameter r for the radius.

Each relation corresponds to a *polynomial constraint*, which is a polynomial equation or inequalities between the parameters of objects. As an example, let us consider the relation ‘externally connected’ for two circles, which corresponds to two discrete circles for which only their boundaries are touching. The polynomial constraint used to define this relation is the following first-order formula over the reals:

$$\forall c_1, c_2 \in \text{Circles} \left((x(c_1) - x(c_2))^2 + (y(c_1) - y(c_2))^2 = (r(c_1) + r(c_2))^2 \rightarrow \text{extConnected}(c_1, c_2) \right)$$

This formula basically says that for any two circles, if the parametric values used to construct them in a coordinate system satisfy the equality (i.e. the left part of the implication), then the two circles will be in a relation that we can characterize as ‘externally connected’ (i.e. the right part of the implication).² As can be seen, binary

² The relation ‘externally connected’ does not bring about semantic issues. However, other kinds of relations, such as ‘left of’ or ‘far from’, can easily accept several polynomial constraints. Discussing the difficulties resulting from the semantics of relations is the topic of next part. In the current part, we only consider relations that cannot be ambiguous.

relations are written in a prefix notation of the form $relation(object, object)$, such as $extConnected(a, b)$. Also, it should be noticed that a polynomial constraint will be required for each different pair of objects since it is not the same parameters that will be used in their definition. For example, $inside(point, circle)$ has a different polynomial constraint than $inside(circle, circle)$.

The definition of a relation is encoded as a *rule* in a program, which corresponds to the second main type of expressions in logic programming. A rule is a clause of the form $h \leftarrow b_1, \dots, b_n$ where the right part (i.e. the *body*) contains the conditions that must be fulfilled in order to make the left part (i.e. the *head*) true. In consequence, the polynomial constraint is in the body of the rule, whereas the qualitative relation is its head, as can be seen below.

```
extConnected(C1,C2)=true <- (x(C1)=X1 & y(C1)=Y1 & r(C1)=R1 & x(C2)=X2 &
y(C2)=Y2 & r(C2)=R2) & (X1-X2)*(X1-X2)+(Y1-Y2)*(Y1-Y2) = (R1+R2)*(R1+R2).
```

Additional expressions are needed to associate parametric values to variables, but the computational translation of the logical polynomial constraint 'externally connected' defined earlier is visible in the bold part of the body.

An Example of Program

Let us consider the two premises "Circle A is inside Circle B" and "Circle C is disconnected from Circle B". Table 7.1 presents a program containing these two pieces of information.³ Not all the lines deserve to be detailed for our purpose, but three sections should be highlighted: The first one is the declaration of the two premises as facts in the program (lines 17-19). The second one is the polynomial constraints

³ This program is an example of a file that can be used as input for ASPMT(QS), as described in Appendix A.

that are defined in order for ASPMT(QS) to understand the meaning of the relations (i.e. lines 27-45). The third one is the declaration of the types and names of objects in the beginning of the program (i.e. lines 1-2 and 4-5).

In more general terms, note that no default specification is declared in this example, mainly for the purpose of keeping the program relatively short. Eventually, each relation declared as a fact will also be shaped by additional relations (i.e. the default specifications) for all the spatial aspects unspecified by the premise. For example, this will prevent ASPMT(QS) from constructing, for the relation "Circle A is smaller than Circle B", a model where Circle A is inside Circle B, whereas we intuitively would place A outside B in our spatial model.⁴ As defined in the representation system for psychology, we consider that at least the position, topology, size and distance of a relation must be specified in order to create a model that seems psychologically plausible.

As can be seen in the input program of Table 7.1 (left), no quantitative values are used to declare the relations or their corresponding polynomial constraints. This evidence confirms that the framework ASPMT(QS) can be used to encode the represented objects of our system (i.e. the premises of the problem) in an adequate way: both the relations declared in ASPMT(QS) and the relations received by the reasoner through language are qualitative.

In what comes next, we will describe the model created by ASPMT(QS), which corresponds to the content of Table 7.1 (right). As will become clear, this output enables the construction of the *representation* of our system for psychology, namely, the model created in the reasoner's mind.

⁴ From a computational point of view, Circle A *inside* Circle B does not represent more or less information than Circle A *disconnected* from Circle B, whereas intuitively, the relation *inside* seems stronger in terms of information.

Table 7.1. On the left, an input program corresponding to the sentence "Circle A is inside circle B, and circle C is disconnected from circle B". On the right, the output generated for this program.

Input program	Output
<pre> 1 :- sorts 2 circle. 3 4 :- objects 5 a,b,c :: circle. 6 7 :- constants 8 x(circle) :: real[0..100]; 9 y(circle) :: real[0..100]; 10 r(circle) :: real[0..100]; 11 inside(circle, circle) :: boolean; 12 disconnected(circle, circle) :: boolean. 13 14 :- variables 15 C,C1,C2 :: circle. 16 17 %—————Facts ————— 18 inside(a,b). 19 disconnected(c,b). 20 21 %—————Default Instantiation ——— 22 23 x(C)=X. 24 y(C)=X. 25 r(C)=X. 26 27 %—————Relations ————— 28 29 %inside 30 inside(C1,C2)=true <- (x(C1)=X1 & y(C1)=Y1 & r(C1) 31 =R1 & x(C2)=X2 & y(C2)=Y2 & r(C2)=R2) & (R1<R2 & 32 (X1-X2)*(X1-X2)+(Y1-Y2)*(Y1-Y2) <= (R1-R2)*(R1-R2)). 33 34 inside(C1,C2)=false <- (x(C1)=X1 & y(C1)=Y1 & r(C1) 35 =R1 & x(C2)=X2 & y(C2)=Y2 & r(C2)=R2) & not (R1<R2 & 36 (X1-X2)*(X1-X2)+(Y1-Y2)*(Y1-Y2) <= (R1-R2)*(R1-R2)). 37 38 %disconnected 39 disconnected(C1,C2)=true <- (x(C1)=X1 & y(C1)=Y1 & 40 r(C1)=R1 & x(C2)=X2 & y(C2)=Y2 & r(C2)=R2) & 41 (X1-X2)*(X1-X2)+(Y1-Y2)*(Y1-Y2) > (R1-R2)*(R1-R2). 42 43 disconnected(C1,C2)=false <- (x(C1)=X1 & y(C1)=Y1 & 44 r(C1)=R1 & x(C2)=X2 & y(C2)=Y2 & r(C2)=R2) & not 45 (X1-X2)*(X1-X2)+(Y1-Y2)*(Y1-Y2) > (R1-R2)*(R1-R2). 46 </pre>	<pre> 1 inside_a_a_ = false 2 inside_a_b_ = true 3 inside_a_c_ = false 4 inside_b_a_ = false 5 inside_b_b_ = false 6 inside_b_c_ = false 7 inside_c_a_ = false 8 inside_c_b_ = false 9 inside_c_c_ = false 10 disconnected_a_a_ = false 11 disconnected_a_b_ = false 12 disconnected_a_c_ = true 13 disconnected_b_a_ = false 14 disconnected_b_b_ = false 15 disconnected_b_c_ = true 16 disconnected_c_a_ = true 17 disconnected_c_b_ = true 18 disconnected_c_c_ = false 19 20 r_a_ = 1.0 21 r_b_ = 2.0 22 r_c_ = 1.0 23 x_a_ = 2.0 24 x_b_ = 2.0 25 x_c_ = 6.0 26 y_a_ = 2.0 27 y_b_ = 2.0 28 y_c_ = 2.0 </pre>

7.1.2 Representation: Values for a Spatial Model

The output of Table 7.1 (right) contains two distinct parts: a first part in lines 1-18, and a second part in lines 20-28.

The second part (lines 20-28) is the most interesting for the psychological purposes of our approach, since it provides the parametric values for a possible diagrammatic representation. Any Graphical User Interface can be used to instantiate these values into objects, and thus create a diagrammatic representation similar to the one in Figure 7.2 (a). In line with the diagrammatic hypothesis, this representation can be considered as the spatial model constructed in the reasoner's mind.

However, it is important to keep in mind that these quantitative values only exist in EPRISM for the purpose of creating this diagrammatic representation. Under no circumstances these exact values can be used in the psychological process. For the latter, it is *only* the spatial model constructed with these values that can be used (i.e. inspected by attentional shifts). This limitation is represented by graying out the axes of the coordinate system.

The first part of the output (lines 1-18) contains the evaluation of the qualitative relations for each pair of objects in the model. As already presented in the previous parts, this relational information can be represented by means of a Qualitative Constraint Network (QCN), such as in Figure 7.2 (b). The advantage of ASPMT(QS) is that it evaluates all the relations defined in the program, and moreover, for each possible pair of objects. For example, without declaring any fact for it, the relation 'smaller' can just as well be defined as another polynomial constraint in the program. It would then appear among the relations of the output as the additional information that a is smaller than b . The label 'smaller' can also be added in the QCN of Figure 7.2 (b) next to the relation 'inside' between a and b .⁵

⁵ This functionality clearly shows how ASPMT(QS) overcomes the limits of previous systems from

Finally, note that no relation can remain indeterminate in this QCN because the model it describes (i.e. the model constructed in the second part of the output by the parametric values) is always specific. Such a property is another reason why ASPMT(QS) is in line with the diagrammatic hypothesis: a model is always a specific diagrammatic representation.

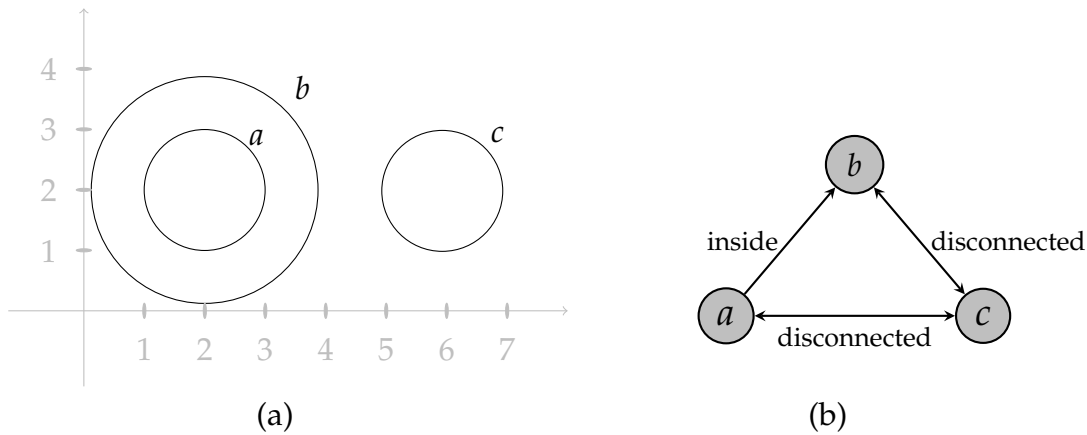


Fig. 7.2. In (a), the Qualitative Constraint Network constructed, based on the constraints given in lines 1-18. In (b), the diagrammatic representation constructed with the values given in lines 20-25.

We have shown that ASPMT(QS) allows us to define the represented objects (i.e. the qualitative relations) and the representations (i.e. the model constructed from parametric values) of our representation system for psychology. A last element that remains to be presented in order to complete the adequacy of ASPMT(QS) is how geometric constraints are taken into account.

7.1.3 Governed by Geometric Constraints

ASPMT(QS) implements the representation system for psychology because the constraints that govern the distribution of types in the representation system also apply to ASPMT(QS) when it constructs a spatial model.

qualitative spatial reasoning since a same pair of objects can be characterized through different spatial aspects at the same time.

Taking into account geometric laws in ASPMT(QS) is in fact possible by means of the qualitative constraints *taken collectively*. Indeed, the asset of ASPMT(QS) is the possibility to calculate whether all the object parameters can receive a value, considering all the constraints that apply to their parameters. In particular, when two objects are declared in a certain relation, the values that can be assigned to their parameters are constrained by the polynomial constraints of this relation. Searching for an assignment for all the parameters at once is possible in ASPMT(QS) by means of efficient solvers for first-order logic over the reals, as described in Appendix B.

In what follows, we will show how the three kinds of constraints introduced in the previous part, namely, *inconsistencies*, *consequential constraints* and *disjunctive constraints*, are respected by ASPMT(QS).

Inconsistencies

The first kind of constraint, *inconsistencies*, is the most simple one. In our representation system, an inconsistency corresponds to a case where two types cannot hold together, which is formalized by $\Gamma \vdash \emptyset$. Since a type is a piece of information, and therefore a relation in ASPMT(QS), an inconsistency corresponds to an input program where at least two declared relations cannot be satisfied at the same time. More specifically, no assignment of values can be found for the objects of these relations considering all the constraints that apply to their parameters. In case of inconsistency, ASPMT(QS) simply creates an output containing the answer 'INCONSISTENT', and no QCN or model is created.

Consequential Constraints

The second kind of constraint is the *consequential constraint*, which formally corresponds to $\Gamma \vdash \{\delta\}$. In ASPMT(QS), δ corresponds to a relation that is always true in the

output QCN, whatever the model constructed as a solution for the input program.⁶

To the best of our knowledge, a functionality that is not directly supported by ASPMT(QS) is the possibility to distinguish the relations that *must* follow from the stated relations, from those that turn out to be true only because of the specific model constructed. These latter relations correspond to the default specifications in our approach, and they might indeed have been false in another model that also satisfies the input program.

The only way to check whether a relation follows by consequential constraint is to negate this relation - with the expression 'not' - in the same input program. This method corresponds to the *reductio ad absurdum* method as is done in many deductive systems such as the analytic tableau. If the program turns out to be inconsistent with this negated fact, it means that the conclusion cannot be false, and therefore follows by means of a consequential constraint. For example, adding the negation of the relation "Circle A is disconnected from circle C" (i.e. *not disconnected(a,c)*) to the program presented in Table 7.1 creates the output INCONSISTENT. Indeed, the negated relation conflicts with the true relation of disconnectedness between A and C (i.e. line 12). This inconsistency occurs because a consequential constraint forces "Circle A is disconnected from circle C" to be always true when "Circle A is inside Circle B", and "Circle C is disconnected from Circle B" are declared.

Disjunctive Constraints

The third kind of constraint is the *disjunctive constraint* and it formally corresponds to

⁶ Recall that we consider the four kinds of types $\{\sigma_{pos}, \sigma_{topo}, \sigma_{size}, \sigma_{dis}\}$ in our representation system. Therefore, the additional precision about consequential constraints can be given: a constraint is a consequential constraint if $\Gamma \vdash \{\delta\}$, and δ corresponds to any of the four aspects of space considered. It is clear that a consequential constraint will generally occur between the types of a same aspect, but this augmented definition allows us to consider consequential constraints that cross spatial aspects. For example, for the single piece of information "Circle A is inside Circle B" ($\{\gamma_{topo}^1\}$), a consequential constraint occurs since circle 'A' being smaller than circle 'B' is the only possible type for the size aspect (i.e. $(a_C, b_C) = \{\sigma_{pos}, inside, smaller, \sigma_{dis}\}$).

$\Gamma^1 \vdash \Gamma^2$, where Γ^2 contains more than one member (e.g. $\{\gamma^3, \gamma^4, \gamma^5\}$).⁷ As already said, ASPMT(QS) does not differentiate the relations that are only over-specific from the ones that follow from a consequential constraint, which prevent us from knowing whether more than one relation is possible for an aspect of space or not. However, the only case where this differentiation will be necessary is for indeterminate problems, where several types are possible *for the aspect of space addressed by the problem*. Since we do not discuss indeterminate problems here, this functionality is currently not necessary. However, it is worth noticing that the framework of ASPMT(QS) can perfectly support this functionality, as some authors of the field have already confirmed.⁸

This last fact completes the proof that the framework ASPMT(QS) can be used for implementing all three kinds of constraints that can occur between types in our representation system. Combined with the possibility to encode the premises of a problem qualitatively (i.e. the represented objects) and the creation of a representation for them (i.e. the representation), the complete adequacy of ASPMT(QS) for implementing the representation system for psychology is finally demonstrated.

7.2 The Roles of the Reasoning Components

In this section, the roles fulfilled by the different modules of EPRISM are described. The elements given in the previous section about ASPMT(QS) allow us to already present an overview of all these roles (Figure 7.3). A detailed description of each module is then given in the rest of this section.

⁷ In order to be adapted to our approach, the disjunctive constraint can receive a similar precision than the one given for the consequential constraint above, that is, the multiplicity must occur for the same aspect of space (e.g. $\{\gamma_{topo}^3, \gamma_{topo}^4, \gamma_{topo}^5\}$).

⁸ A prototype version of a tool with this functionality has already been created by Beidi Li and Carl Schultz (personal communication).

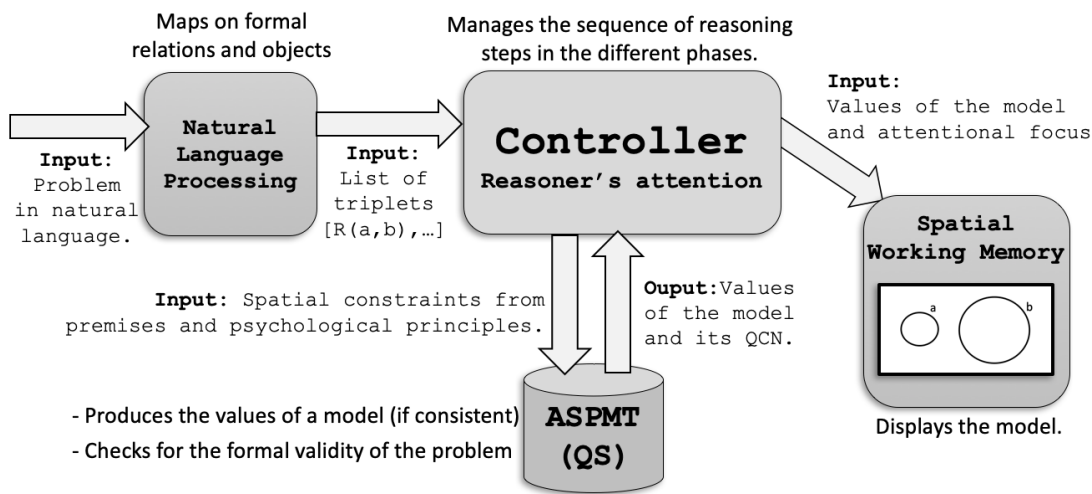


Fig. 7.3. Architecture of EPRISM showing the connections between the four modules, the input and output of each module, and their role in the computational model.

