



UvA-DARE (Digital Academic Repository)

Spatial reasoning : theory and practice

Aiello, M.

Publication date

2002

Document Version

Final published version

[Link to publication](#)

Citation for published version (APA):

Aiello, M. (2002). *Spatial reasoning : theory and practice*. Institute for Logic, Language and Computation.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Spatial Reasoning

Theory and Practice

Marco Aiello

Spatial Reasoning

Theory and Practice

ILLC Dissertation Series 2002-2



INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

For further information about ILLC-publications, please contact

Institute for Logic, Language and Computation

Universiteit van Amsterdam

Plantage Muidergracht 24

1018 TV Amsterdam

phone: +31-20-525 6051

fax: +31-20-525 5206

e-mail: illc@science.uva.nl

homepage: <http://www.illc.uva.nl/>

Spatial Reasoning

Theory and Practice

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof.mr. P.F. van der Heijden
ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in de
Aula der Universiteit
op vrijdag 22 februari 2002, te 12.00 uur

door

Marco Aiello

geboren te Fabriano, Italië.

Promotie commissie:

Promotores:

prof.dr J.F.A.K. van Benthem
prof.dr ir A.M.W. Smeulders

Overige leden:

prof.dr L. Fariñas del Cerro, Université Paul Sabatier, Frankrijk
prof.dr ir F. Giunchiglia, Università di Trento, Italië
prof.dr ir R. Scha, Universiteit van Amsterdam, Nederland
dr M. de Rijke, Universiteit van Amsterdam, Nederland
dr Y. Venema, Universiteit van Amsterdam, Nederland

Faculteit der Natuurwetenschappen, Wiskunde en Informatica
Universiteit van Amsterdam
Nederland

The research was supported by the Institute for Logic, Language and Computation and by the Informatics Institute of the University of Amsterdam.

Copyright © 2002 by Marco Aiello
<http://www.aiellom.it>

Cover design and photography by the author.
Typeset in pdfL^AT_EX.
Printed and bound by Print Partners Ipskamp, Enschede.

ISBN: 90-5776-079-7

A Mario e Ggina.

CONTENTS

Acknowledgments	xi
1 Introduction	1
1.1 Reasoning about space	1
1.2 Theory and practice	3
2 The topo approach: expressiveness	7
2.1 Basic modal logic of space	8
2.1.1 Topological bisimulation	11
2.1.2 Connections with topology	13
2.1.3 Topo-bisimilar reductions	14
2.2 Games that compare visual scenes	14
2.2.1 Strategies and modal formulas	18
2.3 Logical variations	20
3 The topo approach: axiomatics	23
3.1 Topological spaces and Kripke models	23
3.1.1 The basic connection	23
3.1.2 Analogies	25
3.2 General completeness	25
3.2.1 The main argument	26
3.2.2 Topological comments	28
3.2.3 Finite spaces suffice	29
3.3 Completeness on the reals	31
3.3.1 Cantorization	32
3.3.2 Counterexamples on the reals	35
3.3.3 Logical non-finiteness on the reals	39
3.4 Axiomatizing special kinds of regions	42
3.4.1 Serial sets on the real line	42

3.4.2	Formulas in one variable over the serial sets	45
3.4.3	Countable unions of convex sets on the real line	48
3.4.4	Generalization to \mathbb{R}^2	49
3.5	A general picture	51
3.5.1	The deductive landscape	51
4	Logical extensions	53
4.1	Universal reference	54
4.2	Alternative extensions	60
4.2.1	Hybrid reference	60
4.2.2	Until a boundary	62
4.3	Standard logical analysis	65
5	Geometrical extensions	67
5.1	Affine Geometry	67
5.1.1	Basic geometry	67
5.1.2	The general logic of betweenness	69
5.1.3	Modal languages of betweenness	71
5.1.4	Modal logics of betweenness	74
5.1.5	Special logics	76
5.1.6	Logics of convexity	76
5.1.7	First-order affine geometry	81
5.2	Metric geometry	82
5.2.1	The geometry of relative nearness	82
5.2.2	Modal logic of nearness	86
5.2.3	First-order theory of nearness	90
5.3	Linear algebra	91
5.3.1	Mathematical morphology and linear logic	93
5.3.2	Richer languages	96
6	A game-based similarity for image retrieval	101
6.1	Introduction	101
6.2	A general framework for mereotopology	102
6.2.1	Expressiveness	103
6.2.2	Comparison with RCC	105
6.3	Comparing spatial patterns	106
6.3.1	Model comparison games distance	107
6.4	Computing similarities	110
6.4.1	Methodology	110
6.4.2	Polygons of the plane	111
6.4.3	The topo-distance algorithm	115
6.5	The IRIS prototype	119
6.5.1	Implementing the similarity measure	119

6.6	Discussion	122
7	Thick 2D relations for document understanding	125
7.1	Introduction	125
7.2	A logical structure detection architecture	128
7.3	Methodology	129
7.3.1	Document encoding rules	129
7.3.2	Relations adequate for documents	133
7.3.3	Inference	139
7.4	Evaluation	142
7.4.1	Criteria	143
7.4.2	Results	143
7.4.3	Discussion of the results	147
7.5	Concluding remarks	149
8	Conclusions	151
8.1	Where we stand	151
8.2	Final remarks on theory and practice	152
A	A bit of topology	155
B	Sorting transitive directed graphs	159
C	Implementations	163
C.1	Topax	163
C.2	IRIS	169
C.3	SpaRe	171
	Bibliography	181
	Index	193
	Samenvatting	199

ACKNOWLEDGMENTS

I believe that a PhD thesis is not the effort of an individual, but the final outcome of a synergy. Since there is officially one author, I feel the urge to thank in these initial pages the people involved in a way or another with my PhD project.

I arrived in Amsterdam four years ago as a lost soul. Unlike Voltaire's *Candide*, I found generosity, humor, inspiration and, most importantly, solid scientific values to believe in. I realized my conversion to the 'Amsterdam school' was complete when I received an email in which I was addressed as a "modal logician." What a joy. I did not even know what a 'modal logic' was till I moved my first steps in Amsterdam. If today someone may recognize me as a credible scientist, I own it first and foremost to Johan van Benthem.

One of the many qualities of Johan I had the privilege to appreciate, and I am sure I share this feeling with many others, is his natural disposition of putting everyone at ease. I could always raise a question, no matter how silly, and get a simple yet illuminating answer. Every occasion to meet, discuss or even exchange emails with Johan have been pleasurable events which I have and I will be looking forward to. In short, Johan thank you.

I am deeply in debt with Arnold Smeulders for his continuous interest in my work, for his warm supervision, and for an extreme availability. He provided me with visionary questions, while leaving me a considerable amount of freedom in my research. I only wish I could have answered more of his questions. Not to please him, but because if I did, I would be a famous scientist today.

I am thankful to a number of researchers of the University of Amsterdam and CWI. I could always knock on their doors and find an interested and competent mind at my disposal, in particular, Rein van den Boomgaard, Kees Doets, Henk Heijmans, Dick de Jong, Michiell van Lambalgen, Maarten Marx, Maarten de Rijke, and Yde Venema.

Special thanks go to Krzysztof Apt. Interacting with him has taught me a lot about science and, also, on academic life's pleasures and pitfalls. I have appreciated his trust when he proposed to me to become the information director of the new ACM journal he was founding, the *Transactions on Computational Logic* (TOCL). Our occasional

Tuesdays lunches were both filled with precious information and humor. I should have known he was a funny guy from the very beginning, after all on his homepage you can read “Recent Publications (Important notice: I have read them all.)”

I also thank my co-authors. I have learned a great deal in many dimensions from working with all of them: Johan van Benthem, Arnold Smeulders, Guram Bezhanishvili, Maarten de Rijke, Christof Monz, Carlos Areces, the ‘trentini’ Luciano Serafini, Paolo Busetta, Antonia Donà, Alessandro Agostini, and the document analysis folks Marcel Worring, and Leon Todoran.