7.2.1 The Controller

Since the reasoning process only needs the list of triplets to start, the controller only interacts once with the linguistic module. On the contrary, the controller interacts several times with ASPMT(QS) and the spatial working memory during the resolution of a problem. The different roles of the controller in the resolution of a problem are the following:

- Extracting the information from the list of triplets sequentially, shaping them with default specifications and psychological principles, and sending these elements to ASPMT(QS).
- Receiving from ASPMT(QS) the parametric values - and the corresponding QCN - of a model, or else the answer 'INCONSISTENT'.
- Sending the parametric values of the model to the spatial working memory.

- Guiding the attentional shifts of the reasoner among the objects and places of the spatial model created in the spatial working memory.
- Producing the psychological evaluation of the problem based on the information extracted by attentional shifts.
- Sending the negated conclusion as a fact to ASPMT(QS) in order to produce the formal evaluation of a problem.

Additionally to these roles, the controller will also be the module where psychological characteristics are encoded (e.g. the preferences used in indeterminate problems, the limit of memory and so on). These characteristics might not constitute clear commands in the controller, but they rather result from the entire algorithmic process that underlies the module.

7.2.2 ASPMT(QS), its *Semantic, Formal and Psychological* Roles

As we have explained in the last section, ASPMT(QS) is the location where the relations are defined (i.e. the polynomial constraints). Therefore, a first important role of ASPMT(QS) that is not visible in the reasoning process itself, is to contain all this semantic information. Defining these relations in a natural way involves cognitive notions and elements of linguistics that will mainly be discussed in the fourth part.

Within the resolution process itself, the main functions of ASPMT(QS) in EPRISM are the following:

- Creating a model that satisfies all the elements of the premises and the default specifications received from the controller.

- Sending the parametric values of this model - and its QCN - to the controller.
- Detecting when the elements received by the controller are inconsistent.
- Producing the formal evaluation of the problem at the end of the reasoning process by means of a program containing the negated conclusion.

The distinction between the *formal* and the *psychological* resolution of a problem highlighted in the introduction is significant to clearly understand the role of ASPMT(QS) in the reasoning process. Indeed, ASPMT(QS) participates in both resolutions, but is not based on the same parts of its output.

For the psychological resolution, ASPMT(QS) is used as many times as the controller requires the construction of a new spatial model (i.e. an update with a new object). Only the part of the output containing the values of a possible model is used for this purpose (e.g. lines 20-28 in Table 7.1). Detecting incoherent spatial descriptions, as people naturally do, is an additional function of ASPMT(QS) that is inherently carried out in this use.

For the formal resolution, ASPMT(QS) is only used once at the end of the reasoning process. A program that contains the negation of the conclusion along with all the other facts of the problem is sent to ASPMT(QS) by the controller. If the program produces the output 'INCONSISTENT', the inference is valid. If a model is produced, it represents a counterexample to the conclusion, and the inference is invalid.

7.2.3 Representations in the Spatial Working Memory Module

The spatial working memory is the place where the spatial model is constructed and the attentional focus is directed. As already claimed, the spatial working memory does not provide any exact measure about the objects or their relations. In order to

make this limitation real in the computational model, the spatial working memory displays the spatial model *without* any graduated axis or coordinate system. Figure 7.4 shows an example of a spatial model displayed in the working spatial memory.

It is towards the objects and places of this spatial model that the controller directs the reasoner's attention. The limitations of attention cannot be instantiated similarly to PRISM (i.e only one cell can be focused at a time) since the spatial working memory of EPRISM does no longer contains cells. Therefore, in accordance with studies about visual attention ([Slattery et al., 2011](#); [Franconeri et al., 2012](#)), the limitation of attention is implemented by allowing only a small area of the spatial model to be focused at. This attentional focus can be represented by means of a small lens in the spatial model, as shown in Figure 7.4.

Finally, since it remains unclear how exactly complex objects are apprehended by attentional shifts, we will simply consider that if the focus is on one part of an object, it is sufficient to consider that the whole object is aimed. This explains why Circle A is entirely apprehended in Figure 7.4 - which is represented by the fact that the object is in bold - although the focus only cover its upper part.⁹

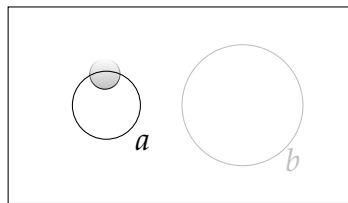


Fig. 7.4. An example of a spatial model displayed in the spatial working memory. The gray lens on Circle A represents the attentional focus. The object is in bold to represent the fact that it is apprehended by the reasoner.

⁹ Regarding this topic, some authors argue for the fact that it is the upper part of objects that is generally gazed at, as if we were about to take them ([Juravle et al., 2015](#)). This is the reason why the focus is placed on top of the circle in our example.

7.2.4 The Roles of the Reasoner's Attention

The reasoner's attention is carried out by the controller. It is the only component that enables the progression of the reasoning process. Two main *psychological operations* can be performed by attention:

1. The *insertion* of an object into the spatial model.
2. A *shift* from one object or place to another in the spatial model. Shifts fulfill several goals, such as searching for an object in the model, checking the availability of a location, or evaluating a conclusion.¹⁰

These two operations are clearly the actions that enable the progression of the reasoning process towards its resolution. In this respect, it should be stressed that any operation performed on the spatial model (i.e. inserting a new object, finding an object, inferring a relation and so on) requires attention.

The way *relations* are inspected in the spatial model does not create any problem when the objects can be reduced to single points, as is the case for the cells in PRISM. For these situations, understanding a positional relation simply amounts to shift in the relevant direction. However, how people check other kinds of relations opens up several questions, which partly results from the fact that we do not know how complex objects are apprehended. For instance, what are the attentional shifts that enable us to understand a relation of inclusion between two circles? How can we compare distances although there are no exact measures in the mind? Or finally, how can we consider a circle to be bigger than another?

Several ideas can be suggested as elements of an answer. For example, understanding the distance in our spatial models might rely on a notion of time (i.e. the

¹⁰ In computational terms, *insertion* and *shift* respectively correspond to the *writing* and *reading* operations of a system. The concept of *reasoning step* described later simply corresponds to a *state* of this system.

reasoner's attention takes more time to shift over a long distance). As for an inclusion relation, an attentional shift that exits or enters a circle must cross a boundary at least once. This feature might be used to explain how we understand that an object is inside another one.

Although these questions are still open in EPRISM, the attentional mechanism seems on the right track to implement the main idea of FREE RIDE highlighted in the previous part. Indeed, the fact that the reasoner can extract some pieces of information by means of simple attentional shifts between the objects of the spatial model constitutes clear cases of the cognitive potential. This also supports the more general idea that not much cognitive material is needed to make mental inferences possible. In fact, the mere capacities of creating a mental representation and inspecting it suffice to account for the way inferences are possible without any rules. These possibilities will become more concrete in the next chapter, where the operations that constitute a reasoning process are described.

Chapter 8

Reasoning Phases: a First Approach

This chapter is a first attempt to detail how a reasoning process might be carried out in EPRISM.¹ The three reasoning phases postulated in Ragni and Knauff's approach (i.e. the construction, inspection and variation phases) continue to structure the resolution of the problem. However, we have explained in the last part that our approach differs from Ragni and Knauff's approach by including a spontaneous variation phase of reasoning, as will be detailed in due course.

8.1 Reasoning Process and its Resolution

An important concept that enables us to structure the whole reasoning process throughout the three phases is the *reasoning step*. This concept is presented in what follows.

¹ To help the comprehension of its content, it might be useful to clearly have in mind the way PRISM carries out the resolution of a problem. We invite the reader to have a look at <http://spatialmentalmodels.appspot.com/>, or at <https://www.modeltheory.org/portfolio/spatial-reasoning/>, where the computational model is visible in action. The two tables that illustrate the processes in EPRISM (i.e. Table 8.2 and 8.3) can also be consulted previously to the reading of the theoretical content that starts this chapter.

8.1.1 Reasoning Steps

The three phases of the reasoning process are realized by means of a succession of *reasoning steps*. Cognitively speaking, a reasoning step can be considered corresponding to *one unit* of the cognitive cost that a problem requires. Psychologically speaking, a reasoning step corresponds to the minimal operation in the reasoning process, that is, one of the two psychological operations previously defined (i.e. the *insertion* of an object or an attentional *shift*).

Each reasoning step consists in characterizing: (i) the state of construction of the spatial model, and (ii) the object or place towards which the attentional focus is directed in this model. These two elements are formally described by means of:

1. The *parametric values* of all the objects in the spatial model, which is equivalent to the formal description of the spatial model.
2. The *qualitative constraint network* (QCN) of this spatial model.
3. The location of the *attentional focus*, which corresponds to the place or object towards which the reasoner's attention is directed in the spatial model (i.e. what is under the gray lens).

Each reasoning step ends by displaying the content of the spatial working memory, that is, a representation of the spatial model with the location where the attentional focus is aiming at. In what follows, the detail of the reasoning steps that compose the different phases is given.

8.1.2 The Construction Phase

At the initial reasoning step, both the model and the QCN are empty, and the attentional focus is at the center of the spatial working memory. When the controller

starts processing the list of triplets, each premise, depending on the objects it contains compared to the objects already placed in the spatial model, is assigned a type. Only the two following types of premises are considered:

Type 1 - (Initial premise) The two objects of the premise are new and no object is already placed in the spatial working memory.

Type 2 - (One new object) One of the objects of the premise is already placed in the model (e.g. "C is to the left of B", when 'B' is already placed in the spatial model).

The first premise is always of Type 1 since its objects are new, and the spatial model is empty. Since only two premises are discussed here, the rest of the premises in the list of triplets will always be of Type 2 in the forthcoming problems.²

In line with the principles highlighted in Ragni and Knauff's approach, the complete processing of a relation does not correspond to a single step of reasoning, but to three steps for a premise of Type 1, and maybe more for a premise of Type 2.³ Each reasoning step composing a Type 1 and Type 2 premise is detailed in what follows.

Type 1 premise

The first step of a Type 1 premise consists in inserting the first object of the relation

² There also exists a Type 3 and 4 premise. A Type 3 premise occurs when two new objects compose the premise although the model is not empty. This results in the creation of a second model in the spatial working memory. The Type 4 premise contains two objects already placed, and is logically used to merge the two existing models that have been created after a Type 3 premise. The problems involving these types of premises are quite rare and they significantly complicate the psychological analysis. For these two reasons, we do not discuss Type 3 and 4 premises here. The computational model PRISM does not fully support such problems either (Ragni and Knauff, 2013).

³ It is precisely because of this level of detail that Ragni and Knauff's computational model is able to account for psychological difficulties resulting from the premises effects (e.g. the difficulty of non-transitive problems), where other computational models cannot.

into the spatial model: the controller simply declares the object in ASPMT(QS).⁴ Although there is no relation yet, ASPMT(QS) creates a model containing this single object, and returns its parametric values. This way of functioning is in line with the formal characteristics of ASPMT(QS) described in the previous chapter, that is, it always constructs a model regardless of the set of relations declared. Then, the controller sets the attentional focus to this object in the spatial model, which has the result of completing all the information needed for this first reasoning step (i.e. the parametric values, the QCN, and the place of the focus are known). This reasoning step can be displayed in the spatial working memory module.

The second step in the processing of the premise consists in shifting attention according to the relation stated by it. This operation might be seen as a computational issue since nothing in the spatial model exists to mathematically specify the place where attention should be shifted. To cope with such a difficulty, the rest of the premise can be sent to ASPMT(QS) in order to obtain the place where the next object would be inserted. However, despite the presence of this object in the model constructed by ASPMT(QS), its values are not sent to the spatial working memory - we call it the *fictitious* object for this reason. On the contrary, its values can be used to set the attentional focus for checking the availability of the location where the object will be inserted.⁵ This terminates the second reasoning step and the spatial working memory can be displayed with the new place of the focus.

As a third and final step, the second object of the relation is concretely inserted into the spatial model, and the attentional focus is set to it. Once again, the spatial working memory displays the model to terminate the reasoning step. In the examples analyzed at the end of this chapter, the three reasoning steps of a Type 1 premise

⁴ In fact, the controller sends the place where the attentional focus is in order to place the object directly under it.

⁵ This check, besides its psychological relevance for the cognitive cost of a task, will turn out crucial to account for indeterminate problems in later versions of EPRISM.

correspond to the three first rows of each table (i.e. Tables 8.2 and 8.3).

Type 2 Premise

By considering only determinate problems, the processing of a premise of Type 2 becomes very similar to that of a Type 1 premise, except that a search for the object already placed in the model might occur for premises of Type 2. Indeed, among the two objects of the premise, the one that is already present in the spatial model might not be the one at which the reasoner's attention was aiming at the last reasoning step. Searching for an object in the model corresponds to make attentional shifts from one object to another until the object of the relation is found. Note that this operation does not require any interaction with ASPMT(QS) since no modification of the model is needed. However, each shift of attention constitutes a reasoning step, and consequently, must be displayed.

Figure 8.1 details the process of Type 2 premises. As can be seen at the bottom left, several attentional shifts may be necessary to find the object already placed in the model. Whenever this object is found, the same operation with the fictitious object enables us to shift the attentional focus to the place where the second object will be inserted. Then, the new object is concretely inserted and the attentional focus is set to it.

Although they are not included in the presentation of premises above, recall that each relation sent by the controller to ASPMT(QS) will always be accompanied with default specifications. This will enable the creation of a default model that is as close as possible to the one created in the reasoner's mind.

Finally, note that for any of these premises, it might happen that ASPMT(QS) produces the output 'INCONSISTENT' if a constraint conflicts with the already

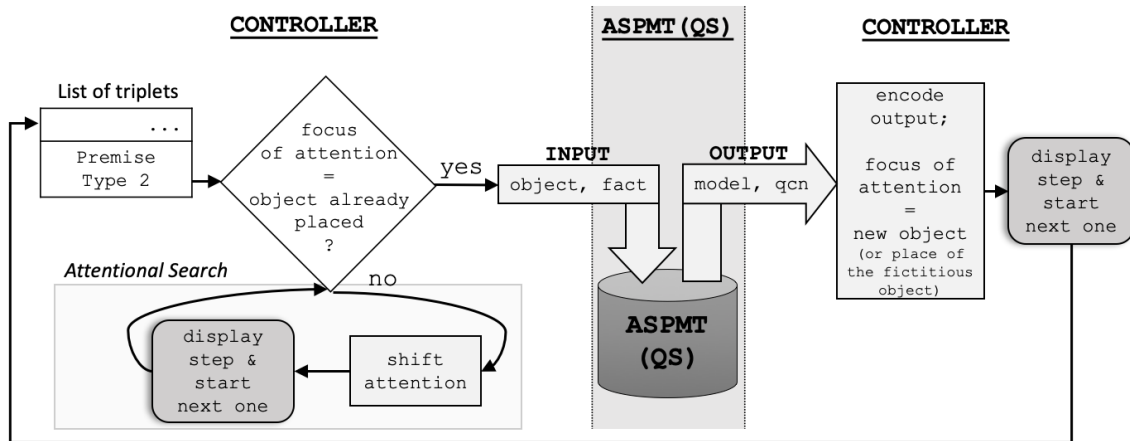


Fig. 8.1. The detailed process of Type 2 premises.

stated ones. In this case, no model can be created for this set of constraints and the problem is simply considered inconsistent.⁶ Several cases of inconsistent models will be discussed in the chapter about logical connectives at the end of the fourth part.

8.1.3 Reasoning on the Constructed Model

The *inspection phase* starts directly after the last premise by the process of the conclusion. ASPMT(QS) is no longer used for the generation of models in this phase since the spatial model is fully constructed, and only attentional shifts among its objects are necessary to inspect the relations.

Evaluating a conclusion is in fact quite similar to the process of a Type 2 premise because it starts with the search for the first object of the conclusion. As seen in the construction phase, several reasoning steps can be necessary if this object is not the one being focused on.

When the first object of the conclusion is found, a shift of attention is made in

⁶ Despite the impossibility to construct the mental representation, it is worth noticing that people can still reason about an inconsistent problem by adjusting its incoherent premises. This psychological capacity is not considered here, but represents an interesting topic of investigation (Legrenzi et al., 2003; Johnson-Laird et al., 2004).

accordance with the relation stated in the conclusion. As explained, a relation is not checked by being *seen* in the model, but instead, is checked by means of attentional shifts. For example, for a conclusion about position, checking that Object A is to the left of Object B is done by the attentional shift itself that goes from 'A' to 'B'. The way the other kinds of relations are inspected remains to be defined precisely, but several possible implementations have already been suggested.