Discussing scientific issues at conferences, during lab visits or even via email, had a huge impact on my ideas for the thesis. I can not cite everybody in such a society of minds. Let me just thank them all, and recall a few: Brandon Bennett, Guram Bezhanishvili, Luis Fariñas del Cerro, Alessandro Cimatti, Anthony Cohn, Alberto Finzi, Dov Gabbay, Valentin Goranko, Oliver Lemon, Carsten Lutz, Donato Malerba, Angelo Montanari, Carla Piazza, Alberto Policriti, Ian Pratt, Dave Randell, Oliviero Stock, Paolo Traverso, Achille Varzi, and Laure Vieu.

Fiora Pirri, Fausto Giunchiglia, Mike Papazoglou, and Henk Heijmans invited me to visit their labs and present my work. The lab visits have been great experiences fostering discussion and collaboration. I spent two months at IRST, Trento as a follow up to one of such invitations. I would like to thank Fausto for making this possible, and all the people at IRST for providing a productive and fun environment.

I would like to thank the participants in the workshop on modal logics of space organized by Yde Venema and myself in Amsterdam on May the 10th 1999, and especially the invited speakers: Philippe Balbiani, Luis Fariñas del Cerro, Volker Haarslev, Oliver Lemon, Ian Pratt, and Vera Stebletsova, for contributing to the success of the event. The spatial reasoning reading group at ILLC, which began its meetings shortly after the workshop, with its regular members Rosella Gennari, Gwen Kerdiles, Vera Stebletsova, and Yde Venema, provided a great learning opportunity. In particular, Yde’s explanations have been fundamental in my understanding of spatial logics.

At the crossing between scientific collaboration and friendship I thank all the colleagues of ILLC and ISIS, in particular: Carlos Areces, Raffaella Bernardi, Rosella Gennari, Eva Hoogland, Gwen Kerdiles, Nikos Massios, Christof Monz, Marc Pauly, Leon Todoran, and Renata Wassermann. Carlos, Christof, and Rosella also volunteered to proofread the thesis. I thank them both for their precious remarks and their masochism. Peter Block, Ingrid van Loon, and Marjan Veldhuisen of the administrative staff at ILLC were there for me whenever needed. Thank to them, I never had to worry about anything different from my PhD.

Amsterdam is a fun city to live in. But it is even better with the right friends. I am grateful to the Dutch friends for introducing me to the Netherlands, and to the non-Dutch friends for exploring things with me: Arantxa, Christiaan en Mirjam, Christof, Elena, Jelle, Jilles, Karen, Paola, Piero, Tom en Annemieke, the basketball teammates of the USC and Schrobbehaar.

During my visits to Stanford, I had the opportunity to enjoy the warm hospitality of Richard and Fran Waldinger, for which I am very thankful.

A number of friends from Italy came to visit me in Amsterdam, giving relief (as well as Parmigiano cheese) to an emigrant's life. I thank them all for the time spent together in the 'low-lands': Simone e Margherita (also thank you for making me feel less lonely on my regular trips to Sassari), Ada, Alberto, Alessandra, Alessandro e Claudia, Cinzia, Franco e Federica, Claudio e Barbara, Daniele, Fabio, Irma, Massimo e Nadia, Mattia e Alessandra, Nicola, Paola, and Paolo e Giulia. In particular, I thank Irma Rosati and Mattia De Rosa for kindly accepting to be my paranimphs during the public defense of this thesis.

Finally, I acknowledge the Italian National Research Council (CNR) for providing extra financial support (part of which I converted in an indispensable companion—woz—an Apple Powerbook G3). In addition, I acknowledge the following fine organizations for providing support through their travel awards: the Italian Association for Artificial Intelligence (AI*IA), the Center for the Study of Language and Information at Stanford (CSLI), and the Organization Committee of the 7th European Workshop on Logics in Artificial Intelligence (Jelia 2000).

The thesis, which would not be there without the love of Claudia, is dedicated to my parents and grandparents. In particular, I dedicated it *in memoriam* to Mario Aiello, the father, but also the computer scientist, on the 25th anniversary of his death.

Amsterdam, October 29th 2001

Marco Aiello

1.1 Reasoning about space

Spatial structures and spatial reasoning are essential to perception and cognition. Much day-to-day practical information is about what happens at certain spatial locations. Moreover, spatial representation is a powerful source of geometric intuitions that underlie general cognitive tasks. How can we represent spatially located entities and reason about them? To take a concrete domestic example: when we are setting a table and place a spoon, what are the basic spatial properties of this new item in relation to others, and to the rest of the space? Not only, there are further basic aspects to perception: we have the ability to compare different visual scenes, and recognize objects across them, given enough ‘similarity’. More concretely: which table settings are ‘the same’? This is another task for which logic provides tools.

Constraining space within the bounds of a logical theory and using related formal reasoning tools must be performed with particular care. One cannot expect the move from space to formal theories of space to be complete. Natural spatial phenomena will be left out of logical theories of space, while non-natural spatial phenomena could try to sneak in (cf. the account of Helly’s theorem implications on diagrammatic theories in [Lemon, 2002]). Paraphrasing Ansel Adams’ concern of space bound in a photograph,¹ one could say that space in nature is one thing; space confined and restricted in the bounds of a formal representation and reasoning system is quite another thing. Connectivity, parthood, and coherence, should be correctly handled and expressed by the formalism, not aiming at a complete representation of space, but focusing on expressing the most perspicuous spatial phenomena.

The preliminary and fundamental step in devising a spatial reasoning framework lays, thus, in the identification of which spatial behaviors the theory should capture and,

¹“Space in nature is one thing; space confined and restricted by the picture edges is quite another thing. Space, scale, and form must be made eloquent, not in imitation of painting arrangements, but in terms of the living camera image.” [Adams, 1981]

possibly, in the identification of which practical uses will be made of the framework. A key factor is in appropriately balancing expressive power, completeness with respect to a specific class of spatial phenomena, and computational complexity.

The blend of expressivity and tractability we are aiming at points us in the direction of modal logics as a privileged candidate for the formalization task. We will not go into details on modal logics or on the reason for which modal logics balance nicely expressive power and computational complexity (one can refer to a number of texts on the subject, including the recent [Blackburn et al., 2001] or the more specific [Vardi, 1997]). To enjoy the theoretical part of the thesis, we assume the reader has some basic knowledge of modal logic and its best-known possible world semantics (also referred to as Kripke semantics). Strangely enough, even though knowledge of Kripke semantics is helpful for better understanding the presented material, we are going to make little use of it, and rather resort to topological semantics, introduced about 30 years earlier than Kripke semantics by [Tarski, 1938]. Modern modal logics of space need old modal logic semantics.

The attention on spatial reasoning stems, in the case of the present thesis, from the interest in applications in the domains of image processing and computer vision, hence, the sub-title *Theory and Practice*. But this is only one of the many motivations for which spatial logics have been considered in the past. These range from the early philosophical efforts [Whitehead, 1929, Lesniewski, 1983] to recent work motivated by such diverse concerns as spatial representation and vision in AI [Shanahan, 1995, Randell et al., 2001], semantics of spatial prepositions in linguistics [Herskovits, 1997, Winter and Zwartz, 1997], perceptual languages [Dastani et al., 1997, Dastani, 1998], or diagrammatic reasoning [Hammer, 1995, Gurr, 1998, Kerdiles, 2001]. The resulting logics are diverse, too. Theories differ in their primitive objects: points, lines, polygons, regions (contrast [Tarski, 1938] against [Tarski, 1959]). Likewise, theories differ in their primitive spatial relations: such as inclusion, overlap, touching, ‘space’ versus ‘place’, and on how these should be interpreted: [Randell et al., 1992, Bennett, 1995, Asher and Vieu, 1995]. There are mereological theories of parts and wholes, topological ones (stressing limit points, and connection) and mereotopological ones (based on parthood and external connection). Systematic accounts of the genesis of spatial vocabulary date back to Helmholtz’ work on invariants of movement, but no generally agreed primitive relations have emerged on the logic side. Moreover, axioms differ across theories: [Clarke, 1981, Clarke, 1985] vs [Pratt and Schoop, 1998] vs [Casati and Varzi, 1999]. Also our modal approach has its predecessors of which we mention [Segerberg, 1970, Segerberg, 1976, Shehtman, 1983, Bennett, 1995, Venema, 1992, Balbiani et al., 1997, Lemon and Pratt, 1998].