The psychological inspection phase can have two cases as a result. In the first case, the constructed default model invalidates the conclusion. It means that the reasoner will find the two objects mentioned in the conclusion in a different relation in the model than the relation stated in this conclusion. The model constitutes a *counterexample* of the problem, which allows the process of reasoning to end. Such a psychological evaluation is formally confirmed by ASPMT(QS) which is able to generate a possible model for the program that contains the negated conclusion.

In the second case, the constructed model validates the conclusion, which corresponds for the reasoner finding the two objects of the conclusion in the same relation as the one stated by it. In such a case, either the reasoner ignores the fact that a counterexample might exist and thus concludes to the validity of the argument. Or, the reasoner starts searching for a counterexample by means of a *variation phase*. As shown at length in the first part, the argument is valid only when all the possible models are found satisfying the conclusion. Formally verifying this fact simply consists in obtaining the 'INCONSISTENT' output for the program with the negated conclusion in ASPMT(QS).

On the psychological side, however, a different procedure is needed to allow concluding to the validity of the problem. As announced before, a lot remains to be done to fully understand the way such a variation phase is carried out psychologically. Several principles were nonetheless already suggested, such as the fact that

people start the variation phase from the model they have in mind, and that they vary only one object at a time in a continuous way ([Hamami et al., 2021](#)). Since these principles suggest that the variation is similar to a kind of mental simulation, additional possibilities might be required on the computational level. In this respect, it is worth mentioning that ASPMT(QS) already has a more developed version that allows spatio-temporal reasoning ([Schultz et al., 2018](#)).

8.2 Two Examples in EPRISM

Two examples are analyzed in this last section in order to illustrate the processes described above. The first example shows how a typical problem from PRISM can be processed in EPRISM, the second example models Claire’s reasoning presented at the end of Part 2.

8.2.1 A Typical PRISM Problem in EPRISM

It is specifically the role of the linguistic process detailed in the fourth part of this dissertation to map the objects and relations of a problem to formal objects and relations understandable by ASPMT(QS). However, in the case of PRISM where so little semantic questions arise, it is already possible to give a correct formalization of the problems.

The cells of the array can be reduced to simple *points* without losing any spatial properties involved in the problem. As for the relations, only the strict positional relations ‘to the left of’, ‘to the right of’, ‘above’ and ‘below’ are permitted. These positional relations for points are obtained by specifying equalities between one axis, while using inequalities for the other axis. The corresponding polynomial constraints

are the following:

Table 8.1. Polynomial definitions of the strict positional relations used in PRISM.

```
%-----Below-----
below(P1,P2)=true <- (x(P1)=X1 & y(P1)=Y1 & x(P2)=X2 &
y(P2)=Y2) & (Y1 < Y2 & X1=X2) .

below(P1,P2)=false <- (x(P1)=X1 & y(P1)=Y1 & x(P2)=X2 &
y(P2)=Y2) & not (Y1 < Y2 & X1=X2) .

%-----Above-----
above(P1,P2)=true <- (x(P1)=X1 & y(P1)=Y1 & x(P2)=X2 &
y(P2)=Y2) & (Y1 > Y2 & X1=X2) .

above(P1,P2)=false <- (x(P1)=X1 & y(P1)=Y1 & x(P2)=X2 &
y(P2)=Y2) & not (Y1 > Y2 & X1=X2) .

%-----Left-----
left(P1,P2)=true <- (x(P1)=X1 & y(P1)=Y1 & x(P2)=X2 &
y(P2)=Y2) & (X1 < X2 & Y1=Y2) .

left(P1,P2)=false <- (x(P1)=X1 & y(P1)=Y1 & x(P2)=X2 &
y(P2)=Y2) & not (X1 < X2 & Y1=Y2) .

%-----Right-----
right(P1,P2)=true <- (x(P1)=X1 & y(P1)=Y1 & x(P2)=X2 &
y(P2)=Y2) & (X1 > X2 & Y1=Y2) .

right(P1,P2)=false <- (x(P1)=X1 & y(P1)=Y1 & x(P2)=X2 &
y(P2)=Y2) & not (X1 > X2 & Y1=Y2) .
```

For this example, we consider Problem 6 again (from Figure 4.1 used to illustrate the preferred model theory in the first part), and simply replace the names of the cars by letters:

- (6) A is to the right of B.
 C is to the left of B.
 D is in front of C.
 E is in front of B.
-















Is D to the left of E?

Table 8.2 details the resolution of the problem. Each row of the table corresponds to a reasoning step whose total can be considered as the cognitive cost required by the task (i.e. fourth column). The spatial model, as well as what the reasoner is focusing on, are displayed for each reasoning step (i.e. third column).

During the construction phase, when a new object has to be inserted into the model, an input of type " $x :: objectType$ " is sent to ASPMT(QS) (i.e. second column) in order to declare it. The symbol "?", represents the attentional check for the availability of a location. As explained before, an attentional shift to a specific place can be made by fictitiously inserting an object at this place - this is the reason why the whole relation is sent to ASPMT(QS) -, and then shifting the attentional focus to the place of this fictitious object. This operation occurs at reasoning steps 2, 4, 6 and 10 in Table 8.2. At step 8, the attentional shift occurs because of a search for Object B, which is found at the next reasoning step.

During the inspection phase, the reasoner evaluates the relation by first finding Object D, and then shifting attention accordingly with the relation. Since Object E is indeed found after this shift, the psychological evaluation results in a positive answer. A formal evaluation in ASPMT(QS) confirms this psychological evaluation.

Table 8.2. Processing steps of a typical determinate problem from PRISM. The grey lens represents the attentional focus.

Premises and Conclusion	ASPMT(QS) input	Spatial model	Reasoning Steps
Model Construction			
A is to the right of B.	a :: point.		1
	right(a,?).		2
	b :: point. right(a,b).		3
C is to the left of B.	left(?,b).		4
	c :: point. left(c,b).		5
D is in front of C.	below(?,c).		6
	d :: point. below(d,c).		7
E is in front of B.			8
			9
	below(?,d).		10
	e :: point. below(e,d).		11
Model Inspection			
Is D to the left of E? (Formal Eval.: Valid).			11
			12
Psychological Eval.: Yes .			13

Since the inspection phase ends with a positive answer, a variation phase may be started in order to check whether a counterexample can be found or not. However, the spatial aspects of the relations are so constrained in PRISM that no variation phase can be started. Although the distance can be changed in our approach, Berkeley's insight

seems relevant here to predict that the reasoner will quickly notice the impossibility to find a counterexample by modifying this spatial aspect, while respecting the premises. For this reason, the variation phase can stop almost immediately - it is not even represented in Table 8.2 - and the reasoner concludes to the validity of the problem.









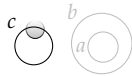
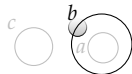
8.2.2 Claire's Reasoning in EPRISM

This second example corresponds to Problem Q_2 solved by Claire at the end of Part 2: "Circle A is inside Circle B, Circle C is disconnected from Circle B. Is Circle C to the right of Circle B?". Table 8.3 presents the possible reasoning steps that lead Claire to assess the conclusion as invalid - 'dc' means 'disconnected' in the table.

During the construction phase, several default specifications are used implicitly, such as the fact that the circle at reasoning step 3 is inserted without touching the boundary of the smaller circle. At step 4, the positional default specification described in the previous part allows us to predict that the attentional shift goes to the right for inserting the new object.

When the inspection phase starts at step 5, an attentional shift enables Claire to see that the default spatial model constructed turns out satisfying the conclusion. A variation phase is thus started. Since the search for a counterexample is not formally supported by EPRISM yet, only one reasoning step is assumed for this search, although it is surely more demanding cognitively. Claire is quickly able to modify her spatial model to a case where the conclusion does not hold, and refuses the conclusion.

Table 8.3. Reasoning steps postulated for the resolution of the problem "Circle A is inside Circle B, Circle C is disconnected from Circle B. Is Circle C to the right of Circle B?". The grey lens represents the attentional focus, and 'dc' means 'disconnected'.

Premise and Conclusion	ASPMT(QS) input	Spatial model	Reasoning Steps
Model Construction			
Circle A is inside Circle B.	<code>a :: circle.</code>		1
	<code>inside(a,?).</code>		2
	<code>b :: circle. inside(a,b).</code>		3
Circle C is disconnected from B.	<code>dc(?,b).</code>		4
	<code>c :: circle. dc(c,b).</code>		5
Model Inspection			
Is C right of B? (Formal Eval.: Invalid).			5
Yes in default model.			6
Model Variation			
			6
	Circle C is moved by continuous change...		7
Counterexample found. Psychological Eval.: No .			8

8.3 Concluding Remarks

The main contributions of this part have been: (i) The realization of the *representation system for psychology* by means of the formal possibilities offered by ASPMT(QS). (ii) The construction of the computational architecture of EPRISM, and (iii), the description of the resolution process for two concrete examples that involve different kinds of relations. These contributions all together also serve a more general goal (iv): the demonstration that, through the cognitive potentials of FREE RIDE and INCONSISTENCIAS, inferences are perfectly possible based on spatial models only.

As we have seen, several aspects of spatial reasoning still require a large amount of research, as is the case for the attentional mechanism for complex objects and relations, or the psychological realization of the variation phase. Despite these lacks, the content of this part has given many arguments for the possible realization of EPRISM, or at least for its use as an adequate framework for psychological studies about spatial reasoning ([Hamami and Mumma, 2013](#); [Shimojima and Katagiri, 2013](#); [Hamami et al., 2021](#)).

The last part of this dissertation investigates the difficulties, but also the benefits, that adding a linguistic module to EPRISM would represent. The main function of such a module would be to create the list of triplets presupposed for each reasoning process described above. Describing such a function will also give us the opportunity to discuss all the linguistic points that we overlooked throughout this part. Finally, it will also be argued that EPRISM might become relevant for other topics of cognitive science, especially by considering some elements of cognitive linguistics that allow us to extend its application to more than only spatial sentences.

Part IV

Natural Language, World Knowledge and Connectives

Overview

This part evaluates the difficulties, but also the benefits, related to the addition of a linguistic module to EPRISM.

Chapter 9 describes the main function that such a linguistic module would fulfill, which is to map the elements of sentences from natural language to formal objects and relations that are understandable by EPRISM. Moreover in this chapter, the image schema theory is presented in order to show that some parts of language that do not seem to be about space at first sight, are yet understood spatially.

Chapter 10 presents the possibilities that would be opened up if the linguistic module could also process logical connectives of natural language, such as *or*, *and*, *if* and *not*. The benefit will be particularly visible when linguistic facts highlighted in Chapter 9, are combined with the analysis of connectives. Indeed, the combination offers the possibility to take a fresh look at several topics of cognitive science, namely, the *modulation* process of the theory of mental models and the logic of conditionals.

Chapter 9

Formalizing Natural Language

This chapter presents the mapping function of the linguistic module, that is, the translating process of sentences from natural language to the formal list of triplets used in the reasoning process. For several points, the presentation that follows does not go into detail since understanding and talking about space in natural language rely on a huge range of cognitive capacities that cannot all be discussed here ([Aunague and Vieu, 2011](#)). Precisely, an additional goal of this chapter is to highlight the huge amount of implicit knowledge that underlies linguistic capacities in order to emphasize the difficulties awaited in a concrete implementation of the linguistic module in EPRISM. These difficulties, however, should not be discouraging considering the large amount of theories that end up grounding cognitive capacities on spatial thinking.

The three main roles that are carried out by the linguistic module are: (i) the mapping of objects from a description to simple geometric objects in accordance with the laws of diagrammatic simplification, (ii) the mapping of relations from natural language to formal relations, while deciding those that cannot be used in the resolution of the problem, (iii) the detection of the non-spatial sentences that, however, produce

a spatial mental representation in the reaonser's mind. The mapping on objects (i) and relations (ii) are described in the first section of this chapter, and the detection of non-spatial sentences (iii) in the second section.

9.1 Mapping to Formal Objects and Relations

Considering the working spatial memory as a two-dimensional continuous space enables us, theoretically, to process any spatial descriptions. Nonetheless, each object and relation in a description must find a formal object and relation in ASPMT(QS) that correctly captures its semantics. Of course, some existing natural language processing systems will be used to achieve this goal ([Khan et al., 2013](#)), but the content of what follows aims at highlighting the additional requirements that will be specific to EPRISM.

9.1.1 Mapping to Formal Objects

To map the objects of a description to geometric objects in ASPMT(QS), the requirements highlighted for the construction of a *diagrammatic representation* in the introduction hold: the geometric objects have to conserve the spatial properties of the elements of the description that enable the comprehension and reasoning on the problem. This task is not as easy as it sounds because a same object can be used differently by several problems. For example, the 'cars' in the problems from PRISM can be treated as mere points since no inferences about inclusion are required. However, these same cars can just as well be used for their capacity of containment in other problems, for example, when a group of people has to be distributed among them. It is precisely one of the great human capacities to quickly understand the

spatial aspects a problem is about, as well as the type of spatial entities that must be used adequately ([Aurnague et al., 2007](#)).

None of these capacities are natural from a computational point of view, and each of them requires an exhaustive formal procedure. Based on the car example alone, it is already clear that a mere database containing a one-to-one mapping between objects from common life and geometric shapes will not fulfill the task, but that a more complex process of simplification will be required. In this respect, the relations used in a description will clearly play a more important role than the objects themselves for the purpose of deciding the relevant geometric shapes for the problem.

9.1.2 Formal Definitions of Relations from Natural Language

Each relation from natural language requires a polynomial constraint on which it can be mapped in order to be understood by EPRISM. The difficulty here is to decide the exact meaning of the relation, that is, how the polynomial constraint should be defined. Moreover, recall that a relation can only be true or false in EPRISM, which means that it cannot be almost/slightly/partly satisfied in the final QCN.

Consequently to this last remark, a relation that accepts several gradients for its meaning will require a polynomial constraint for each of these gradients. For instance, the strict relation ‘to the left of’ used in the previous part (Table 8.1), cannot be used to assess that an object is to the left of another one without being *wholly* to its left. Another relation that is less strict than the previous one (e.g. using $X1 < X2$ without the requirement that $Y1 = Y2$) is needed to capture this other meaning.

Interestingly, some spatial aspects of relations seem to have more clear-cut gradients than others for their relations ([Vieu, 1991](#); [Schleipen et al., 2007](#); [Égré, 2017](#)). These differences follow from the fact that no exact measures can be calculated in the

mind, which directly impacts some relations more than others. This is particularly the case for positional relations where more than one polynomial constraints can be associated with a single word from natural language, as the predicate ‘left’ above already illustrates. The relations about distance are not easier since they must be defined in comparison to one another, that is, two objects can be considered *far* from each other only in comparison to objects that are defined as *close* to each other. The size relations seem to have more clear-cut gradients through the ‘smaller’, ‘equisize’ or ‘larger’ relations. As for the topological aspect, it presents the most interesting properties since a topological relation can be clearly distinguished from another, and moreover, based on features that do not involve distance.

In complement to these investigations, an additional question that directly follows from the diagrammatic hypothesis occurs: since the meaning of some relations clearly covers several possibilities, one of these possibilities has to be chosen to represent this relation. A perfect example where this question occurs is for the predicate ‘next to’, which clearly requires to make a choice between at least two more basic relations. A topological relation such as ‘inside’, where a circle can be included into another while touching its boundary or not, also presents this ambiguity.¹ For each of these relations, EPRISM will have to make a choice when constructing the model. Interestingly, Ragni and Knauff have already investigated these preferences for the case of topological relations ([Ragni and Knauff, 2011](#)).

¹ Of course, these questions are linked with the default specifications presented in the second part. For example, the topological default specification stated in the list of Chapter 6 has allowed us to represent the relation ‘inside’ throughout the rest of the dissertation, as an inclusion of circles where their boundaries do not touch.

9.1.3 Constraining the Set of Possible Relations

An additional difficulty regarding the relations in EPRISM is to define the relations that *can* be used for a problem, but most importantly, those that *cannot*. These choices are crucial since a small change in this set of relations might impact the evaluation of the problem.

Fortunately, for the relations that can be used, nothing has to be specified in EPRISM. Indeed, it has been shown that ASPMT(QS) is not a relation algebra, which means that there is no need to define all the converse relations or even to make the set of relations jointly exhaustive ([Ligozat and Renz, 2004](#)). For instance, a program that only contains the relation 'to the left of' will perfectly function, even without the relation 'to the right of'. The only specificity of such a program is that the QCN associated with the possible model will only contain the relation 'left of'. Despite this flexibility, it is common to define the set of relations with all the possible ones expected in the problem, especially for *production* problems where the conclusion is in the form "what is the relation between X and Y?" or "what can be deduced?".

Regarding the relations that *cannot* be used in the resolution of a problem, they have to be specified in the program. Deciding for impossible relations is easily done by humans, mainly based on their semantic knowledge. For example, people will intuitively reject the relation 'same place' for the problems in PRISM, since most of the examples involve cars that cannot pile up in parking lots. Specifying impossible relations has to be done similarly in EPRISM, otherwise, ASPMT(QS) will inevitably place the elements of the model in these relations if they participate in finding a solution.

So far, all the characteristics of the mapping function have been presented for objects and relations used in descriptions that involve spatial reasoning in a rather clear way. Additionally, the linguistic module will also fulfill a more subtle function,

which is to detect the sentences that are understood by means of a spatial model even if they are not about space at first sight. It is mainly based on theories of cognitive linguistics that such a detection is made possible, as explained in the next section.

9.2 Structures and Non-Spatial Sentences

It is common in cognitive linguistics to suppose that our comprehension of expressions and concepts rely on structures in the mind. These structures are of an abstract nature and often involve bodily notions in their composition. Despite this abstractness, some of these most simple structures can be considered through their spatial dimension only.