The above references have no pretense of being a complete overview of the literature on spatial formalisms and, even less, on applications of spatial formalisms. We shall refer, discuss and compare our work with the literature, with previous approaches and systems on a ‘local basis’. That is, relevant literature is discussed in each chapter where appropriate in order to set the context, compare our approach with previous ones, and identify future extensions of our own work based on previous efforts.

1.2 Theory and practice

Our contribution with this thesis is twofold. On the one hand, we investigate new and existing spatial formalisms with the explicit goal of identifying languages nicely balancing expressive power and tractability. On the other hand, we study the feasibility of practical applications of such qualitative languages of space, by investigating two symbolic approaches to pattern recognition.

The structure of the thesis reflects the two sub-tasks. The first part reflects the ethereal nature of our theoretical approach to space. The second part reflects a more practical task, that is, applying spatial theories to real world problems.

Modal formalisms are the thread of the thesis. We walk through a family of modal languages of space for topological, affine, metric and vector spaces. The task is not that of compiling a drudging taxonomy of modal spatial languages, but rather to design languages with specific expressive tasks. ‘Expressivity in balance’ is the motto here.

While walking through modal logics of space some steps will be mandatory. Some basic languages are needed as they form the basic for any subsequent analysis. This is the case of **S4**: a poor language in terms of expressivity, but, as it turns out, the minimal normal modal language with respect to topological interpretations. In fact, this language will be our first test. On this language we shall introduce the topological semantics (after Tarski), define adequate notions of bisimulations and model comparison games, analyze completeness in modern terms (via canonical models), and more.

Our subsequent investigation concerns some striking facts about **S4**. First, we consider completeness with respect to general topological spaces, to Cantor space, to the real line, and further to serial sets of the real line and plane. Spatial finiteness arises as a result. Then, we look at logical extensions. A typical example of this kind of language is that of **S4_u**, an extension by a universal modal operator. **S4_u** is known in the literature of spatial reasoning as Bennett [1995] used it to encode a decidable fragment of the region connection calculus of Randell, Cui and Cohn [1992]. Further examples comprise the spatial extension of the temporal Since and Until logic of [Kamp, 1968].

Our next move is from topology to geometric structures. This involves a major semantic change. Topological interpretation is abandoned, and more custom possible worlds semantics is used. In this context, modal logics tend to either be sorted (typical example is that of having sorts for points and lines, and an incidence relation) or to adopt dyadic modal operators. Our focus will be on logics of the second kind.

In [Tarski, 1959], Tarski introduces the notion of elementary geometry and provides a first order axiomatization in terms of two fundamental relations, that is, betweenness and equidistance. These are sufficient for any affine or metric construction. For instance, one can define parallelism, convexity, or the notion of an equilateral triangle. But what happens if one considers betweenness in isolation? Further, what is the modal fragment of languages of betweenness? And, are there alternative relations for axiomatizing elementary geometry?

We answer these questions in our investigation of geometrical extensions to our basic modal approach to space. At the end of our journey in this realm of modal logics,

we arrive at a vector theory of shape: mathematical morphology. This mathematical theory of shape lends itself naturally to modal representations, as its two basic operators, which mimic Minkowski's operations in vector spaces, are easily axiomatized in terms of modal 'arrow logics'. It will be harder to maintain the balance between expressivity and tractability as small deviations from the minimal axiomatization force trespassing the limits of decidability. As a compensation, interesting new axiomatizations and open questions arise. All in all, we shall discover a number of intriguing facts about topological and geometric spaces, thanks to a modal analysis of space.

When considering applications, the point of view on the logics of space analyzed in our theoretical promenade shifts. Now interesting logics become those which can express region properties, rather than those merely referring to points, model comparison games become interesting only if turned into distance measures, and boundaries of regions play an even greater role.

There are even more general concerns when applying symbolic approaches to pattern recognition problems: spatial coherence and brittleness. Spatial coherence regards the way nature presents itself to observation, that is in a manner intrinsically hard to capture symbolically. Elsewhere we have spelled out our personal concerns for spatial coherency in the context of formal perceptual languages [Aiello and Smeulders, 1999]. We refer to [Florack, 1997] for an authoritative point of view.

Brittleness regards a risk ran by strict symbolic approaches when applied to real world domains: they might break. There are various reasons for which a system can show a brittle behavior. Little variations present in nature may result in misclassifications at the symbolic level. Thus, the misclassification propagates on to a wrong analysis. The problem occurred in one of the practical systems we present, forcing the introduction of a 'less brittle' interpretation of region relations.

We choose two significant problems in image processing and pattern matching as our testing grounds: image retrieval and document image analysis.

Image retrieval is achieved by matching a description or a query image on a collection of images. Symbolic approaches are successful in this field to the extent that symbolic segmentation of the images is available. The matching process between a query and a collection of images is a matter of comparison. When analyzing modal logics of space we encounter a tool performing precisely this task: model comparison games, which we apply to measure image similarity.

We believe that the field of document image analysis is ripe for symbolic approaches. Various decades of research in pattern matching have solved most of the problems involved in basic document image processing. For example, current technology for skew estimation or optical character recognition is very accurate. One of the present challenges lies in the management and grouping of all the basic layout information in order to achieve document understanding. Symbolic approaches are of interest here, as there is formal structure to be detected in printed documents. One may even argue, as we do, that the structure present in documents has the form of precise formal rules. These are the rules followed, most often without awareness, by document

authors and, with awareness, by composers. It is by reverse engineering these rules and by using them to analyze documents that we can achieve document understanding.

The overall conclusion over our practical experiences will help us understand where they are effective and where not. Practical issues also prompt for interesting theoretical questions, thus, closing the ‘vicious circle’ theory and practice—practice and theory.

The thesis is organized in seven technical chapters, plus an introductory and a conclusions chapters, and three appendices. The chapters from 2 to 5 form the theoretical core of the dissertation, while Chapters 6 and 7 are the practical component.

The first two chapters set the boundaries of our framework: Chapter 2 from the expressive point of view, and Chapter 3 from the axiomatization one. Then, we analyze two sorts of extensions of the framework. Logical extensions are presented in Chapter 4, while geometrical ones are introduced in Chapter 5.

In Chapter 2, we revive the topological interpretation of modal logics, turning it into a general language of patterns in space. In particular, we define a notion of bisimulation for topological models that compares different visual scenes. We refine the comparison by introducing Ehrenfeucht-Fraïssé style games played on patterns in space.

In Chapter 3, we investigate the topological interpretation of modal logic in modern terms, using the notion of bisimulation introduced in Chapter 2. We look at modal logics with interesting topological content, presenting, amongst others, a new proof of McKinsey and Tarski’s theorem on completeness of $S4$ with respect to the real line, and a completeness proof for the logic of finite unions of convex sets of the reals.

In Chapter 4 we consider logical extensions to the topological modal approach to space. The introduction of universal and hybrid modalities is investigated with respect to the added logical expressive power. A spatial version of the tense Since and Until logic is also examined. A brief comparison with higher-order formalisms gives a more general perspective on (extended) modal logics of space.

In Chapter 5, we proceed with the modal investigation of space by moving to affine and metric geometry, and vector algebra. This allows us to see new fine-structure in spatial patterns suggesting analogies across these mathematical theories in terms of modal, tense and conditional logics. Expressive power is analyzed in terms of language design, bisimulations, and correspondence phenomena. The result is both unification across the areas visited, and the uncovering of interesting new questions.