In this section, we will show how this spatial dimension can be captured in EPRISM. The basic idea is that when a sentence is supposed to trigger such a structure in people's mind, the linguistic module maps this sentence on the relevant spatial relation in order to correctly account for the inference made with it. To present this mapping, it is first necessary to describe the image schema theory and the main ideas of cognitive linguistic from which it originates.

9.2.1 The Image Schema Theory

The field of cognitive linguistics was born in the 70s, when linguists started to argue against the idea that our comprehension and use of language could be studied as an independent system ([Lakoff and Johnson, 1980](#)). Instead, they claimed that language had to be explained by taking into account the experiences that come from having a body with various sensorimotor capacities. In other words, the cognition cannot be studied dissociated from its embodied condition.



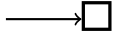

Within this line of research, the image schema theory claims that humans possess mental structures that organize everyday experiences, and ground our comprehension of abstract linguistic expressions and concepts. These structures, called *image schemas* in the theory, are formed during infancy through our capacity to extract recurrent features of physical configurations of objects and bodily experiences.

One of the most famous image schemas is CONTAINMENT², which is formed by all the situations where an object is placed inside a bigger one, such as a toy in a bucket or a ball in a cup. The repetitive experiences of this configuration of objects enable the child to extract the abstract pattern of *an object within another one* detached from any particular situation.

Later in development, image schemas are claimed to underlie our comprehension of linguistic expressions and abstract concepts. For the example of CONTAINMENT, it underlies our understanding of expressions such as ‘to be *in* the garden’, ‘to be *in* love’, or abstract concepts such as a *set* in mathematical logic. Many image schemas have been discovered and named over the years, such as SOURCE-PATH-GOAL, BLOCKED MOVE, CENTER-PERIPHERY, CYCLE, ATTRACTION, CONTACT, BALANCE (Geeraerts et al., 2005). Some of the most prominent ones are presented in more detail in Table 9.1, along with a possible representation (second column), examples of experiences from which they can be extracted (third column), and sentences or mathematical concepts in which they are used (fourth column).

² Image schemas are often written in uppercase letters.

Table 9.1. Some of the most common image schemas along with a possible representation (second column), the type of experiences from which they might originate (third column), and possible concepts in which they can be reused (fourth column).

Name	Representation	Possible Origin	Underlying
CONTAINMENT		Games with cups and balls	"to be in a country", "to be in love", a <i>set</i> in set theory
SOURCE -PATH -GOAL		Physically moving from one location to a goal	Steps of a process, the mental number line
BLOCKED MOVE		Encountering a blocking element while moving	to put a process to an end
BALANCE		The acquisition of stability in the walk	"tilt in his favor", equations in maths

Along with CONTAINMENT, the image schema of SOURCE-PATH-GOAL is also a structure that has been postulated in many cases. One of the uses for which it is most famous is our capacity of counting. Early in the study of mental representations, authors have found evidence for the fact that our comprehension of numbers consists in placing elements on a line. Dehaene highlighted the effect produced by such a phenomenon through many experiments, and following ([Restle, 1970](#)), called this structure the *mental number line*:³

I argue that tasks such as measurement, comparison of prices, or approximate calculations, solicit an "approximation mode" in which we access and manipulate a mental model of approximate quantities similar to a mental "number line" ([Dehaene, 1992](#), p.20).

Dehaene is not working within the framework of the image schema theory, but authors from the field have picked up on this idea and developed it into the more general purpose of accounting for all our mathematical capacities through image schemas ([Lakoff and Núñez, 2000](#)). Besides its use for counting, the image schema

³ It is worth mentioning that numbers are not placed at a constant distance from one another in the mental number line, and that the line is not always straight. For example, some participants represent the twelve first numbers similarly to a clock dial's configuration, and then use a vertical line departing from twelve for the other numbers ([Ranzini et al., 2009](#)).

SOURCE-PATH-GOAL seems to also support our comprehension of time ([Tenbrink, 2008](#); [Moore, 2014](#)). Expressions such as "back in time" or "in a distant time" can be considered as the signs of a spatial comprehension of this notion in the mind.

Despite these interesting signs for the presence of image schemas in cognition, it remains unclear, at this point, how a system such as EPRISM can be used as a framework to describe them. If a spatial model is diagrammatic, an image schema seems more complex than a single diagrammatic representation. Shedding some light on this complexity has mainly been undertaken by ([Mandler and Cánovas, 2014](#)). The authors proposed to distinguish three components in image schemas.

9.2.2 Three Different Components

It is certain that the graphical representation in the second column of Table 9.1 cannot legitimately represent an image schema. Firstly, because an image schema transcends any specific graphical representation (e.g. the container of the image schema of CONTAINMENT can have an infinite amount of shapes as long as it contains another object). The second reason for the impossibility to graphically represent an image schema comes from its sensorimotor components (e.g. the notions of 'push', 'force' or 'balance'), which do not have a graphical nature at all. At best, bodily notions can be represented by conventional artefacts, such as an arrow to depict some 'force' applied in the image schema.

Faced with this entanglement of elements of different nature that compose image schemas, some authors proposed the following tripartite distinction ([Mandler and Cánovas, 2014](#)). The most basic elements composing image schemas are movements made by entities (e.g. departing from a position to reach another place, or going in and out of a bigger object and so on). Due to their pure spatial nature, authors

called these basic elements *spatial primitives* in their tripartition. Developmental studies corroborate the independence of this basic level of complexity in cognition by showing that babies first pay attention to uninterpreted dynamic movements before understanding a scene, or even remembering the objects involved (Newcombe et al., 1999).⁴ The second level of complexity is obtained when spatial primitives are combined together in order to form *image schemas* such as CONTAINMENT or SOURCE-PATH-GOAL. The essence of image schemas is that they *structure* the combinations of spatial primitives. A third and final level of complexity can be reached by the addition of more abstract bodily notions to these image schemas, such as the notion of 'push', 'force', or 'balance', in order to form what authors call a *schematic integration*.

9.2.3 Capturing the Spatial Dimension of Simple Image Schemas

This tripartite distinction of levels seems to shed some light on the representational issues highlighted above, more specifically, on the reasons why these issues occurred. For the case of schematic integration, it is mainly because bodily notions are by nature different from graphical representations that they cannot be represented. The issue for the level of image schemas is different. Indeed, it is reasonable to consider the core meaning of an image schema as representable by means of a diagram, and this, even without artifacts. For example, the single relation of inclusion between the entities captures the meaning of CONTAINMENT. The representational issue, for its part, comes from the concreteness of the diagram compared to the abstract nature of the image schema since one image schema corresponds to an infinity of representations.

⁴ It is in fact quite common in the literature to find works that acknowledge this fundamental role of space for language and even cognition in general (e.g. *The Fundamental System of Spatial Schemas in Language* (Talmy, 2005), *The spatial foundations of cognition and language: Thinking through space* (Mix et al., 2010), *The spatial Foundation Of the Conceptual System* (Mandler, 2012)).

This problem has already been solved in our approach since the first aim of qualitative formalism is precisely to overcome the specificity imposed by concrete representations. In this view, the core meaning of an image schema can find a suitable translation into a qualitative constraint.

This adequacy of qualitative methods to describe image schemas has already been noticed and exploited by several authors. Two approaches are briefly presented in what follows, mainly through the image schemas of CONTAINMENT and SOURCE-PATH-GOAL. In a second time, their definitions will be described in terms of polynomial constraints suitable for EPRISM.

9.2.4 Existing Formal Approaches

One of the most recent and general approaches is the Image Schema Language ISL^{FOL} . This formal characterization of image schemas has the main purpose of categorizing them into families in order to account for concept invention in the framework of conceptual blending ([Hedblom, 2020](#)).

The language ISL^{FOL} is composed by calculi from the field of Qualitative Spatial Reasoning, and combines them with temporal and first-order logic. For example, the formula corresponding to the image schema of CONTAINMENT is the following:⁵

$$\begin{aligned} \forall x, y : Region(PP(x, y) \leftrightarrow TPP(x, y) \vee NTPP(x, y)) \\ \forall O_1, O_2 : Object(Contained_Inside(O_1, O_2) \leftrightarrow PP(O_1 \text{ inside } (O_2))) \end{aligned}$$

Due to the limits pointed out for qualitative calculi in the first part, such an approach is not adapted for reasoning. Moreover, the use of additional temporal and first-order logic increases the difficulty to make any kind of inference.

⁵ 'PP', 'TPP' and 'NTPP' respectively mean 'Proper Part', 'Tangential Proper Part', and 'Non Tangential Proper Part'. They correspond to topological relations between regions in the Region Connection Calculus ([Cohn et al., 1997](#)).

Another formal approach of image schemas turns out to be made in CLP(QS), an alternative tool to ASPMT(QS) also mentioned previously. The first purpose of formalizing image schemas in this approach is to enable the analysis of visuo-auditory perceptual signals (e.g. cognitive vision in robotics, human-centered interpretation of geospatial images and so on) in order to automatically generate linguistic descriptions (Suchan et al., 2015). For example, CONTAINMENT is defined as a rule similar to the one below:⁶

```
containment(entity(E), container(C)) :-
    region(entity(E), E_region),
    region(container(C), C_region),
    topology(nttp, E_region, C_region).
```

As for the image schema of SOURCE-PATH-GOAL, authors characterize it as an entity, called the *trajector*, that goes along a *path* delimited by a *source* and a *goal*:

```
source_path_goal(Trajector, Source, Path, Goal):-
    entity(Trajector), location(Source), location(Goal),
    path(Path, Source, Goal),
    at_location(Trajector, Source, at_time(T_1)),
    at_location(Trajector, Goal, at_time(T_2)),
    move(Trajector, Path, between(T_1, T_2)).
```

As can be noticed in the SOURCE-PATH-GOAL implementation, a notion of time is used to make the trajector movement sequential. On the contrary, CONTAINMENT is only composed by static elements.

This approach based on CLP(QS) does not intend to make a holistic classification similar to the approach with ISL, and for this reason, only a few image schemas are clearly described in the approach. Moreover, it should be noticed that the image schemas are only used to enable a mapping from visuo-spatial input to linguistic descriptions. This direction is opposite to the one used in our approach, which is

⁶ This definition is in fact a bit simplified compared to the definition used by the authors for their specific purposes.

from a linguistic input containing an image schema to a specific instance of this image schema. Consequently, a different characterization of CONTAINMENT and SOURCE-PATH-GOAL that is suitable for our goal is presented in what follows.

9.2.5 Defining CONTAINMENT for EPRISM

In EPRISM, the image schema of CONTAINMENT can be defined as the inclusion of a point (P) or circle (C), into another circle (C). Table 9.2 presents the polynomial constraints corresponding to these spatial relations.⁷

Table 9.2. Polynomial constraints for CONTAINMENT.

```
inside(P,C)=true <- (x(P)=X1 & y(P)=Y1 & x(C)=X2 & y(C)=Y2 & r(C)=R2)
& ((X2-X1)*(X2-X1)+(Y2-Y1)*(Y2-Y1) < R1*R1) .

inside(P,C)=false <- (x(P)=X1 & y(P)=Y1 & x(C)=X2 & y(C)=Y2 & r(C)=R2)
& not ((X2-X1)*(X2-X1)+(Y2-Y1)*(Y2-Y1) < R1*R1) .

inside(C1,C2)=true <- (x(C1)=X1 & y(C1)=Y1 & r(C1)=R1 & x(C2)=X2 &
y(C2)=Y2 & r(C2)=R2) & (R1<R2 & (X1-X2)*(X1-X2)+(Y1-Y2)*(Y1-Y2) <=
(R1-R2)*(R1-R2)) .

inside(C1,C2)=false <- (x(C1)=X1 & y(C1)=Y1 & r(C1)=R1 & x(C2)=X2 &
y(C2)=Y2 & r(C2)=R2) & not (R1<R2 & (X1-X2)*(X1-X2)+(Y1-Y2)*(Y1-Y2) <=
(R1-R2)*(R1-R2)) .
```

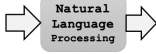



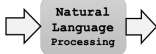

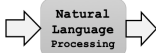

For evident reasons, the image schema of CONTAINMENT also enables us to understand the situation where an object is outside a container. Therefore, the polynomial constraints of *outside(P, C)* and *disconnected(C, C)* might also be defined in link with this image schema.

⁷ Recall that suffixes such as *PC* and *CC*, which are not written here, might be needed to differentiate the relation names.

Whenever the relation in a premise is considered understood by means of the image schema of CONTAINMENT in the reasoner's mind, the linguistic module maps this relation into one of the formal relations of Table 9.2. The exact choice of the relation depends on the geometric shape into which the contained object can be reduced to (i.e. a point or a circle), which is decided by the mapping function described in the previous section.

Table 9.3 presents sentences from natural language (first column) that are typically considered involving the image schema of CONTAINMENT in their comprehension. By means of the linguistic module, these sentences are mapped to constraints that can be used as input to ASPMT(QS) (second column), which in turn, is able to produce a possible spatial model for the original sentence (third column).

Table 9.3. In the first column, examples of sentences that are understood based on the image schema of CONTAINMENT. The second column contains the input for ASPMT(QS) obtained after the translation by the linguistic module. In the third column, examples of possible spatial models constructed for the original sentences are depicted.

Natural Language Description		ASPMT(QS) input	Spatial Model
"Alice is in Germany"		a :: point. g :: circle. inside(a,g).	
"Paris is in France"		p :: circle. f :: circle. inside(p,f)	
"The element e is in set S "		e :: point. s :: circle. inside(e,s).	
"All humans are mortal"		h :: circle. m :: circle. inside(h,m)	

The variety of sentences that seem to rely on an image schema of CONTAINMENT are worth being emphasized. Indeed, it is not only the geographical information of our everyday life that seems to rely on it, but also several important mathematical notions such as the concept of *sets*, or *statements* from syllogistic reasoning (Lakoff


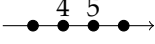

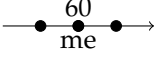

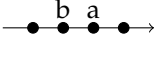
and Núñez, 2000; Geeraerts et al., 2005). As we will see later in this chapter and in the conclusion of this dissertation, postulating such mathematical concepts to be spatially understood opens up great questions about the foundation of our deductive capacities.

9.2.6 Defining SOURCE-PATH-GOAL for EPRISM

The most appropriate way to represent the SOURCE-PATH-GOAL image schema is to declare a segment on which points are placed, which is done by means of a 'collinear' relation. Then, postulating attentional shifts along this segment suffices to explain how an enumeration of its elements is psychologically possible. To ease the forthcoming formalization, the strict positional relations 'left' and 'right' defined in Table 8.1 are used to automatically place these objects in line, which consequently dispenses us to formalize an actual line in the input.

Typical sentences that are considered triggering a structure of SOURCE-PATH-GOAL in the reasoner's mind are presented in Table 9.4 (first column) along with a possible spatial model for each of them (third column) and the corresponding input in ASPMT(QS) (second column).

Table 9.4. In the first column, examples of sentences that are understood based on the image schema of SOURCE-PATH-GOAL. The second column contains the ASPMT(QS) input obtained after the translation by the linguistic module. In the third column, some examples of possible spatial models constructed with these constraints are presented.

Natural Language Description		ASPMT(QS) input	Spatial Model
"4 is less than 5"	⇒  ⇒	<pre>p4 :: point. p5 :: point. left(p4,p5).</pre>	
"I weigh 60 Kilos"	⇒  ⇒	<pre>me :: point. p60 :: point. samePlace(me,p60).</pre>	
"Runner A is ahead of Runner B"	⇒  ⇒	<pre>a :: point. b :: point. right(a,b).</pre>	

9.2.7 Difficulties with the Image Schema Theory and Promising Trends

Theoretically, the linguistic module is supposed to detect any use of image schemas within the natural language descriptions in order to produce the correct spatial model for the reasoning process. Unfortunately, no systematic approach has been found so far to automatically detect image schemas in sentences. In her work, Hedblom also notices this lack by saying that "One of the most challenging parts of using image schemas in formal systems and artificial intelligence is that there currently exists no comprehensive method to identify them in natural language" ([Hedblom, 2020](#), p.167).

However, the automatic detection of image schemas in language may also be seen as an ill-defined goal since the criteria that define an image schema are already blurred from the start. Indeed, making an exhaustive list of image schemas generates several difficulties: A first difficulty is to know the frequency with which a structure has to be found in natural language in order to be considered as an image schema. Besides, even for structures that are clear cases of image schemas, a second difficulty is to decide what variation constitutes a different image schema from the original one. For example, several small variations of the spatial properties of CONTAINMENT can be considered, such as a container that only has one or two openings, or a tight/loose type of CONTAINMENT and so on. The explosion of spatial primitives combinations should not all result in a new image schema in the theory ([Tseng, 2007](#)). Finally, these difficulties grow in complexity when it comes to naming all the structures elected as image schemas.

In view of these difficulties, it might be more relevant to suppose that a large part of language is understood by means of spatial representations, and that some of these

representations (i.e. the image schemas) are just more recurrent than others. Without the intermediate step of image schemas, a sentence would be directly mapped on the qualitative relation used to construct the spatial model of this sentence ([Mani and Pustejovsky, 2012](#); [Suchan et al., 2014](#); [Richardson et al., 2001](#)).

Before closing this chapter, it should be noticed that several structures learned through education, such as the mental number line or geographical knowledge, have been used implicitly in the examples above. It will also be the role of the linguistic module to detect when such background knowledge takes part in the resolution of a problem. From a formal point of view, each structure will have to be associated with a predefined set of constraints. When a problem presupposes some elements of background knowledge, the set of constraints corresponding to the structure invoked in the mind is loaded as additional facts into the program, and therefore constraint the search for a solution.