In Chapter 6, we take a different look at model comparison games for the purpose of designing an image similarity measure for image retrieval. Model comparison games can be used not only to decide whether two specific models are equivalent or not, but also to establish a measurement of difference among a whole class of models. We show how this is possible in the case of the spatial modal logic $S4_u$. The approach results in a spatial similarity measure based on topological model comparison games. We move towards practice by giving an algorithm to effectively compute the similarity measure for a class of topological models widely used in computer science applications: polygons of the real plane. At the end of the chapter, we briefly overview an implemented system based on the game-similarity measure.

In Chapter 7, we use a propositional language of qualitative rectangle relations to detect the reading order from document images. To this end, we define the notion of a document encoding rule and we analyze possible formalisms to express document encoding rules such as \LaTeX , SGML languages, and others. Document encoding rules expressed in the propositional language of rectangles are used to build a reading order detector for document images. In order to achieve robustness and avoid brittleness when applying the system to real life document images, the notion of a thick boundary interpretation for a qualitative relation is introduced. The system is tested on a collection of heterogeneous document images showing recall rates up to 89%.

The presentation ends with three appendices. Appendix A is a brief recall of basic topological notions, useful for reading Chapters 2, 3, and 4. Appendix B presents an algorithm for sorting directed transitive cyclic graphs in relation to the system presented in Chapter 7. Appendix C overviews three implementations related to the thesis.

Material related to the thesis has been presented in various contexts. The contributions are to be considered joint with the respective co-authors.

Chapter	Co-authors	reference
2	Johan van Benthem	[Aiello and van Benthem, 1999], a short version is to appear in a <i>CSLI</i> volume [Aiello and van Benthem, 2002a]
3	Johan van Benthem Guram Bezhanishvili	[Aiello et al., 2001], submitted to the <i>Journal of Logic and Computation</i>
4, 5	Johan van Benthem	[Aiello and van Benthem, 1999, Aiello and van Benthem, 2002b], submitted as one paper to the <i>Journal of Applied non-Classical Logics</i>
6		[Aiello, 2000, Aiello, 2001a], to appear in the <i>Journal of the Interest Group in Pure and Applied Logic</i> [Aiello, 2002b]
7	Arnold Smeulders	manuscript submitted to " <i>Information Sciences</i> "

The material of Chapter 7 describes a component of a larger architecture. The latter has been presented in various contexts: [Aiello et al., 2000, Todoran et al., 2001a], and [Todoran et al., 2001b] which has been submitted to the *Journal of Document Analysis and Recognition*.

CHAPTER 2

THE TOPO APPROACH: EXPRESSIVENESS

We begin our investigation of representations of space from a simple modal logic. Our primary goals here are that of identifying the appropriate tools we need in the rest of the thesis and instantiating them for the simplest modal spatial logic.

Perhaps we are already running too fast. We have assumed an agreement on the meaning of the term ‘space’ and we have started to refer to spatial languages talking about a simplest one. But the goal of assigning a unique meaning to the term space is really open-ended and under-determined. Mathematicians have developed many different formal accounts, ranging from less or more fine-grained geometries (affine, metric) to more coarsely-grained topologies. Philosophers have even added formal theories of their own, such as ‘mereology’, cf. [Casati and Varzi, 1999]. Qualitatively different levels of description also arise naturally in computer science, viz. mathematical morphology [Serra, 1982]. A similar diversity of grain levels arises in logic, which provides many different spatial languages for talking about objects and their locations. Our general paradigm is this hierarchy of levels, even though we develop our methods mainly at the level of topology, cf. [Singer and Thorpe, 1967] or [Engelking, 1989]. Inside the topological level, one can identify a sub-hierarchy of languages of increasing expressive power and logical complexity. We begin at the bottom of this hierarchy with the simplest language. Simplest here means less expressive language, both from a syntactic and a semantic point of view. The syntactic evidence to the claim of simplicity will be provided in the present chapter.

The simple language is **S4**. The name will not surprise the modal logician since **S4** is a well known modal logic: the logic of partial orders. Maybe the surprise lies in the fact that it is the simplest spatial logic, in place of **K**, which is the simplest normal modal logic for possible worlds semantics. Again, explanations will follow.

In the present chapter, we recall the syntax and state the truth definition for **S4** in the spatial context. We proceed by providing the two fundamental tools tied to our modal approach to space which keep us company for most of the thesis: topological bisimulations and topological games.

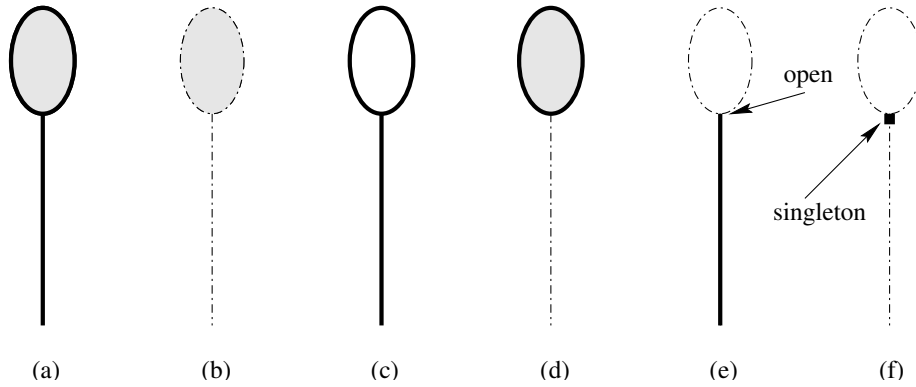


Figure 2.1: A formula of the language $S4$ identifies a region in a topological space. (a) a spoon, p . (b) the containing part of the spoon, $\Box p$. (c) the boundary of the spoon, $\Diamond p \wedge \Diamond \neg p$. (d) the container part of the spoon with its boundary, $\Diamond \Box p$. (e) the handle of the spoon, $p \wedge \neg \Diamond \Box p$. In this case the handle does not contain the junction point handle-container. (f) the joint point handle-container of the spoon, $\Diamond \Box p \wedge \Diamond (p \wedge \neg \Diamond \Box p)$: a singleton in the topological space.

The chapter is rich in visual examples that should help in grounding intuitions of the logic and of the tools we define. The images of the chapter—and of the following ones—borrow from the daily activity of eating, in particular cutlery is the running example in the figures. Unless stated otherwise, all depicted items are to be considered subsets of \mathbb{R}^2 equipped with the standard topology (that defined by the unitary disks). Closed contours indicate that the set is not only the contour, but also all the points inside. Of course, these spoons and forks should be taken with a grain of salt: our framework is completely general.

For the convenience of the reader, and to make the thesis as much as possible self-contained, we recall the basic topological definitions in Appendix A.

2.1 Basic modal logic of space

In the 30s, Tarski provided a topological interpretation and various completeness theorems ([McKinsey and Tarski, 1944, Rasiowa and Sikorski, 1963]) making modal $S4$ the basic logic of topology. In the topological interpretation of a modal logic, each propositional variable represents a region of the topological space, and so does every formula. Boolean operators such as negation \neg , or \vee , and \wedge are interpreted as complement, union and intersection, respectively. The modal operators diamond and box, become the topological closure and interior operators. More precisely, the modal logic $S4$ consists of:

- a set of proposition letters P ,

Formula	Interpretation
\top	the universe
\perp	the empty region
$\neg\varphi$	the complement of a region
$\varphi \wedge \psi$	intersections of the regions φ and ψ
$\varphi \vee \psi$	union of the regions φ and ψ
$\Box\varphi$	interior of the region φ
$\Diamond\varphi$	closure of the region φ

Figure 2.2: Formulas of **S4** and their intended meaning.

- two constant symbols \top, \perp ,
- Boolean operators $\neg, \wedge, \vee, \rightarrow$, and
- two unary modal operators \Box, \Diamond .

Formulas are built by means of the following recursive rules:

- p such that $p \in P$ is a well formed formula,
- \top, \perp are well formed formulas,
- $\neg\varphi, \varphi \vee \psi, \varphi \wedge \psi$ are well formed formulas if φ and ψ are well formed formulas,
- $\Box\varphi$ and $\Diamond\varphi$ are well formed formulas if φ is well formed formula.