All the theoretical elements that have been defined in this chapter are used more concretely in the next one to analyze sentences with logical connectives (i.e. *or*, *and*, *if*). The translations described for CONTAINMENT and SOURCE-PATH-GOAL will enable us to predict the spatial models constructed in the reasoner's mind, and consequently, to account for modifications on connectives.

Chapter 10

Logical Connectives

As we have seen throughout this dissertation, language conveys pieces of information (i.e. the types) that describe the content of spatial models. Additionally to this function, language also enables us to state logical constraints between these pieces of information by linking them with different connectives, such as *and*, *or*, and *if*. Such connectives are very common in reasoning and their study represents a large amount of psychological and formal research.

In this chapter, we show that the formal and psychological elements defined previously might be used to propose an innovative analysis of problematic cases of connectives discussed in the literature. Of course the chapter neither provides the details of their implementation in EPRISM, nor the exact description of their psychological process, but it nevertheless gives the premises for making all of this possible.

The next section explains what are *sets of spatial models* and how geometric constraints can apply to them. In the last section, concrete examples from both the theory of mental models and the logic of conditionals are used to illustrate what the effects of connectives on these sets of spatial models are.

Since the following approach is mainly formal, we will represent spatial models without the psychological default specifications of the previous parts. For example, the spatial model constructed for the sentence "Point A is to the left of Point B" (θ^1) and "Point B is to the left of Point C" (θ^2), can be represented by the set $\{\sigma^1, \sigma^2\}$. Of course, all the other spatial aspects of these relations are still assumed to be instantiated in the reasoner's mind.

10.1 Constraints and Sets of Spatial Models

10.1.1 Sets of Spatial Models

As we have seen repeatedly, the conjunction *and* between two premises tells the reasoner that the constructed spatial model must support the two types conveyed by these premises.¹ Besides the connective *and*, other connectives allow to state possibilities between the pieces of information that must be supported by the spatial model. It follows that several models will be created in the reasoner's mind, each model being a possibility allowed by the connective. For example, let us consider the following disjunctive sentence:²

(8) Circle A is to the left of Circle B (θ^1) *or* Circle C is to the right of Circle B (θ^2).

¹ Without developing it, we have used this condition to stand for the notion of a 'true' relation. Since this notion gains in importance in this chapter, it might be relevant to recall that: from a formal point of view, 'true' means that the object parameters satisfy the qualitative constraint defined in ASPMT(QS), and from a psychological point of view, that the reasoner judges by means of attentional shifts that the two objects are in that relation. Despite these clarifications, it is clear that the definition of 'true' for the model-based approach should receive more attention (Elqayam, 2005).

² Throughout this chapter, we are using a general numeration of types in order to keep track of them (e.g. θ^2 stands for "Circle C is to the right of Circle B" in the entire chapter).

This sentence creates a first spatial model where a circle labeled 'A' is to the left of a circle labeled 'B' (σ^1), but also a spatial model where a circle labeled 'C' is to the right of a circle labeled 'B' (σ^2). Finally, even a third model can be created where both pieces of information are supported, that is, a spatial model where three circles labeled 'A', 'B' and 'C' are placed side by side in the spatial model. In total, three possible spatial models exist in the reasoner's mind, and this can be represented by a *set of spatial models*, such as $\{\{\sigma^1\}, \{\sigma^2\}, \{\sigma^1, \sigma^2\}\}$.

Our capacity to handle reasoning on multiple models at the same time has not directly been discussed in this dissertation, but it has been presupposed for indeterminate problems (e.g. "A is to the left of B, and A is to the left of C" creates two different spatial models in the reasoner's mind). However, the understanding of 'at the same time' needs precaution in this last sentence. In line with many arguments previously given, and especially the diagrammatic hypothesis, only one spatial model at a time can actually be considered. By saying that multiple models can mentally exist, we mean that people are able to think of a model while keeping in mind the possibilities to (re)deploy the other models. The limitations of attentional capacities would have prevented us anyway from inspecting several models *exactly* at the same time.

10.1.2 Constraints Apply To Each Spatial Model of a Set

As explained in the second part by means of the *integration principle*, people integrate the premises of a problem into a single spatial model (Johnson-Laird, 2008, 124). From a formal point of view, integration simply corresponds to representing two pieces of information in the same set (i.e. the same spatial model). For example, the spatial model $\{\sigma^1, \sigma^2\}$ indicates that the reasoner has managed to construct *one* mental

representation that supports both types σ^1 and σ^2 at the same time. As we have emphasized before, it is specifically because spatial models integrate these pieces of information together that geometric constraints (i.e. the consequential constraint $\{\sigma^1, \sigma^2\} \vdash \sigma^3$ or inconsistencies $\{\sigma^1, \sigma^2\} \vdash \emptyset$) can apply to them, and consequently permit FREE RIDE and CONSISTENCY CHECK.

Now, it should be clearly kept in mind that the psychological operation of integration described above - which simply corresponds to the comma between the types of a set -, differs from the connective *and* in language. This distinction between *integration* and *conjunction* should be made for at least two reasons: The first is that integration is at the level of our mental representations and takes the geometric constraints into account, whereas conjunction is before anything else a logical connective that occurs at the level of language.

The second reason is that although understanding a conjunction always corresponds to making an integration operation, an integration operation corresponds to more than an occurrence of the conjunction *and*. For example with the disjunctive Sentence 8 above, although the connective *or* links the two pieces of information, the resulting set of models contains a spatial model with two types (i.e. the third one with $\{\sigma^1, \sigma^2\}$). For this third model, the reasoner has carried out an integration operation of the two pieces of information into a single model, but it would be confusing to name this psychological operation 'a conjunction' in this case, since we are analyzing a disjunction. Therefore, the integration operation corresponds to all the cases - we could say truth table lines - where a connective allows the two pieces of information it links to be possible at the same time: the connective *and*, but also the third model of *or* and so on. However, it only resembles the conjunction in a fortuitous way, and they should not be confused with one another.

This distinction made, it is clear that geometric constraints apply to all the spatial

models that compose a set of spatial models, and not simply to premises that are linked by a conjunction. For example, a reasoner that constructs spatial models for Sentence 8 is perfectly able to deduce the new information within the third model that "Circle A is to the left of Circle C" (θ^3). This inference corresponds to the cognitive potential of FREE RIDE permitted by the consequential constraint $\{\sigma^1, \sigma^2\} \vdash \sigma^3$, where σ^3 means that the circle labeled 'A' must be to the left of the circle labeled 'C'. In a similar way, the constraint of *inconsistencies* also applies to these sets, and the cognitive potential of CONSISTENCY CHECK can also be used.³

Equipped with these theoretical clarifications, the main idea of the two sections that follow is to show how *consequential constraints* and *inconsistencies* play a role in the way a reasoner treats a connective in reasoning.⁴ More specifically, we will show that geometric constraints can modify the spatial models of a set in order to make them correspond to the spatial models naturally constructed by a reasoner. The first section illustrates this claim through the *modulation* process used in the theory of mental models, and the second chapter by means of logical principles mainly discussed in the field of conditional logic.

³ Note that, interestingly, a consequential constraint can even be applied on a spatial model that only supports a single type since the converse relation always follows from the declaration of a relation (e.g. in the spatial model constructed for "A is to the left of B", one can always deduce the new information that "B is to the right of A", since a consequential constraint exists between the two types). On the contrary, the constraint of inconsistency requires at least two types in the spatial model, and a piece of information alone can never be inconsistent.

⁴ Note that neither Shimojima, nor the theories of Knauff or Ragni discuss connectives. Moreover, investigating the effects of constraints during a process of reasoning is not a topic directly addressed by Shimojima, who mainly keeps a synthetic and static point of view on constraints (although see (Shimojima and Katagiri, 2013)). In fact, the diagrams used by Shimojima can be seen as gradually constructed by the reasoner in our approach, since we are using linguistic descriptions that impose a sequential presentation of the information.

10.2 Connectives in the Construction Phase

10.2.1 The *Modulation Process* in the Theory of Mental Models

Within the model-based approach, it is typically the role of the classical theory of mental models presented in the first part to study how connectives are understood psychologically. The set of mental models predicted for each connective has been shown in Table 2.1 of this first part.

Since one of the original goals of the theory of mental models is to take into account the semantics of sentences, a mechanism enables their meaning to modify the set of mental models predicted. This mechanism is called *modulation* and is described by the authors as "the process in the construction of models in which content, context, or knowledge can prevent the construction of a model and can add information to a model" (Johnson-Laird et al., 2015, p.2). In order to make modulation possible, the authors create a database with all the impossible models of a problem prior to its analysis. Since the theory represents models by means of words, impossible models simply correspond to a tuple in the database. During the construction phase, if a model turns out to be in the database of impossible models, it is simply removed from the set of predicted models. Table 10.1 presents some examples commonly used to show how modulation changes the set of predicted models for a disjunctive connective (Quelhas and Johnson-Laird, 2017).⁵

By means of several experiments, authors have confirmed that people indeed reason based on the set of modulated models of Table 10.1 instead of keeping an unmodified set of spatial models (Quelhas and Johnson-Laird, 2017). Despite the correctness of this prediction, the use of a database to obtain such a modulated set

⁵ Note that modulation is also used for other connectives such as the conditional, but we do not include such examples here since we dedicate a section to them afterwards.

Table 10.1. Examples of modulated sentences from (Quelhas and Johnson-Laird, 2017).

Assertion	Modulated models		Descriptions and explanations
<i>Ana is in Portugal or Rui is in Spain</i>	Ana in Portugal	Rui in Spain	Modulation does not block any model.
	Ana in Portugal	Rui in Spain	
<i>Ana is in Portugal or she is in Spain</i>	Ana in Portugal	Ana in Spain	The disjunction becomes exclusive by modulation.
<i>Ana is in Madrid or she is in Spain</i>	Ana in Madrid	Ana in Spain	'Ana is in Spain' is always true by modulation.
		Ana in Spain	

cannot be considered as a real psychological operation. This artificial aspect of the method is acknowledged by the authors of the theory who say that "this mechanism is more akin to a conventional search through a lookup table rather than a realistic retrieval from semantic memory" (Khemlani et al., 2018, p.23).

10.2.2 Replacing the Modulation Process by Geometric Constraints

In line with the criticisms already put forward in the first part, we claim that such an artificial process is required due to the incapacity of the theory to characterize models in a relevant way, that is, by taking into account the specific features on which our reasoning is based. This criticism seems to be shared by other approaches from the literature, especially when they go beyond the atomistic view of mental models and formalize aspects of their content that are significant in reasoning (Koralus and Mascarenhas, 2013).

When the relevant features of our reasoning turn out to be the spatial aspects of our mental representations, we believe EPRISM to be the right kind of framework to explain our reasoning. The linguistic module, informed by theories of cognitive linguistics, is able to judiciously decide if a sentence triggers a spatial model, even if it does not seem spatial at first sight. This is exactly what would happen for the examples of Table 10.1. Indeed, the image schema theory has shown that such sentences are understood through the image schema of CONTAINMENT. Let us

consider the second sentence of Table 10.1:

(9) Ana is in Portugal (θ^4) or she is in Spain (θ^5).

Based on the facts highlighted above, the linguistic module maps the two pieces of information on the qualitative relations *inside*($Ana_P, Portugal_C$) for θ^4 , and *inside*($Ana_P, Spain_C$) for θ^5 .⁶ Additionally, the linguistic module adds the piece of information from background knowledge that "Portugal and Spain are two adjacent countries" (θ^6), which corresponds to the relation *extConnected*($Portugal_C, Spain_C$). These three relations compose the list of triplets that is sent to the controller.

Since the sentence is a disjunction, the set of spatial models is supposed to be $\{\{\sigma^4, \sigma^6\}, \{\sigma^5, \sigma^6\}, \{\sigma^4, \sigma^5, \sigma^6\}\}$. However, a constraint of inconsistency occurs in the spatial information since a point cannot be in two externally connected circles at the same time (i.e. $\{\sigma^4, \sigma^5, \sigma^6\} \vdash \emptyset$). When this constraint is applied to the set of spatial models, it turns out that the third model presents the set of types corresponding to this inconsistency. For this reason, this set is removed from the set of predicted spatial models, which is now reduced to $\{\{\sigma^4, \sigma^6\}, \{\sigma^5, \sigma^6\}\}$. Eventually, these two models correspond the ones predicted in Table 10.1.

The same kind of process can be applied on the third sentence, except that it is not an inconsistency that removes spatial models this time, but a consequential constraint that adds pieces of information to them. The sentence considered is:

(10) Ana is in Madrid (θ^7) or she is in Spain (θ^5).

With the background knowledge that "Madrid is in Spain" (θ^8), the original set of spatial models created is $\{\{\sigma^7, \sigma^8\}, \{\sigma^5, \sigma^8\}, \{\sigma^7, \sigma^5, \sigma^8\}\}$. However, a consequential

⁶ Recall that the indexes 'P' and 'C' stand for 'Point' and 'Circle'.

constraint that corresponds to $\{\sigma^7, \sigma^8\} \vdash \sigma^5$ occurs. Indeed, if a point is inside a circle labeled 'Madrid', and this circle is inside a circle labeled 'Spain', then the point must also be in this circle labeled 'Spain'. Each spatial model in the set must be updated by this constraint, which has the effect of transforming the first spatial model $\{\sigma^7, \sigma^8\}$ into the spatial model $\{\sigma^7, \sigma^8, \sigma^5\}$. Since this new spatial model becomes similar to the third one, they can be merged, and only the final set of spatial models is produced: $\{\{\sigma^5, \sigma^8\}, \{\sigma^7, \sigma^8, \sigma^5\}\}$. Once again, this final set of spatial models corresponds to the one predicted in Table 10.1.

In conclusion, there is no need for a database containing the impossible models encoded prior to the analysis of the sentences. The translation based on image schemas and the geometric constraints inherently respected by ASPMT(QS) suffice to immediately produce the right set of spatial models. In this respect, it should be clear that the application of constraints is presented sequentially here, but it is only for the purpose of the argumentation. In EPRISM, these constraints are inherently carried out by the computational model, that would not even be able to produce a model that is inconsistent, or in which information following from consequential constraints is missing. It is precisely because the same laws apply to the models created by ASPMT(QS) and to the human capacity of spatial reasoning that we believe EPRISM to be a more adequate approach for psychology.

The effects of inconsistencies and consequential constraints were only studied for the construction of models here. In the last section, their effects on inferences are investigated.

10.3 Reasoning with Connectives

10.3.1 Some Valid Arguments In Classical Logic

One of the advantages of logical systems is to account for the validity of an argument only based on its logical form. For example, the two logical forms of *Disjunctive Syllogism* and *Modus Ponens* presented in Table 10.2 (second column) are valid in classical logic (third column).⁷ Arguments 11 and 13 (first column) have these logical forms, and we can reasonably consider them as acceptable (fourth column).

Table 10.2. All the examples have a logical form (second column) that is valid in classical logic (third column). However, Arguments 11-13-15 are intuitive, whereas Arguments 12-14-16 go against our intuition (fourth column).

Arguments	Logical form	Validity	Acceptability
Disjunctive Syllogism			
11 Circle A is to the left or to the right of Circle B. Circle A is not to the left of Circle B, so Circle A is to the right of Circle B.	$\phi \vee \psi, \neg\phi \models \psi$	Valid	Accepted
12 Circle A is inside Circle B or it is smaller than Circle B. Circle A is not smaller than B, so Circle A is inside B.	$\phi \vee \psi, \neg\phi \models \psi$	Valid	Refused
Modus Ponens			
13 If I weigh more than 150 kg, I weigh more than 100 kg. I weigh more than 150 kg, so I weigh more than 100 kg.	$\phi \rightarrow \psi, \phi \models \psi$	Valid	Accepted
14 If I weigh more than 150 kg, I weigh less than 25 kg. I weigh more than 150 kg, so I weigh less 25 kg.	$\phi \rightarrow \psi, \phi \models \psi$	Valid	Refused
True Consequent			
15 John is in London now. So, if John is in England now, John is in London now.	$\phi \models \psi \rightarrow \phi$	Valid	Accepted
16 John is in London now. So, if John is in Paris now, John is in London now.	$\phi \models \psi \rightarrow \phi$	Valid	Refused

Modus ponens involves the conditional sentence "If I weigh more than 150 kg, I weigh more than 100 kg", in which the first part is called the *antecedent*, and the second

⁷ However, see (McGee, 1985) for modus ponens and (Priest, 2008; Priest et al., 2018) for disjunctive syllogism.

part is the *consequent*. Undoubtedly, conditional sentences, whether composing a sentence alone or embedded with other logical connectives, are a central topic of investigation into logic and in the psychology of reasoning (Douven, 2016; Égré and Rott, 2021). In this respect, we only consider non-embedded connectives in this section, mainly because embedded connectives add a level of complexity for their analysis that exceeds the scope of this dissertation. Indeed, a connective that applies to another one does no longer state possibilities between pieces of information, but relates to the evaluation of this other connective. Knowing how their interconnection should be defined is often the source of debates in the literature, as is the case for negated conditional sentences such as "it is not the case that if Mary comes to the party, then Peter will come too". For the interaction between these two connectives, a large amount of logical and psychological studies have been made (Handley et al., 2006; Cantwell, 2008; Espino and Byrne, 2012; Égré and Politzer, 2013), such as testing which sentence is considered equivalent to a negated conditional sentence (Olivier, 2019; Skovgaard-Olsen et al., 2019). Although these questions about embedded connectives will not be opened here, they represent an important aspect of our comprehension of language that can certainly not be ignored in the perspective of developing EPRISM to more than simple sentence problems.