In Figure 2.2, the intended meaning of some basic formulas is summarized. These are pictured more vividly in Figure 2.1 with a spoon-shaped region. The intuitions about the language are reflected in its semantics, which involves the idea of special regions denoted by proposition letters. Topological models (*topo-model*) $M = \langle X, O, \nu \rangle$ are topological spaces (X, O) plus a valuation function $\nu : P \rightarrow \mathcal{P}(X)$. Conversely, we will sometimes strip the valuation from a topo-model, and just consider its underlying topological space. This is like working with *frames* in the usual Kripke semantics.

2.1.1. DEFINITION (TOPOLOGICAL SEMANTICS OF S4). Truth of modal formulas is defined inductively at points x in topological models M :

$M, x \models \perp$	never
$M, x \models \top$	always
$M, x \models p$	iff $x \in \nu(p)$ (with $p \in P$)
$M, x \models \neg\varphi$	iff not $M, x \models \varphi$
$M, x \models \varphi \wedge \psi$	iff $M, x \models \varphi$ and $M, x \models \psi$
$M, x \models \varphi \vee \psi$	iff $M, x \models \varphi$ or $M, x \models \psi$
$M, x \models \varphi \rightarrow \psi$	iff if $M, x \models \varphi$, then $M, x \models \psi$
$M, x \models \Box\varphi$	iff $\exists o \in O : x \in o \wedge \forall y \in o : M, y \models \varphi$
$M, x \models \Diamond\varphi$	iff $\forall o \in O : \text{if } x \in o, \text{ then } \exists y \in o : M, y \models \varphi$

As usual we can economize by defining $\varphi \vee \psi$ as $\neg\varphi \rightarrow \psi$, and $\diamond\varphi$ as $\neg\Box\neg\varphi$.

One of Tarski's early results was this. Universal validity of formulas over topological models has the modal logic **S4** as a sound and complete proof system. The standard axiomatization is:

$$\begin{aligned} \diamond A &\leftrightarrow \neg\Box\neg A && \text{(Dual.)} \\ \Box(A \rightarrow B) &\rightarrow (\Box A \rightarrow \Box B) && \text{(K)} \\ \Box A &\rightarrow A && \text{(T)} \\ \Box A &\rightarrow \Box\Box A && \text{(4)} \end{aligned}$$

Modus Ponens and Necessitation are the rules of inference:

$$\frac{\varphi \rightarrow \psi \quad \varphi}{\psi} \text{ (MP)} \qquad \frac{\varphi}{\Box\varphi} \text{ (N)}$$

For a closer fit to topological reasoning, however, it is better to work with an equivalent axiomatization:

$$\begin{aligned} \Box\top &&& \text{(N)} \\ (\Box\varphi \wedge \Box\psi) &\leftrightarrow \Box(\varphi \wedge \psi) && \text{(R)} \\ \Box\varphi &\rightarrow \varphi && \text{(T)} \\ \Box\varphi &\rightarrow \Box\Box\varphi && \text{(4)} \end{aligned}$$

Modus Ponens and Monotonicity are the only rules of inference

$$\frac{\varphi \rightarrow \psi \quad \varphi}{\psi} \text{ (MP)} \qquad \frac{\varphi \rightarrow \psi}{\Box\varphi \rightarrow \Box\psi} \text{ (M)}$$

In addition, consider the following derived theorem of **S4**:

$$\Box A \vee \Box B \leftrightarrow \Box(\Box A \vee \Box B) \qquad \text{(or)}$$

Axiom (Dual.) reflects the topological duality of interior and closure. Axiom (K) does not have an immediate interpretation, but it is equivalent to theorems (N) and (R), which do (cf. [Bennett, 1995]). (N) says the whole space is open. (R) is the finite intersection condition on a topological space. Next, (or) says that open sets are closed under finite unions. (Closure under *arbitrary* unions requires an infinitary extension of the modal language.) Finally, axiom (T) says every set contains its interior, and (4) expresses inflationarity of the interior operator. Further principles of **S4** may define special notions in topology. For instance, the derived rule

$$\text{if } \Box(\varphi \leftrightarrow \diamond\Box\varphi), \text{ then } \Box(\Box\neg\varphi \leftrightarrow \Box\diamond\Box\neg\varphi)$$

says that if a set is closed regular, so is its 'open complement'.

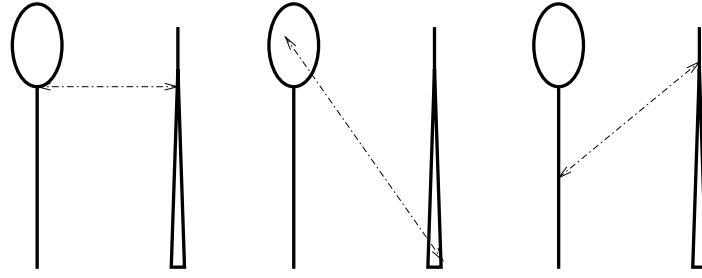


Figure 2.3: A spoon is bisimilar to a ‘chop-stick’. The relation among points that match is highlighted via the double headed arrows.

2.1.1 Topological bisimulation

Once we have a language for expressing properties of visual scenes, we can also formulate *differences* between such scenes. This brings us to the notion of ‘sameness’ for spatial configurations associated with our language, and hence to techniques of comparison. The following is the topological version of a well-known notion from modal logic and computer science ([van Benthem, 1976, Park, 1981]).

2.1.2. DEFINITION (TOPOLOGICAL BISIMULATION). Consider the language **S4** and two topological models $\langle X, O, \nu \rangle, \langle X', O', \nu' \rangle$. A *topological bisimulation* is a non-empty relation $\rightleftharpoons \subseteq X \times X'$ such that if $x \rightleftharpoons x'$ then:

- (i) $x \in \nu(p) \Leftrightarrow x' \in \nu'(p)$ (for any proposition letter p)
- (ii) (*forth condition*): $x \in o \in O \Rightarrow \exists o' \in O' : x' \in o'$ and $\forall y' \in o' : \exists y \in o : y \rightleftharpoons y'$
- (iii) (*back condition*): $x' \in o' \in O' \Rightarrow \exists o \in O : x \in o$ and $\forall y \in o : \exists y' \in o' : y \rightleftharpoons y'$

We call a bisimulation *total* if it is defined for all elements of X and of X' . We overload the symbol \rightleftharpoons extending it to models with points: $\langle X, O, \nu \rangle, x \rightleftharpoons \langle X', O', \nu' \rangle, x'$ requires also that $x \rightleftharpoons x'$. If only the atomic clause (i) and the forth condition (ii) hold, we say that the second model *simulates* the first one.

To motivate this definition, one can look at the ‘topological dynamics’ of the back and forth clauses, seeing how they make x, x' lie in the same ‘modal setting’. Further motivations come from a match with modal formulas, and basic topological notions.

2.1.1. EXAMPLE (SPOON AND CHOP-STICK). *Is a spoon the same as a chop-stick?* The answer depends of course on how we define this cutlery. Suppose we let the spoon be a closed ellipse plus a touching straight line and the chop-stick a straight line touching a closed triangle (cf. Figure 2.3). Let us regard both as the interpretation of some fixed proposition letter p in their respective models. Then we do have a topobisimulation by matching up (a) the two ‘junction points’, (b) all points in the two

handles, and likewise for (c) the interiors, (d) the remaining boundary points, and (e) all exterior points in both models.

Many more examples and cutlery related pictures of topologically bisimilar and not spaces can be found in the technical report [Aiello and van Benthem, 1999].

Crucially, modal spatial properties are *invariant for topo-bisimulations*:

2.1.3. THEOREM. *Let $M = \langle X, O, \nu \rangle$, $M' = \langle X', O', \nu' \rangle$ be models with bisimilar points $x \in X$, $x' \in X'$. For all modal formulas φ , $M, x \models \varphi$ iff $M', x' \models \varphi$.*

Proof Induction on φ . The case of a proposition letter p is the first condition on \simeq . As for conjunction, $M, x \models \varphi \wedge \psi$ is equivalent by the truth definition to $M, x \models \varphi$ and $M, x \models \psi$, which by the induction hypothesis is equivalent to $M', x' \models \varphi$ and $M', x' \models \psi$, which by the truth definition amounts to $M', x' \models \varphi \wedge \psi$. The other Boolean cases are similar. For the modal case, we do one direction. If $M, x \models \Box\varphi$, then by the truth definition we have that $\exists o \in O : x \in o \wedge \forall y \in o : M, y \models \varphi$. By the forth condition, corresponding to o , there must exist an $o' \in O'$ such that $\forall y' \in o' \exists y \in o y \simeq y'$. By the induction hypothesis applied to y and y' with respect to φ , then $\forall y' \in o' : M', y' \models \varphi$. By the truth definition of the modal operator we have $M', x' \models \Box\varphi$. Using the back condition one proves the other direction likewise. QED

To clinch the fit, we need a converse. In general this fails, and matters become delicate (see [Blackburn et al., 2001]). The converse does hold when we use an *infinitary* modal language—but also for our finite language over special classes of models. Here is a nice illustration: finite modally equivalent pointed models are bisimilar.

2.1.4. THEOREM. *Let $M = \langle X, O, \nu \rangle$, $M' = \langle X', O', \nu' \rangle$ be two finite models, $x \in X$, and $x' \in X'$ two points in them such that for every φ , $M, x \models \varphi$ iff $M', x' \models \varphi$. Then there exists a bisimulation between M and M' connecting x and x' .*

Proof To get a bisimulation between the two finite models, we stipulate that $u \simeq u'$ if and only if u and u' satisfy the same modal formulas. The atomic preservation condition for a bisimulation holds since the modal φ include all proposition letters. We now prove the forth condition. Suppose that $u \simeq u'$ where $u \in o$. We must find an open o' such that $u' \in o'$ and $\forall y' \in o' \exists y \in o : y \simeq y'$. Now, suppose there is no such o' . Then for every o' containing x' $\exists y' \in o' : \forall y \in o : \exists \varphi_y : y \not\models \varphi_y$ and $y' \models \varphi_y$. In words, every open o' contains a point y' with no modally equivalent point in o . Taking the finite conjunction of all formulas φ_y , we get a formula $\Phi_{o'}$ such that $y' \models \Phi_{o'}$ and $\neg\Phi_{o'}$ is true everywhere in o . Slightly abusing notation, we write $o \models \neg\Phi_{o'}$. This line of reasoning holds for any open o' containing x' as chosen. Therefore, there exists a collection of formulas $\neg\Phi_{o'}$ for which $o \models \bigwedge_{o'} \neg\Phi_{o'}$. Since $x \in o$, by the truth definition we have $x \models \Box \bigwedge_{o'} \neg\Phi_{o'}$. By the fact that x and x' satisfy the same modal formulas, it follows that $x' \models \Box \bigwedge_{o'} \neg\Phi_{o'}$. But then, there exists an open o^* (with $x' \in o^*$) such that

$o^* \models \bigwedge_{o'} \neg \Phi_{o'}$. Since o^* is an open containing x' , is one of the o' , i.e. $o^* \models \neg \Phi_{o^*}$. But we had supposed that for all opens o' there was a point $y' \models \Phi_{o'}$, so in particular the y' of o^* satisfies Φ_{o^*} . We have thus reached a contradiction: which shows that some appropriate open o' must exist. The back clause is proved analogously. QED

2.1.2 Connections with topology

The preceding results provide a match with logical definability. But topo-bisimulations are also related to purely topological notions. Let us consider only topological frames now, without valuations. Clearly, we have the following implication:

2.1.5. THEOREM. *Homeomorphism implies total topo-bisimulation.*

But not vice-versa! Homeomorphisms provide much more ‘analogy’ between two spaces than topo-bisimulations. A trivial way of seeing this is as follows. *Any two topological spaces are bisimilar.* One can just take the full Cartesian product of their points. Nevertheless, this is not a trivialization of the notion. First, specific topo-bisimulations may be of independent interest – e.g., those preserving additional properties’ of points (encoded in topo-models), where no similar trivial example exists. Second, the back clause of topo-bisimulation resembles the characteristic property of *continuous maps*. This fact provides a foothold for a systematic ‘modal logic analysis’ of topological behavior. E.g., *existential* modal formulas constructed from literals, conjunction, disjunction and box only are preserved under simulations.

2.1.6. THEOREM. *Let $M = \langle X, O, \nu \rangle$, $M' = \langle X', O', \nu' \rangle$ be two models, with a simulation \rightarrow from M to M' , such that $x \rightarrow x'$. Then, for any existential modal formula φ , $M, x \models \varphi$ only if $M', x' \models \varphi$.*

This result explains how continuous maps preserve basic topological properties. The following fact is just one typical illustration:

2.1.7. COROLLARY. *Let f be a surjective continuous map from $\langle X, O \rangle$ to $\langle X', O' \rangle$. If the space $\langle X, O \rangle$ is connected, then so is $\langle X', O' \rangle$.*

We leave the proof of Corollary 2.1.7 for Section 4.1. The reason for postponing the proof is the need of extra logical power at the language level, more precisely, one needs universal quantification over points. The origin of this need comes from the topological component of the theorem which expresses a global property. In fact, a surjectiveness claim is a claim of involvement for *all* points of the codomain space.

2.1.2. REMARK (INFORMATION TRANSFER). Various (bi-)simulations transfer logical information across topological spaces. A case in point are ‘Chu morphisms’ relating topological spaces that are ‘adjoint’ in an abstract sense (cf. [van Benthem, 1998]). Existential modal formulas are then mirrored in general first-order ‘flow formulas’.

2.1.3 Topo-bisimilar reductions

In many contexts, bisimulations and simulations are used to find minimal models. This is useful, for instance, to find minimal representations for labeled transition systems having certain desired properties modally expressible. Topo-bisimulation can be used for finding a minimal representation for a determined spatial configuration. For example, consider a spoon with two handles, as depicted in Figure 2.6.a. The spoon has 7 ‘salient’ points, these satisfy the formulas reported in Figure 2.4.

Point	Formula
1	$\Box p$
2	$\Diamond p \wedge \Diamond \neg p$
3	$\Box \neg p$
4	$p \wedge \neg \Diamond \Box p \wedge \Diamond \Box \neg p$
5	$\Diamond \Box p \wedge \Diamond (p \wedge \neg \Diamond \Box p)$
6	$p \wedge \neg \Diamond \Box p \wedge \Diamond \Box \neg p$
7	$\Diamond \Box p \wedge \Diamond (p \wedge \neg \Diamond \Box p)$

Figure 2.4: Formulas true at points of the model in Figure 2.5.

It is easy to find an **S4** Kripke model satisfying the 7 formulas above, for instance, the one in Figure 2.5.a. By a bisimulation one ‘reduces’ it to a minimal similar one. The topo-bisimilar reduction is presented in the table on the right of Figure 2.6.

From the reduced model one can ‘reconstruct’ the pictorial example, that is, a spoon with only one handle, Figure 2.6.b. Checking the topo-bisimilarity of Figure 2.6.a and Figure 2.6.b is an easy task to perform. We do not spell out the general method used here for transforming topological models into Kripke ones (and back); but it should be fairly clear from the example.

The claim is not that one should move back and forth from topological and Kripke semantics to find minimal models. Our goal is to show that topo-bisimulations enable the reduction of spatial models in the same way that bisimulations enable the reduction of Kripke models. A general algorithm for deciding topo-bisimulation is still missing, but one for a specific class of models will be presented and used in Chapter 6.