As a third example of valid logical form, let us consider Argument 15 in Table 10.2. Contrary to disjunctive syllogism and modus ponens, this logical form, called *True Consequent*, is generally less intuitive (Égré and Rott, 2021). However, classical logic validates it (third column), and we believe the particular instance Argument 15 to be also accepted (fourth column).

Despite the fact that disjunctive syllogism, modus ponens and true consequent are valid in classical logic, it seems possible to find instances of them that are intuitively rejected. For example, few changes on Arguments 11, 13 and 15 suffice to

create Arguments 12, 14 and 16 presented in Table 10.2. Although the three pairs of arguments have the same logical form (second column), it is clear that Arguments 12, 14 and 16 are intuitively rejected.⁸ However, since they are also valid in the logical system, these arguments can be considered problematic. In what follows, we suggest an hypothesis to account for the reasons why these arguments are problematic.

10.3.2 The Satisfiability Requisite

How can we explain the contrast between instances that are recognized to be valid unproblematically, and instances that are problematic? A possible hypothesis to account for the reasons why Arguments 12, 14 and 16 are problematic is the following:⁹

Satisfiability requisite: independently of spatial reasoning, an argument is judged problematic when there is no way in which the premises, the conclusion, or both, can be satisfied.

This hypothesis enables us to shed light on the problematic aspect of Arguments 12, 14 and 16. Indeed, in each of these arguments, an impossibility prevents the reasoner from constructing a mental representation on which judging the inference can be done. This mental representation cannot be constructed because two spatial pieces of information cannot be satisfied at the same time. Due to this impossibility, the reasoner must consider the argument odd and reject it. At the level of logical systems, issues arise when they are not expressive enough to track this impossibility of constructing a model, and furthermore, validate the inference.¹⁰

⁸ Unlike in (Olivier, 2019), we did not collect data to confirm these judgements, but they seem very clear by introspection and based on the judgements of the people with whom we discussed these examples.

⁹ This hypothesis is from P. Égré.

¹⁰ It is possible to find several names in the literature for arguments that are problematic, and yet, are

In what follows, we analyze Arguments 12, 14 and 16 one after the other, and show that a spatial approach enables us to detect this impossibility of constructing a model. Once again, this detection is enabled by highlighting the consequential constraints and inconsistencies that apply to our set of spatial models constructed for each argument.

10.3.3 Accounting For Problematic Arguments

Clearly, Arguments 12, 14 and 16 are problematic because a spatial inconsistency prevents the reasoner from constructing a model. Let us demonstrate how this inconsistency occurs in each argument.

Disjunctive Syllogism

- (12) Circle A is inside Circle B (θ^9) or it is smaller than Circle B (θ^{10}).
Circle A is not smaller than B (θ^{11}), so Circle A is inside B (θ^9).

Since Argument 12 is a disjunction, the set of spatial models is $\{\{\sigma^9\}, \{\sigma^{10}\}, \{\sigma^9, \sigma^{10}\}\}$. In the two spatial models supporting the type that a circle labeled 'A' is inside a circle labeled 'B' (σ^9), a consequential constraint forces them to also support the type that the circle labeled 'A' is smaller than the circle labeled 'B' (σ^{10}) (i.e. $\sigma^9 \vdash \sigma^{10}$). By updating the spatial model $\{\sigma^9\}$ accordingly, only the two different spatial models $\{\{\sigma^{10}\}, \{\sigma^9, \sigma^{10}\}\}$ remain. The next information is that "Circle A is not smaller than B" (θ^{11}) and each remaining model should be updated with this type. However, θ^{11} is also a negation, which can be considered as producing an inconsistency with its positive form: a circle cannot be smaller and not smaller than another at the same

valid in classical logic. For example, *repugnant validities* are defined as "classically valid inference patterns that reasoners often reject" (Mascarenhas, 2014, p.58). For the specific case of conditionals, some of these validities are named the *paradoxes of material implication* (Égré and Rott, 2021, p.3).

time (i.e. $\{\sigma^{10}, \sigma^{11}\} \vdash \emptyset$). Since the two remaining spatial models contain σ^{10} , they have to be updated by σ^{11} , but considering the inconsistency $\{\sigma^{10}, \sigma^{11}\} \vdash \emptyset$, no spatial model remains possible. As a result, the reasoner cannot create any spatial model to evaluate the argument and consider it problematic.¹¹

Modus Ponens

- (14) If I weigh more than 150 kg (θ^{12}), I weigh less than 25 kg (θ^{13}). I weigh more than 150 kg (θ^{12}), so I weigh less than 25 kg (θ^{13}).

As already said, the first part of Argument 14 is a conditional, where the left element (i.e. "If I weigh more than 150 kg") is called the *antecedent*, and the right element (i.e. "I weigh less than 25 kg") is the *consequent*. To shed light on the way a conditional is mentally understood, Ramsey made the following suggestion in *General Propositions and Causality* (Ramsey, 1929):

If two people are arguing "If p will q ?" and are both in doubt as to p , they are adding p hypothetically to their stock of knowledge and arguing on that basis about q ;[...]. We can say they are fixing their degrees of belief in q given p . (Ramsey, 1929, p.247)

This psychological simulation to assess a conditional is called the *Ramsey test*, and it has been one of the most fertile sources of inspiration for later theories of conditionals (Stalnaker, 1968; Douven and Verbrugge, 2010; Égré and Rott, 2021). According to Ramsey, people start the evaluation of a conditional by "adding p hypothetically to their stock of knowledge". For Argument 14, this stock of knowledge only contains

¹¹ The way a negation should be defined in this analysis is a great topic of investigation. An important question is to know whether people negate the information linguistically, or if they create a spatial model to understand it semantically (e.g. constructing a circle 'A' bigger than a circle 'B' to understand "Circle A is not smaller than Circle B"). Experimental results show that people often construct a spatial model to understand the negated element (Schleipen et al., 2007). Considering this result, it becomes the task of psychologists to define the meaning of each negated relation. The notion of *conceptual distances* that was presented earlier (Hamami et al., 2021) as well as the definition of *conceptual spaces* (Zwarts and Gärdenfors, 2016; Osta Vélez, 2020) will undeniably be of a great help in this endeavor (e.g. the negation of a relation might be calculated as the farthest conceptual distance).

that "150 kg is more than 25 kg". Based on the image schema of SOURCE-PATH-GOAL, this knowledge can be represented by a spatial model containing a mental number line directed to the right, on which a point labeled '150' is placed to the right of a point labeled '25' (σ^{14}). When the antecedent "I weigh more than 150kg" (θ^{12}) is added to our spatial model, this corresponds to add a point labeled 'Me' to the right of the point labeled '150' (σ^{12}). By placing the point labeled 'Me' to the right of '150', we also place it to the right of '25' (σ^{15}) due to the consequential constraint $\{\sigma^{12}, \sigma^{14}\} \vdash \sigma^{15}$. Therefore, the spatial model becomes $\{\sigma^{12}, \sigma^{14}, \sigma^{15}\}$.

When the consequent "I weigh less than 25 kg" (σ^{13}) is read, the point labeled 'Me' should be placed to the left of '25' (σ^{13}). However, this last requirement is not possible since an inconsistency prevents a point labeled 'Me' from being both to the left *and* to the right of a point labeled '25' (i.e. $\{\sigma^{13}, \sigma^{15}\} \vdash \emptyset$). Since no spatial model can be constructed for these pieces of information, the argument becomes problematic in virtue of the Satisfiability Constraint.¹²

True Consequent

For the case of Argument 16, the information is understood based on the image schema of CONTAINMENT, which means that the background knowledge corresponds to a spatial model containing two disconnected circles labeled 'London' and 'Paris'. The inconsistency arises more directly for this argument since the consequent of the conditional is not even needed to produce it. Indeed, adding the information that "John is in Paris" to our spatial model suffices to create the inconsistency with the information that he is also in London: a point labeled 'John' cannot be in the

¹² In fact, this small simulation of integrating the antecedent to our stock of knowledge is similar to the one triggered by a modus ponens argument, where the antecedent of the conditional is asserted as a fact. For this reason, applying the Ramsey test on the conditional of Argument 13 and the inference of modus ponens in this same argument might be seem a bit redundant. However, they enable us to discuss our approach on principles that are commonly used in the literature.

circle labeled 'Paris' and in the circle labeled 'London' at the same time in our spatial model.

Besides the fact that Argument 16 is problematic, its logical form is not intuitive from the start. Indeed, even the valid version of true consequent (i.e. Argument 15) seems odd. This example of true consequent has not been presented with the purpose of solving all its difficulties since this last problem follows from the classical definition of the conditional connective, and not from its spatial information. Instead, we argue that our spatial analysis enables us to single out the problem that no spatial model can be created, from the second difficulty, which is that no meaningful connection exists between the premises of the argument and its conclusion. Despite the inadequacy of our approach to explain this second problem for the case of Argument 15,¹³ a spatial analysis might become useful to account for the acceptance of other conditionals, as is presented in what follows.

10.3.4 A Case of Acceptability

Let us consider the accepted instance of modus ponens in Argument 13:

- (13) If I weigh more than 150 kg (θ^{12}), I weigh more than 100 kg (θ^{17}).
I weigh more than 150 kg (θ^{12}), so I weigh more than 100 kg (θ^{17}).

Combined with the background knowledge that '150' is placed to the right of a point labeled '100' (σ^{18}), our mental representation created by the Ramsey test becomes $\{\sigma^{12}, \sigma^{18}\}$. A consequential constraint applies on this set of information because placing a point 'Me' to the right of a point labeled '150', which is itself to the right of a point labeled '100', will undeniably place the point 'Me' to the right of '100'

¹³ A plausible way in which the oddness of Argument 15 can be explained is by using the constraint on the acceptability of indicative conditionals proposed in (Douven, 2016, p.108), namely, the *Evidential Support Thesis*.

(i.e. $\{\sigma^{12}, \sigma^{18}\} \vdash \sigma^{17}$). For this reason, the spatial model constructed corresponds to $\{\sigma^{12}, \sigma^{18}, \sigma^{17}\}$. When the reasoner reads the consequent of the conditional "I weigh more than 100 kg", it turns out that this information is already contained in the spatial model constructed. Indeed, whatever the exact spatial model constructed for the first two pieces of information, the consequent constraint $\{\sigma^{12}, \sigma^{18}\} \vdash \sigma^{17}$ forces σ^{17} to be also supported by the spatial model.

For this reason, the reasoner accepts the conditional based on the idea that $\{\sigma^{12}, \sigma^{18}\}$ necessarily implies σ^{17} . In other words, it is *because* the two pieces of information are supported by the spatial model that the consequent appears in the spatial model. In more general terms, one can even consider the geometric consequential constraint itself (i.e. \vdash) to ground the inferability from the antecedent to the consequent, and it is exactly this inferential connection we want to assert when we accept the conditional sentence (Mill, 1874; Sellars, 1953; Douven et al., 2018).

10.3.5 Concluding Remarks

The main contributions of this part have been: (i) The enumeration of the principles that a linguistic module would have to follow in order to be compatible with the formal characteristics of ASPMT(QS). (ii) The conceptual study of image schemas and the formalization of CONTAINMENT and SOURCE-PATH-GOAL adapted to EPRISM. (iii) The formal derivation of the modulated sets of spatial models, without using an external database containing the impossible models. (iv) The spatial analysis of arguments in order to explain why they become problematic. And finally (v), the possibility to consider geometric consequential constraints for grounding the inferential connection of certain conditionals.

Although all the functions described above remain at a conceptual level, several

elements already exist to initiate their implementation: first and foremost, the large amount of natural language processing systems will considerably help the implementation of the linguistic module. As we have seen though, some adjustments are required for the specific purposes of EPRISM. Secondly, even if a detection system for image schemas in natural language does not exist in the literature ([Hedblom, 2020](#)), many theories, that are not necessarily based on the image schema theory, obtain successful results for the same goals ([Mani and Pustejovsky, 2012](#)).

Finally, it is worth mentioning that ASPMT(QS) supports logical connectives (i.e. *and*, *or*, *not*, and *if* correspond to the symbols ‘&’, ‘|’, ‘not’ and ‘<-’ in an input program). However, it is necessary to point out that these connectives are at the level of constraints in the program, whereas EPRISM will have to account for connectives at the level of models. More specifically, it is among the models handled by the controller that a psychological process of connectives will have to be implemented. But it should be kept in mind that a fully developed implementation of this process will only be possible based on linguistic studies of connectives. Ragni and Knauff remain silent about these topics, but the theory of mental models, as well as all the logical approaches of connectives, will clearly be of help in this endeavor.

Chapter 11

General Conclusion

11.1 To Sum Up

In this dissertation, the conceptual and computational elements necessary for extending the computational model PRISM to all kinds of reasoning have been presented. After some chapters about the main principles and theories of spatial reasoning, Chapter 5 presented the inferential argument that enabled us to state that spatial models must be diagrammatic in nature. This hypothesis opened the possibility to formalize spatial models in an exhaustive way. Then, several arguments were given to refuse the idea that spatial models are clear images in the mind. In Chapter 6, a representation system was described in order to characterize spatial models in a psychological way, but moreover, to link our mental representations with the sentences that are used to create them.

In Chapter 7, the adequacy of the computational tool ASPMT(QS) for implementing the representation system for psychology was demonstrated. This enabled us to build a possible architecture for EPRISM in Chapter 8, and to model in detail the reasoning process of two examples of spatial reasoning: one example from PRISM

and one involving the topology of circles.

Chapter 9 detailed the requirements that a linguistic module has to fulfill in order to be adapted to EPRISM, but also, the possibility of formalizing image schemas into our approach. Finally, in Chapter 10, the formalization of image schemas were used to shed new light on the study of connectives, particularly for the modulation process of the theory of mental models and the logic of conditionals.

In summary, all the conceptual and computational elements necessary for the creation of EPRISM have been given: a clear description of mental models, the concept of default specifications, the representation system for psychology and the adequacy of ASPMT(QS) to implement this system. Although the concrete creation of EPRISM remains to be done, its elements can already be used to support the study of psychological capacities that are still unclear, such as how a reasoner understands a complex relation by means of attentional shifts, or carries out a variation phase. As already mentioned, the more advanced version of ASPMT(QS) for spatio-temporal reasoning constitutes a promising option that confirms the possible creation of EPRISM ([Schultz et al., 2018](#)).

To conclude this dissertation, it might be relevant to discuss the extent to which EPRISM can be used to explain cognition. In particular, a question arose between the lines throughout the chapters, and it probably became more apparent at the end of Chapter 10: is deductive reasoning spatial? In what follows, some works that explore the plausibility of this question are briefly presented in order to highlight the relevance of the topic.

11.2 Deductive Reasoning as Spatial Reasoning

Without naming it, this dissertation clearly endorses a pluralistic view of mental representations, that is, more than only one type of representation exists in the mind. This idea of heterogeneity of mental representations seems more and more investigated by cognitive scientists ([Horst, 2016](#)), and defining all their types is even seen as the next step of cognitive science by some authors:

Is all information stored in propositional, language-like, symbolic internal representations, or can humans use at least two different types of representations (and possibly many more)? Here, in historical context, we describe recent evidence that humans do not always rely on propositional internal representations but, instead, can also rely on at least one other format: depictive representation. We propose that the debate should now move on to characterizing all of the different forms of human mental representation. ([Pearson and Kosslyn, 2015](#), abstract).

The content of this dissertation might be seen as a step forward in this characterization, especially by defining mental representations for spatial cognition more concretely. If we accept that spatial models constitute a distinct form of representations in the mind, and given the universality of the constraints that apply to them (i.e. geometric constraints apply objectively to our capacity to think spatially), one can legitimately wonder to which extent spatial models can be considered as a possible foundation for deductive reasoning. In this view, the foundation of formal theories could be seen as a generalization of FREE RIDE and INCONSISTENCIES that are first made on basic spatial models. The idea of grounding deductive capacities on spatial reasoning, of course not expressed with these words, is already found in Peirce's writings:

Deduction is that mode of reasoning which examines the state of things asserted in the premisses, forms a diagram of that state of things, perceives in the parts of that diagram relations not explicitly mentioned in the premisses, satisfies itself by mental experiments upon the diagram that these relations would always subsist, or at least would do so in a certain proportion of cases, and concludes their necessary, or probable, truth. ([Peirce, 1938](#), 1.66)

The possibility of reconstructing logical systems through a generalization of FREE RIDE and INCONSISTENCIES is an ongoing investigation which has already produced successful results ([Stapleton et al., 2018](#); [Sáenz-Ludlow, 2018](#); [Smessaert et al., 2020](#)).¹ For example, it has been shown that "given any finite collection of set-theoretic sentences, S , from the class just described, there exists an Euler diagram, d , from which any set-theoretic sentence, σ , inferrable from S can be observed [obtained by FREE RIDE]" ([Stapleton et al., 2018](#), p.2). Of course, these studies require accepting the translation into diagrams, and moreover for our claims, that we have these same representations in the mind. However, the intuitive character of such translations has been highlighted in several cases with the image schema theory, such as in our comprehension of sets by means of CONTAINMENT ([Lakoff and Núñez, 2000](#)).² Although it is clear that some symbolic formulas do not accept a diagrammatic representation, the investigation seems promising. EPRISM, combined with the reasoning possibilities offered by ASPMT(QS), will definitely be part of such research.

¹ All these articles do not use the same terminology, but they aim at the same goal, namely, investigating how much of logic can be explained through inferences made on diagrams.

² An interesting question to investigate in line with these ideas is the extent to which Russell's paradox for set-theory can be considered as a case of cognitive INCONSISTENCY.

Appendices

Appendix A

Installation of ASPMT(QS)

The computational tool ASPMT(QS) is distributed in the form of a Docker image and therefore requires the installation of the Docker application. The Docker desktop application is freely available on the website <https://www.docker.com>. Once Docker is installed, interactions with it are made through the terminal of your computer. The Docker image of ASPMT(QS) can be obtained by the command "docker pull spatialreasoning/aspmtqs" in the terminal, or by following the instructions available at <https://hub.docker.com/r/spatialreasoning/aspmtqs>. This website also contains some examples that are already implemented in ASPMT(QS).

The command below enables the user to give a document '.txt' to ASPMT(QS). For doing so, replace the first underlined part by the path to the folder (without the name of the file at the end) and the second underlined part by the name of the file. Some detailed input programs are given in the appendices of (Walega et al., 2017), and Table 7.1 (left) in Part 2 of this dissertation can be used as an input program too.

```
docker run -v /path/to/your/folder/ :/data/ spatialreasoning/aspmtqs  
-p /data/nameOfTheFile.txt
```

Appendix B

Basic Principles and Logical Foundation of ASPMT(QS)

The framework Answer Set Programming Modulo Theories for the Qualitative Spatial domain ASPMT(QS) mainly results from the combination of research in Answer Set Programming (ASP), with the solvers used for Satisfiability Modulo Theories (SMT) problems. The specific developments of these components can be seen in Figure B.1.

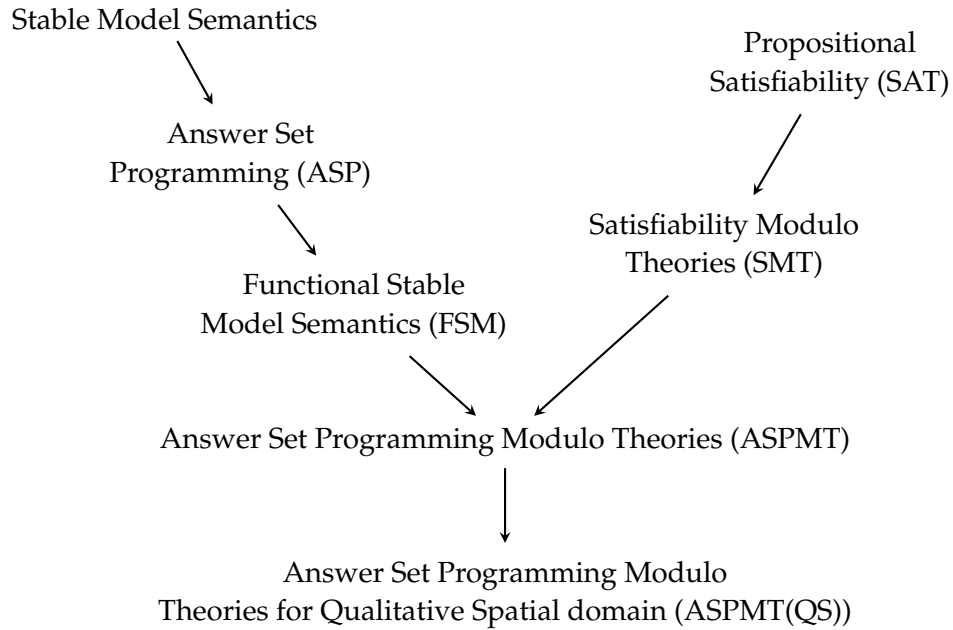


Fig. B.1. From the top to the bottom, development of the different techniques that permit the construction of ASPMT(QS).

Each branch of this development is described in what follows.

The Answer Set Programming (ASP)

Answer Set Programming (ASP) is a form of declarative programming based on the stable model semantics (M. Gelfond, 1988). The theoretical particularity of this semantics is the non-monotonic inference of *negation as failure*, which consists in deriving *not p* from the failure to derive *p*. Let us consider the propositional formula below on the left, and its translation into a declarative form on the right (i.e. also called a *program*). The program of a formula is obtained by associating each conjunct to a line, inverting the "not" with the negation, and the " \rightarrow " with the " \leftarrow ". A line without a symbol " \leftarrow " is considered a *fact*. A line of the form $\phi \leftarrow \psi$ is a *rule*, where ϕ is its *head*, and ψ , which is a conjunction of formula(s), is its *body*. As can be seen, the program below has one fact and two rules.

$p \wedge ((p \wedge q) \rightarrow r) \wedge ((p \wedge \neg q) \rightarrow s)$	p
	$r \leftarrow p, q$
	$s \leftarrow p, \text{not } q$

Classical propositional logic produces four models for this formula, which are represented in Table B.1. The *stable model* of a program is the smallest set (i.e. minimal relative to set inclusion) that satisfies all its facts and rules. Compared with the classical approach, it means that only the model in bold of the fourth line of Table B.1 is a stable model.

The calculation of such a result goes as follows: p is obtained since it is a fact. Since q cannot be obtained in the program, it is considered false due to the negation as failure principle. Therefore, despite the truth of p , the body p, q fails since q is considered false. It follows that r is false and the rule $r \leftarrow p, q$ is satisfied. On the contrary, the second rule $s \leftarrow p, \text{not } q$ allows to obtain s , since p is true from the first fact, and $\text{not } q$ is true from the absence of q . The model $\{p, s\}$ is the only resulting model.

Table B.1. Models of the formula $p \wedge ((p \wedge q) \rightarrow r) \wedge ((p \wedge \neg q) \rightarrow s)$

p	q	r	s
1	1	1	1
1	1	1	0
1	0	1	1
1	0	0	1

As can be deduced, the set of stable models will always be a subset of the set of classical models, and generally the one with the smallest cardinality.¹ Due to these interesting properties, the procedures for finding stable models are considerably less demanding than those used to find classical models.

¹ Since stable models are obtained on the basis of set inclusion, the stable models of a program are not always those with the smallest cardinality. For example, the program $\{p \leftarrow q, q \vee r\}$ has the two models $\{r\}$ and $\{p, q\}$, even if the latter has more elements than the former.

Describing a situation by means of facts and rules constitutes the declarative programming paradigm in computer science. The language PROLOG is probably the most famous language in this paradigm. Based on the stable model semantics, ASP is often considered as a convenient alternative to certain limitations encountered in PROLOG. An integrated development environment is accessible online to test the language ASP with some already encoded examples², and a first practice book has also been released recently (Lifschitz, 2019).

One limitation of the stable model semantics, and consequently ASP, is the difficulty to represent conditions that can change during the processing (e.g. representing the level of water in a tank that fluctuates over time). A step towards a solution was first made in (Ferraris et al., 2007, 2011), and then generalized in the so-called Functional Stable Model Semantics (FSM) (Bartholomew and Lee, 2019). As will be seen, this generalization plays a central role for the combination that made ASPMT(QS) possible.

The Solvers for First-Order Theories

The other series of discoveries (i.e. the right branch of Figure B.1) permitting the fruitful combination mentioned above comes from techniques for decision problems of satisfiability. Boolean satisfiability problem (SAT) is the problem of determining whether a bivalent propositional formula is satisfiable or not.

The Satisfiability Modulo Theories (SMT) (De Moura and Bjørner, 2009) can be seen as a generalization of SAT since it searches for the satisfiability of formulas of first-order logic. This is generally done with theories of first-order logic that constrain the space of solutions. Many theories of first-order logic have already been implemented in the SMT approach, including linear and non-linear real arithmetic,

² <https://potassco.org/clingo/run/>

arrays, the empty theory and so on. This trend of research covers a huge part of the computer science field currently and gives rise to multiple events and applications in a wide number of domains (De Moura and Bjørner, 2011).

A fruitful Combination that is used for the Spatial Domain

Thanks to the generalization mentioned earlier (Bartholomew and Lee, 2019), it is possible to define Answer Set Programming Modulo Theories (ASPMT) (Bartholomew and Lee, 2014), which enables us to use solvers of SMT (e.g. z3) to check for the satisfiability of a formula under FSM. The architecture of ASPMT is shown in Figure B.2.

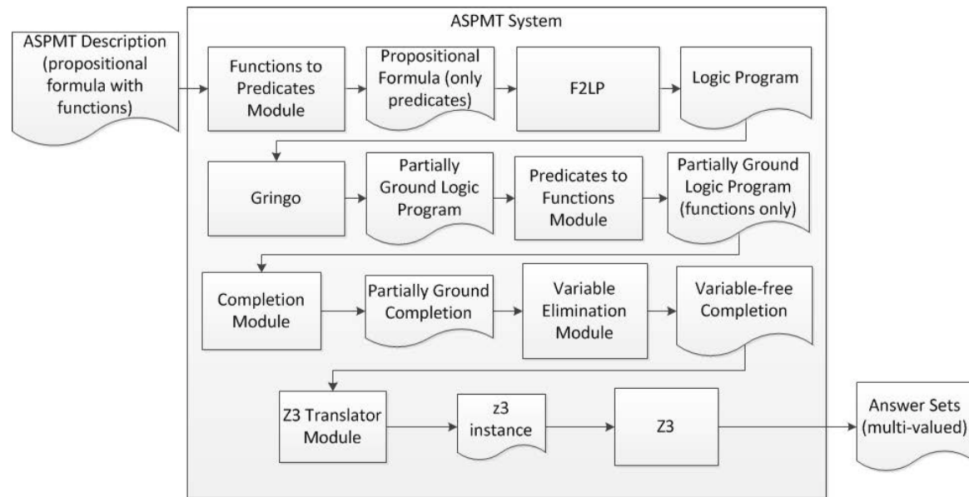


Fig. B.2. Architecture of the different techniques composing the ASPMT system. Figure from (Bartholomew and Lee, 2014, p.538)

As a final layer of programming, Wałęga, Schultz and Bhatt developed a specific module of ASPMT dedicated to the Qualitative Spatial domain using non-linear arithmetic over reals. This program is called ASPMT(QS) (Wałęga et al., 2017) and

is available online on the website of the Spatial Reasoning group.³ Complementary information for the installation is given in Appendix A of this dissertation.

³ <http://hcc.uni-bremen.de/spatial-reasoning/>

Bibliography

- K. L. Alfred, A. C. Connolly, J. S. Cetron, and D. J. Kraemer. Mental models use common neural spatial structure for spatial and abstract content. *Communications biology*, 3(1):1–11, 2020.
- P. Amorapanth, A. Kranjec, B. Bromberger, M. Lehet, P. Widick, A. J. Woods, D. Y. Kimberg, and A. Chatterjee. Language, perception, and the schematic representation of spatial relations. *Brain and language*, 120(3):226–236, 2012.
- M. Aurnague and L. Vieu. A three-level approach to the semantics of space. In *The Semantics of Prepositions*, pages 393–440. De Gruyter Mouton, 2011.
- M. Aurnague, M. Hickmann, and L. Vieu. *The categorization of spatial entities in language and cognition*, volume 20. John Benjamins Publishing, 2007.
- M. Bartholomew and J. Lee. System aspmt2smt: Computing aspmt theories by smt solvers. In *European Workshop on Logics in Artificial Intelligence*, pages 529–542. Springer, 2014.
- M. Bartholomew and J. Lee. First-order stable model semantics with intensional functions. *Artificial Intelligence*, 273:56–93, 2019.
- J. Barwise and J. Seligman. *Information flow: the logic of distributed systems*. Cambridge University Press, 1997.

- G. Berkeley. *A treatise concerning the principles of human knowledge*. JB Lippincott & Company, 1881.
- M. Bhatt, J. H. Lee, and C. Schultz. CLP (QS): A declarative spatial reasoning framework. In *International Conference on Spatial Information Theory*, pages 210–230. Springer, 2011.
- G. Borst and S. M. Kosslyn. Visual mental imagery and visual perception: Structural equivalence revealed by scanning processes. *Memory & cognition*, 36(4):849–862, 2008.
- M. D. Braine and D. P. O’Brien. *Mental logic*. Psychology Press, 1998.
- J. Cantwell. The logic of conditional negation. *Notre Dame Journal of Formal Logic*, 49(3):245–260, 2008.
- J. Carter. “Free rides” in mathematics. *Synthese*, pages 1–24, 2021.
- A. G. Cohn, B. Bennett, J. Gooday, and N. M. Gotts. Qualitative spatial representation and reasoning with the region connection calculus. *GeoInformatica*, 1(3):275–316, 1997.
- K. Craik. W. *The Nature of Explanation*. Cambridge UK: Cambridge University Press, 1943.
- L. De Moura and N. Bjørner. Satisfiability modulo theories: An appetizer. In *Brazilian Symposium on Formal Methods*, pages 23–36. Springer, 2009.
- L. De Moura and N. Bjørner. Satisfiability modulo theories: introduction and applications. *Communications of the ACM*, 54(9):69–77, 2011.
- R. Dechter and al. *Constraint processing*. Morgan Kaufmann, 2003.

- S. Dehaene. Varieties of numerical abilities. *Cognition*, 44(1-2):1–42, 1992.
- S. Dehaene, S. Bossini, and P. Giraux. The mental representation of parity and number magnitude. *Journal of experimental psychology: General*, 122(3):371, 1993.
- D. C. Dennett. *Content and consciousness*. Routledge, 2002.
- I. Douven. *The epistemology of indicative conditionals: Formal and empirical approaches*. Cambridge University Press, 2016.
- I. Douven and S. Verbrugge. The Adams family. *Cognition*, 117(3):302–318, 2010.
- I. Douven, S. Elqayam, H. Singmann, and J. van Wijnbergen-Huitink. Conditionals and inferential connections: A hypothetical inferential theory. *Cognitive Psychology*, 101:50–81, 2018.
- F. I. Dretske. *Knowledge and the Flow of Information*. MIT Press, 1981.
- F. Dylla, J. H. Lee, T. Mossakowski, T. Schneider, A. V. Delden, J. V. D. Ven, and D. Wolter. A survey of qualitative spatial and temporal calculi: Algebraic and computational properties. *ACM Computing Surveys (CSUR)*, 50(1):1–39, 2017.
- P. Égré. Vague judgment: a probabilistic account. *Synthese*, 194(10):3837–3865, 2017.
- P. Égré and G. Politzer. On the negation of indicative conditionals. In *Proceedings of Amsterdam Colloquium*, pages 10–18. M. Aloni, M. Franke and F. Roelofsen, 2013.
- P. Égré and H. Rott. The Logic of Conditionals. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2021 edition, 2021.
- S. Elqayam. Mental models, model-theoretic semantics, and the psychosemantic conception of truth. *Philosophia scientiae*, 9(2):259–278, 2005.

- O. Espino and R. M. Byrne. It is not the case that if you understand a conditional you know how to negate it. *Journal of Cognitive Psychology*, 24(3):329–334, 2012.
- M. J. Farah, K. M. Hammond, D. N. Levine, and R. Calvanio. Visual and spatial mental imagery: Dissociable systems of representation. *Cognitive psychology*, 20(4): 439–462, 1988.
- P. Ferraris, J. Lee, and V. Lifschitz. A new perspective on stable models. In *IJCAI*, volume 7, pages 372–379, 2007.
- P. Ferraris, J. Lee, and V. Lifschitz. Stable models and circumscription. *Artificial Intelligence*, 175(1):236–263, 2011.
- J. A. Fodor. *The language of thought*. Harvard university press, 1975.
- K. D. Forbus. *Qualitative representations: How people reason and learn about the continuous world*. MIT Press, 2019.
- S. L. Franconeri, J. M. Scimeca, J. C. Roth, S. A. Helseth, and L. E. Kahn. Flexible visual processing of spatial relationships. *Cognition*, 122(2):210–227, 2012.
- C. Freksa. Temporal reasoning based on semi-intervals. *Artificial intelligence*, 54(1-2): 199–227, 1992.
- D. Geeraerts, R. Dirven, J. R. Taylor, R. W. Langacker, and B. Hampe. *From perception to meaning: Image schemas in cognitive linguistics*. Mouton de Gruyter, 2005.
- A. Gerevini and J. Renz. Combining topological and size information for spatial reasoning. *Artificial Intelligence*, 137(1-2):1–42, 2002.
- V. Giardino. Behind the diagrams: Cognitive issues and open problems. In *Thinking with Diagrams*, pages 77–101. Mouton de Gruyter Boston, 2016.

- R. W. Gibbs and H. L. Colston. The cognitive psychological reality of image schemas and their transformations. *Cognitive Linguistics (includes Cognitive Linguistic Bibliography)*, 6(4):347–378, 1995.
- N. Goodman. *Languages of art: An approach to a theory of symbols*. Hackett publishing, 1976.
- L. M. Gurtner, M. Hartmann, and F. W. Mast. Eye movements during visual imagery and perception show spatial correspondence but have unique temporal signatures. *Cognition*, 210:104597, 2021.
- Y. Hamami and J. Mumma. Prolegomena to a cognitive investigation of euclidean diagrammatic reasoning. *Journal of Logic, Language and Information*, 22(4):421–448, 2013.
- Y. Hamami, J. Mumma, and M. Amalric. Counterexample search in diagram-based geometric reasoning. *Cognitive Science*, 45(4):e12959, 2021.
- S. J. Handley, J. S. B. Evans, and V. A. Thompson. The negated conditional: A litmus test for the suppositional conditional? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(3):559, 2006.
- M. M. Hedblom. *Image schemas and concept invention: cognitive, logical, and linguistic investigations*. Springer Nature, 2020.
- W. Hodges. Functional modelling and mathematical models: A semantic analysis. In *Philosophy of technology and engineering sciences*, pages 665–692. Elsevier, 2009.
- S. Horst. *Cognitive pluralism*. Mit Press, 2016.
- A. Inants and J. Euzenat. So, what exactly is a qualitative calculus? *Artificial Intelligence*, 289:103385, 2020.