2.2 Games that compare visual scenes

Topo-bisimulation is a global notion of comparison. But in practice, we are interested in *fine-structure*: what are the ‘simplest differences’ that can be detected between two visual scenes? For this purpose, we introduce *topo-games* topological game that generalize *Ehrenfeucht-Fraïssé* comparison games between first-order models, see [Doets, 1996]. Similarity and difference between visual scenes will then have to do with *strategies* for players comparing them.

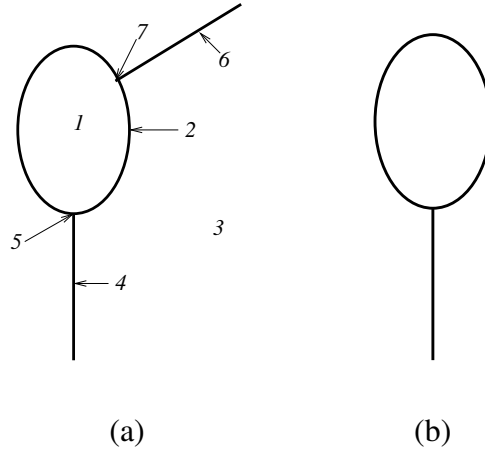


Figure 2.5: The reduction of a topological model to a minimal topo-bisimilar one. From a spoon with two handles to one with only one.

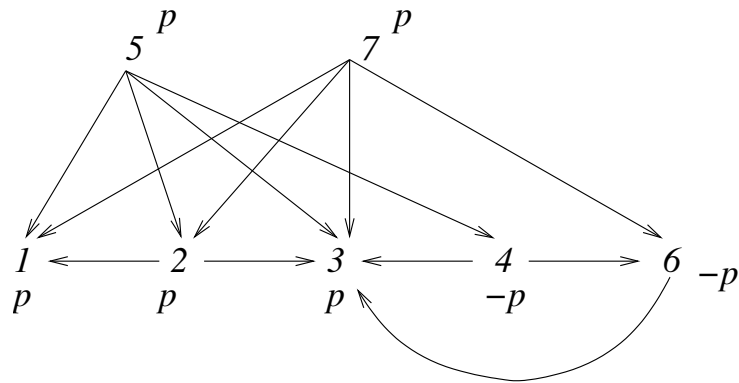
2.2.1. DEFINITION (TOPOLOGICAL GAME). Consider two topo-models $\langle X, O, \nu \rangle$, and $\langle X', O', \nu' \rangle$, a natural number n and two points $x_1 \in X, x'_1 \in X'$. A *topological game* of length n , with starting points x_1, x'_1 —notation $TG(X, X', n, x_1, x'_1)$ —consists of n rounds between two players: Spoiler and Duplicator. Each round proceeds as follows:

- (i) Spoiler chooses a model X_s and an open o_s containing the current point x_s of that model
- (ii) Duplicator chooses an open o_d in the other model X_d containing the current point x_d of that model
- (iii) Spoiler picks a point \bar{x}_d in Duplicator's open o_d in the X_d model
- (iv) Duplicator finally picks a point \bar{x}_s in Spoiler's open o_s in X_s

The points \bar{x}_s and \bar{x}_d become the new current points of the X_s and X_d models, respectively. After n rounds, two sequences have been built:

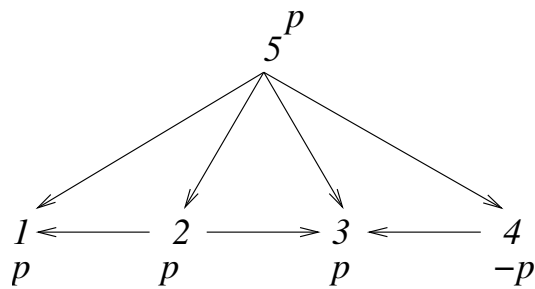
$$\{x_1, o_1, x_2, o_2, \dots, o_{n-1}, x_n\} \quad \{x'_1, o'_1, x'_2, o'_2, \dots, o'_{n-1}, x'_n\}$$

with $x_i \in o_i$, and $o_i \in O$ (analogously for the second sequence). After n rounds, if x_i and x'_i (with $i \in [1, n]$) satisfy the same atoms, Duplicator *wins*. (Note that Spoiler already wins ‘en route’, if Duplicator fails to maintain the atomic match.) A *winning strategy* (‘w.s.’ for short) for Duplicator is a function from any sequence of moves by Spoiler to appropriate responses which always ends in a win for Duplicator. The same notion applies to Spoiler. An *infinite topological game* is one without a finite limit to the number of rounds. In this case, Duplicator wins if the matched points continue to satisfy the same atoms.



(a)

(a)	(b)
1	1
2	2
3	3
4	4
5	5
6	4
7	5



(b)

Figure 2.6: The reduction of the spoons of Figure 2.5 via a bisimulation on the corresponding Kripke models. In the table, the bisimulation relation.

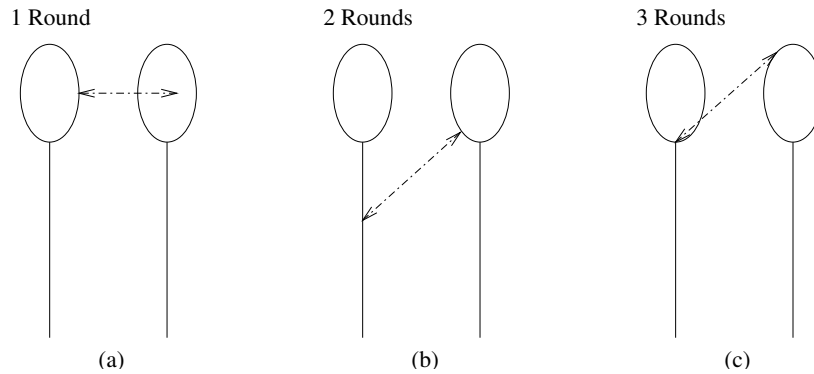


Figure 2.7: Games on two spoons with two different starting points. On top, the number of rounds needed by Spoiler to win.

The opens in the game sequence do not play any role in determining which player wins, but they visually guide the development of the game. For instance, the following intuitive ‘Locality Principle’ holds. Players lose no winning strategies if we restrict their moves to choosing opens that are *contained in the previous open*.

2.2.1. EXAMPLE (PLAYING ON SPOONS). Consider the three configurations in Figure 2.7. (a) The leftmost game starts with a point on the boundary of the spoon versus an interior point of the other spoon. Spoiler can win this game in one round by simply choosing an open set on the right spoon completely contained in its interior. Duplicator’s open response must always contain a point not in the spoon, which Spoiler can then pick, giving Duplicator no possible response. (b) In the central game, a point on the handle is compared with a boundary point of the spoon’s container. Spoiler can again win the game, but needs two rounds this time. Here is a winning strategy. First, Spoiler chooses an open on the left spoon containing the starting point but without interior points. Any open chosen by Duplicator on the other spoon must contain an interior point. Spoiler then picks such an interior point. Duplicator’s response to that can only be a boundary point of the other model (on the handle) or a point outside of the spoon. In the latter case, she loses at once – in the former, she loses in one round, by reduction to the previous game. (c) Finally, on the left the junction between handle and container is compared with a boundary point of the container. In this game, Spoiler will choose an open on the right model, avoiding points on the handle of the spoon. Duplicator is forced to choose an open on the left containing points on the handle. Spoiler then picks such a handle point. Duplicator replies either with an interior point, or with a boundary point of the right spoon. Thus we are back with game (b), and Spoiler can win in the remaining two rounds.

The topological dynamics of these games is appealing. E.g. it is instructive to check that other initial choices for Spoiler may very well lead to his losing the game! (E.g., let Spoiler start in the right-hand model in (b)). A strategy guarantees a win only for

those who follow it... One can also make some more general mathematical observations here. In particular, topo-games are always *determined*: either Duplicator has a winning strategy, or Spoiler has one.