- P. N. Johnson-Laird. *Mental models*. Harvard University Press, 1983.
- P. N. Johnson-Laird. Imagery, visualization, and thinking. *Perception and cognition at century's end*, pages 441–467, 1998.
- P. N. Johnson-Laird. Models and heterogeneous reasoning. *Journal of Experimental & Theoretical Artificial Intelligence*, 18(2):121–148, 2006.
- P. N. Johnson-Laird. *How we reason*. Oxford University Press, 2008.
- P. N. Johnson-Laird. Models and rational deductions. In R. Viale, editor, *Routledge Handbook of Bounded Rationality*. Routledge, 2020.
- P. N. Johnson-Laird and R. M. Byrne. *Deduction*. Lawrence Erlbaum Associates, Inc, 1991.
- P. N. Johnson-Laird and R. M. Byrne. *Precis of deduction*. *Behavioral and brain sciences*, 16(2):323–333, 1993.
- P. N. Johnson-Laird, V. Girotto, and P. Legrenzi. Reasoning from inconsistency to consistency. *Psychological Review*, 111(3):640, 2004.
- P. N. Johnson-Laird, S. S. Khemlani, and G. P. Goodwin. Logic, probability, and human reasoning. *Trends in cognitive sciences*, 19(4):201–214, 2015.
- G. Juravle, C. Velasco, A. Salgado-Montejo, and C. Spence. The hand grasps the center, while the eyes saccade to the top of novel objects. *Frontiers in psychology*, 6: 633, 2015.
- A. Khan, M. Vasardani, and S. Winter. Extracting spatial information from place descriptions. In *COMP '13: Proceedings of The First ACM SIGSPATIAL International Workshop on Computational Models of Place*, pages 62–69, 2013.

- S. S. Khemlani and P. N. Johnson-Laird. The processes of inference. *Argument & Computation*, 4(1):4–20, 2013.
- S. S. Khemlani, R. M. Byrne, and P. N. Johnson-Laird. Facts and possibilities: A model-based theory of sentential reasoning. *Cognitive Science*, 42(6):1887–1924, 2018.
- K. C. Klauer and Z. Zhao. Double dissociations in visual and spatial short-term memory. *Journal of experimental psychology: General*, 133(3):355, 2004.
- M. Knauff. *Space to reason: A spatial theory of human thought*. MIT Press, 2013.
- M. Knauff. Visualization, reasoning, and rationality. In *International Conference on Conceptual Structures*, pages 3–10. Springer, 2019.
- M. Knauff and E. May. Mental imagery, reasoning, and blindness. *Quarterly Journal of Experimental Psychology*, 59(1):161–177, 2006.
- M. Knauff, R. Rauh, and J. Renz. A cognitive assessment of topological spatial relations: Results from an empirical investigation. In *International Conference on Spatial Information Theory*, pages 193–206. Springer, 1997.
- R. Koenig-Robert and J. Pearson. Why do imagery and perception look and feel so different? *Philosophical Transactions of the Royal Society B*, 376(1817):20190703, 2021.
- P. Koralus and S. Mascarenhas. The erotetic theory of reasoning: bridges between formal semantics and the psychology of deductive inference. *Philosophical Perspectives*, 27:312–365, 2013.
- S. M. Kosslyn, W. L. Thompson, and G. Ganis. *The case for mental imagery*. Oxford University Press, 2006.

- P. Kozak. The diagram problem. In *International Conference on Theory and Application of Diagrams*, pages 216–224. Springer, 2020.
- Y. Kurata. 9-intersection calculi for spatial reasoning on the topological relations between multi-domain objects. In *Workshop on Spatial and Temporal Reasoning*, page 34. Citeseer, 2009.
- B. Laeng, I. M. Bloem, S. D’Ascenzo, and L. Tommasi. Scrutinizing visual images: The role of gaze in mental imagery and memory. *Cognition*, 131(2):263–283, 2014.
- G. Lakoff and M. Johnson. *Metaphors We Live By*. University of Chicago Press, 1980.
- G. Lakoff and R. Núñez. *Where mathematics comes from*, volume 6. New York: Basic Books, 2000.
- P. Legrenzi, V. Girotto, and P. N. Johnson-Laird. Models of consistency. *Psychological Science*, 14(2):131–137, 2003.
- D. N. Levine, J. Warach, and M. Farah. Two visual systems in mental imagery: Dissociation of “what” and “where” in imagery disorders due to bilateral posterior cerebral lesions. *Neurology*, 35(7):1010–1010, 1985.
- V. Lifschitz. *Answer set programming*. Springer Berlin, 2019.
- G. Ligozat and J. Renz. What is a qualitative calculus? a general framework. In *Pacific Rim International Conference on Artificial Intelligence*, pages 53–64. Springer, 2004.
- G. Lupyan. The paradox of the universal triangle: Concepts, language, and prototypes. *Quarterly journal of experimental psychology*, 70(3):389–412, 2017.

- V. L. M. Gelfond. The stable model semantics for logic programming. In R. Kowalski, K. Bowen (Eds.), *Proceedings of the Fifth Logic Programming Symposium*, MIT Press, Cambridge, MA,, pages pp. 1070–1080, 1988.
- J. M. Mandler. On the spatial foundations of the conceptual system and its enrichment. *Cognitive science*, 36(3):421–451, 2012.
- J. M. Mandler and C. P. Cánovas. On defining image schemas. *Language and Cognition*, 6(4):510–532, 2014.
- I. Mani and J. Pustejovsky. *Interpreting motion: Grounded representations for spatial language*. Oxford University Press, 2012.
- S. Mascarenhas. *Formal Semantics and the Psychology of Reasoning: Building new bridges and investigating interactions*. PhD thesis, New York University, 2014.
- F. W. Mast and S. M. Kosslyn. Eye movements during visual mental imagery. *Trends in Cognitive Sciences*, 6(7):271–272, 2002.
- V. McGee. A counterexample to modus ponens. *The Journal of Philosophy*, 82(9): 462–471, 1985.
- J. S. Mill. *A System of Logic*, 8th. New York: Harper and Brothers, 1874.
- M. Mishkin, L. G. Ungerleider, and K. A. Macko. Object vision and spatial vision: two cortical pathways. *Trends in neurosciences*, 6:414–417, 1983.
- K. S. Mix, L. B. Smith, and M. Gasser. *The spatial foundations of cognition and language: Thinking through space*. Oxford University Press, 2010.
- K. E. Moore. *The Spatial Language of Time: Metaphor, metonymy, and frames of reference*, volume 42. John Benjamins Publishing Company, 2014.

- N. Newcombe, J. Huttenlocher, and A. Learmonth. Infants' coding of location in continuous space. *Infant Behavior and Development*, 22(4):483–510, 1999.
- F. Olivier. Testing the equivalence between wide-scope and narrow-scope negations in conditionals. In *Student Session of the 30th European Summer School in Logic, Language, and Information*, pages 190–200. ESSLLI, 2019. URL http://esslli2019.folli.info/wp-content/uploads/2019/08/tentative_proceedings.pdf.
- F. Olivier. A logical framework for spatial mental models. In J. Škilters, N. Newcombe, and D. Uttal, editors, *German Conference on Spatial Cognition XII*, pages 268–280. Springer, 2020.
- M. Osta Vélez. *Inference and the structure of concepts*. PhD thesis, LMU, 2020.
- C. Peacocke. The inaugural address: Analogue content. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 60:1–17, 1986.
- J. Pearson. The human imagination: the cognitive neuroscience of visual mental imagery. *Nature Reviews Neuroscience*, 20(10):624–634, 2019.
- J. Pearson and S. M. Kosslyn. The heterogeneity of mental representation: Ending the imagery debate. *Proceedings of the National Academy of Sciences*, 112(33):10089–10092, 2015.
- C. S. Peirce. *Collected Papers of CS Peirce, 8 Bände*. Harvard University Press, Boston, 1938.
- B. R. Postle, C. Idzikowski, S. D. Sala, R. H. Logie, and A. D. Baddeley. The selective disruption of spatial working memory by eye movements. *Quarterly Journal of Experimental Psychology*, 59(1):100–120, 2006.

- G. Priest. *An introduction to non-classical logic: From if to is*. Cambridge University Press, 2008.
- G. Priest, K. Tanaka, and Z. Weber. Paraconsistent Logic. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2018 edition, 2018.
- Z. W. Pylyshyn. What the mind’s eye tells the mind’s brain: A critique of mental imagery. *Psychological bulletin*, 80(1):1, 1973.
- A. C. Quelhas and P. N. Johnson-Laird. The modulation of disjunctive assertions. *The Quarterly Journal of Experimental Psychology*, 70(4):703–717, 2017.
- M. Ragni and M. Knauff. Cross-cultural preferences in spatial reasoning. *Journal of Cognition and Culture*, 11(1-2):1–21, 2011.
- M. Ragni and M. Knauff. A theory and a computational model of spatial reasoning with preferred mental models. *Psychological review*, 120(3):561, 2013.
- M. Ragni, D. Brand, and N. Riesterer. The predictive power of spatial relational reasoning models: A new evaluation approach. *Frontiers in Psychology*, 12, 2021.
- F. P. Ramsey. General propositions and causality. In D. H. Mellor, editor, *Foundations of Mathematics and other Logical Essays*, pages 237–257. Routledge and Kegan Paul, London, 1929.
- D. A. Randell, Z. Cui, and A. G. Cohn. A spatial logic based on regions and connection. *KR*, 92:165–176, 1992.
- M. Ranzini, S. Dehaene, M. Piazza, and E. M. Hubbard. Neural mechanisms of attentional shifts due to irrelevant spatial and numerical cues. *Neuropsychologia*, 47(12):2615–2624, 2009.

- G. Repovš and A. Baddeley. The multi-component model of working memory: Explorations in experimental cognitive psychology. *Neuroscience*, 139(1):5–21, 2006.
- F. Restle. Speed of adding and comparing numbers. *Journal of Experimental Psychology*, 83(2, Pt.1):274, 1970.
- D. C. Richardson, M. J. Spivey, S. Edelman, and A. J. Naples. " language is spatial": Experimental evidence for image schemas of concrete and abstract verbs. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 23, 2001.
- C. C. Ruff, M. Knauff, T. Fangmeier, and J. Spreer. Reasoning and working memory: common and distinct neuronal processes. *Neuropsychologia*, 41(9):1241–1253, 2003.
- A. T. Sack. Parietal cortex and spatial cognition. *Behavioural brain research*, 202(2):153–161, 2009.
- A. Sáenz-Ludlow. Iconicity and diagrammatic reasoning in meaning-making. In *Signs of Signification*, pages 193–215. Springer, 2018.
- A. P. Saygin and M. I. Sereno. Retinotopy and attention in human occipital, temporal, parietal, and frontal cortex. *Cerebral cortex*, 18(9):2158–2168, 2008.
- S. Schleipen, M. Ragni, and T. Fangmeier. Negation in spatial reasoning. In *Annual Conference on Artificial Intelligence*, pages 175–189. Springer, 2007.
- H. Schultheis and T. Barkowsky. Casimir: an architecture for mental spatial knowledge processing. *Topics in cognitive science*, 3(4):778–795, 2011.
- H. Schultheis, S. Bertel, T. Barkowsky, and I. Seifert. The spatial and the visual in mental spatial reasoning: an ill-posed distinction. In *International Conference on Spatial Cognition*, pages 191–209. Springer, 2006.

- C. Schultz, M. Bhatt, J. Suchan, and P. A. Wałęga. Answer set programming modulo 'space-time'. In *International Joint Conference on Rules and Reasoning*, pages 318–326. Springer, 2018.
- W. Sellars. Inference and meaning. *Mind*, 62(247):313–338, 1953.
- P. Shaver, L. Pierson, and S. Lang. Converging evidence for the functional significance of imagery in problem solving. *Cognition*, 3(4):359–375, 1974.
- A. Shimojima. The graphic-linguistic distinction exploring alternatives. *Artificial Intelligence Review*, 13(4):313–335, 1999.
- A. Shimojima. *Semantic properties of diagrams and their cognitive potentials*. Center for the Study of Language and Information, 2015.
- A. Shimojima and Y. Katagiri. An eye-tracking study of exploitations of spatial constraints in diagrammatic reasoning. *Cognitive science*, 37(2):211–254, 2013.
- J. F. Sima. *A computational theory of visuo-spatial mental imagery*. PhD thesis, Universität Bremen, 2014.
- J. F. Sima, H. Schultheis, and T. Barkowsky. Differences between spatial and visual mental representations. *Frontiers in psychology*, 4:240, 2013.
- N. Skovgaard-Olsen, P. Collins, K. Krzyżanowska, U. Hahn, and K. C. Klauer. Cancellation, negation, and rejection. *Cognitive psychology*, 108:42–71, 2019.
- T. J. Slattery, B. Angele, and K. Rayner. Eye movements and display change detection during reading. *Journal of Experimental Psychology: Human Perception and Performance*, 37(6):1924, 2011.

- H. Smessaert, A. Shimojima, and L. Demey. Free rides in logical space diagrams versus aristotelian diagrams. In *International Conference on Theory and Application of Diagrams*, pages 419–435. Springer, 2020.
- R. C. Stalnaker. A theory of conditionals. In *Ifs*, pages 41–55. Springer, 1968.
- G. Stapleton, A. Shimojima, and M. Jamnik. The observational advantages of Euler diagrams with existential import. In *International Conference on Theory and Application of Diagrams*, pages 313–329. Springer, 2018.
- K. Stenning. Distinctions with differences: comparing criteria for distinguishing diagrammatic from sentential systems. In *International Conference on Theory and Application of Diagrams*, pages 132–148. Springer, 2000.
- K. Stenning and J. Oberlander. A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive science*, 19(1):97–140, 1995.
- J. Suchan, M. Bhatt, and P. E. Santos. Perceptual narratives of space and motion for semantic interpretation of visual data. In *European Conference on Computer Vision*, pages 339–354. Springer, 2014.
- J. Suchan, M. Bhatt, and H. Jhavar. Talking about the moving image: A declarative model for image schema based embodied perception grounding and language generation. *arXiv preprint arXiv:1508.03276*, 2015.
- L. Talmy. The fundamental system of spatial schemas in language. *From perception to meaning: Image schemas in cognitive linguistics*, 3, 2005.
- T. Tenbrink. *Space, time, and the use of language: An investigation of relationships*, volume 36. Walter de Gruyter, 2008.

- R. B. Tootell, M. S. Silverman, E. Switkes, and R. L. De Valois. Deoxyglucose analysis of retinotopic organization in primate striate cortex. *Science*, 218(4575):902–904, 1982.
- M.-Y. Tseng. Exploring image schemas as a critical concept: Toward a critical-cognitive linguistic account of image-schematic interactions. *Journal of Literary Semantics*, 2007.
- B. Tversky. Cognitive maps, cognitive collages, and spatial mental models. In *European conference on spatial information theory*, pages 14–24. Springer, 1993.
- L. Vieu. *Sémantique des relations spatiales et inférences spatio-temporelles: une contribution à l'étude des structures formelles de l'espace en langage naturel*. PhD thesis, Toulouse 3, 1991.
- P. A. Wałęga, C. Schultz, and M. Bhatt. Non-monotonic spatial reasoning with answer set programming modulo theories. *Theory and Practice of Logic Programming*, 17(2): 205–225, 2017. doi: 10.1017/S1471068416000193.
- J. Zwarts and P. Gärdenfors. Locative and directional prepositions in conceptual spaces: The role of polar convexity. *Journal of Logic, Language and Information*, 25 (1):109–138, 2016.

RÉSUMÉ

Les études en psychologie du raisonnement spatial se concentrent généralement sur des problèmes de positionnement d'objets, comme par exemple "A est à gauche de B, C est à droite de B, où est A par rapport à C ?". Le modèle computationnel PRISM simule en détail les processus psychologiques qui nous permettent de résoudre ce genre de problèmes, notamment en intégrant des éléments cognitifs pertinents comme les limites attentionnelles. Toutefois, aucun modèle computationnel ne semble exister pour modéliser nos raisonnements sur d'autres types de relations, comme l'inclusion de cercles, ou celles utilisées en géométrie. Cette thèse propose un tel modèle, notamment en combinant les principes psychologiques déjà présents dans PRISM avec les possibilités offertes par le récent outil computationnel *Answer Set Programming Modulo Theories for the Qualitative Spatial domain* ASPMT(QS). En outre, des théories de linguistique cognitive nous autorisent à utiliser ce modèle sur une plus large partie du langage. Cela permet d'apporter un éclairage nouveau sur certains sujets en sciences cognitives, comme par exemple l'étude des connecteurs logiques.

MOTS CLÉS

Psychologie du Raisonnement, PRISM, Modèles Mentaux Spatiaux, ASPMT(QS), Schèmes Imagés, Connecteurs.

ABSTRACT

Studies in the psychology of spatial reasoning often focus on positional problems, such as "A is to the left of B, C is to the right of B, where is A compared to C?". The computational model PRISM developed within the model-based approach, simulates in detail the psychological processes that enable us to solve this type of problem, especially by including relevant cognitive elements such as attentional limits. However, no computational model seems to exist for our reasoning about other types of relations, such as circle inclusion, or those used in geometry. This dissertation aims to fill this gap, in particular by combining the psychological principles already present in PRISM with the possibilities offered by the recent computational tool *Answer Set Programming Modulo Theories for the Qualitative Spatial domain* ASPMT(QS). Moreover, some theories of cognitive linguistics allow us to use the model on a larger part of language. This offers the possibility to shed new light on certain cognitive science topics, such as the study of logical connectives.

KEYWORDS

Psychology of Reasoning, PRISM, Spatial Mental Models, ASPMT(QS), Image Schemas, Connectives.