2.2.1 Strategies and modal formulas

The fine-structure provided by games measures differences in terms of the minimum number of rounds needed by Spoiler to win. These same differences may also be formulated in terms of our modal language. To see this, we need the notion of *modal rank*, being the maximum number of nested modal operators in a formula. For instance, the modal ranks of the formulas in Figure 2.1: p , $\Box p$, $p \wedge \neg \Box p$, $\Diamond \Box p$, $p \wedge \neg \Diamond \Box p$, $\Diamond \Box p \wedge \Diamond(p \wedge \neg \Diamond \Box p)$, are 0, 1, 1, 2, 2, and 3, respectively. We are now ready for our main result.

2.2.2. THEOREM (ADEQUACY). topological game!adequacy Duplicator has a w.s. in $TG(X, X', n, x, x')$ iff x and x' satisfy the same formulas of modal rank up to n .

Proof The left to right direction is proven by induction on the length n of the game $TG(X, X', n, x, x')$. If $n = 0$ and Duplicator has a winning strategy, this means that the points x, x' satisfy the same proposition letters, and hence the same Boolean combinations of proposition letters, i.e., the same modal formulas of modal rank 0. Now for the inductive step. Suppose that Duplicator has a winning strategy σ in $TG(X, X', n, x, x')$. We want to show that $X, x \models \varphi$ iff $X', x' \models \varphi$, when the modal rank of φ is n . By simple syntactic inspection, φ must be a Boolean combination of formulas of the form $\Box \psi$ where ψ has modal rank less or equal to $n - 1$. Thus, it suffices to prove that $X, x \models \Box \psi$ iff $X', x' \models \Box \psi$. Without loss of generality, let us consider the first model. Suppose that $X, x \models \Box \psi$. By the truth definition there exists an open o (with $x \in o$) such that $\forall u \in o : X, u \models \psi$. Now, assume that the n -round game starts with Spoiler choosing o in X . Using the strategy σ , Duplicator can pick an open o' such that $x' \in o'$ and $\forall u' \in o' : X, u' \models \psi$. Now Spoiler can pick any point u' in o' . Duplicator can use the information in σ to respond with a point $u \in o$, concluding the first round, so that the remaining strategy σ' is still winning for Duplicator in $TG(X, X', n - 1, u, u')$. By the inductive hypothesis, the fact that $X, u \models \psi$ (where ψ has modal rank $n - 1$) implies that $X', u' \models \psi$. Thus we have shown that all $u' \in o'$ satisfy ψ , and hence $X', x' \models \Box \psi$. The other direction is analogous.

The right to left direction is again proven by induction on n . If $n = 0$, then x and x' satisfy the same non-modal formulas. In particular, they satisfy the same atoms, which is winning for Duplicator, by the definition of topological game. Now for the inductive step. Without loss of generality, let us assume that Spoiler picks an open set o containing x in X in the first round of $TG(X, X', n, x, x')$ game. Now, take the set $\{\text{DES}_{n-1}(z) \mid z \in o\}$, where $\text{DES}_{n-1}(z)$ denotes all the formulas up to modal rank $n - 1$ satisfied at z . This set is not finite per se, but we can simply prove the following

2.2.3. FACT (LOGICAL FINITENESS). There are only finitely many modal formulas of depth k up to logical equivalence.

Therefore, we can write one Boolean formula to describe this open set o , namely $\bigvee \bigwedge \text{DES}_{n-1}(z)$. Since this is true for all $z \in o$, by the truth definition we have that $X, x \models \Box \bigvee \bigwedge \text{DES}_{n-1}(z)$ (a formula of modal rank n). By hypothesis, x and x' satisfy the same modal formulas of modal rank n , so $X', x' \models \Box \bigvee \bigwedge \text{DES}_{n-1}(z)$. This last fact, together with the truth definition implies that there exists an open o' such that $\forall z' \in o' : X', z' \models \bigvee \bigwedge \text{DES}_{n-1}(z)$. This is the open that Duplicator must choose to reply to Spoiler's move. Now Spoiler can pick any point u' in o' . Such a point satisfies at least one disjunct $\bigwedge \text{DES}_{n-1}(z)$, and we let Duplicator respond with $z \in o$. As a result of this first round, z, u' satisfy the same modal formulas up to modal depth $n - 1$. Hence by the inductive hypothesis, Duplicator has a winning strategy for $TG(X, X', n - 1, z, u')$. Putting this together with our first instruction, we have a winning strategy for Duplicator in the n -round game. QED

This is the usual version of adequacy: slanted towards similarity. But in our pictorial examples, we rather looked at Spoiler. One can also set up the proof of Theorem 2.2.2 so as to obtain an *effective correspondence* between (a) winning strategies for Spoiler, (b) modal 'difference formulas' for the initial points. Here is an illustration.

2.2.2. EXAMPLE (MATCHING STRATEGIES WITH FORMULAS). Look again at Figure 2.7. The strategies described for Spoiler are immediately linked to modal formulas that distinguish the two models. Suppose the spoons are denoted by the proposition letter p and hence the background by $\neg p$. In the game on the left, $\Box p$ is true of the starting point of the right spoon, and its negation $\Diamond \neg p$ is true of the starting point of the other spoon. The modal depth of these formulas is one and therefore Spoiler can win in one round. In the central case, a distinguishing formula is $\neg \Diamond \Box p$, which holds for the starting point on the left spoon, but not for that on the right. The modal depth is 2, which is the number of rounds that Spoiler needed to win the game. Finally, a formula of modal depth 3 that is only true of the point on the left spoon of the leftmost game is: $\Diamond(p \wedge \neg \Diamond \Box p)$. The negation of this formula is true on the other starting point, thus justifying Spoiler's winning strategy in 3 turns.

There is still more fine-structure to these games. E.g., visual scenes may have several modal differences, and hence more than one winning strategy for Spoiler. Also, recall that topo-games can be played *infinitely*. Then the winning strategies for Duplicator (if any) are precisely the various *topo-bisimulations* between the two models. For further details, see [Aiello and van Benthem, 1999], [van Benthem, 1999]—and also [Barwise and Moss, 1996].

Before considering completeness of **S4** with respect to topological spaces in the next chapter, we remark an alternative modal approach to axiomatizing topology.

2.3 Logical variations

Tarski's interior modality \Box iff $\exists o \in O : x \in o \wedge \forall y \in o : M, y \models \varphi$ is actually a mixture of elements of different sorts. A $\Box\varphi$ formula is true in a point x whenever there exist an open set containing the point x itself and such that all points of the set satisfy φ . The definition quantifies at the same time over points and over sets of points connected by the incidence relation of set membership. Naturally, there is an alternative take on the basic topological approach to topological reasoning: a 'stepwise' approach separating points from open sets, thus splitting Tarski's modality into two separate modal quantifiers. The resulting modal logic was studied in [Dabrowski et al., 1996] and in Georgatos' PhD thesis [Georgatos, 1993]. The main motivation of their work is that of modeling, with "weak logical systems whose primitives are appropriately chosen," logics of knowledge. In particular, with such a logic one can focus on the notion of effort in contraposition with that of view. The authors also explicitly mention the added motivation of having devised a tool of potential use for visual reasoning. We share the motivation and here place their language in our map of spatial logics to tour.

The definition of a model is analogous to that of topological models presented in Section 2.1 and the truth definition for the new modal operators becomes:

$$\begin{aligned} M, x, o \models \Box\varphi & \text{ iff } \forall y \in o : M, y, o \models \varphi \\ M, x, o \models \Diamond\varphi & \text{ iff } \forall o' \subseteq o \in O : x \in o' \wedge M, x, o' \models \varphi \end{aligned}$$

where $x, y \in X$ are points and $o, o' \in O$ are open sets. The relation with Tarski's interior modality is quite straightforward:

$$\Box\varphi \text{ if } \Diamond\Box\varphi$$

Proof The truth definition of the formula $\Diamond\Box\varphi$ states $M, x, o \models \forall o' \subseteq o \in O : x \in o' \wedge \forall y \in o' : M, y, o \models \varphi$. On the other hand, in the truth definition of \Box there is no reference to an open set, so the previous truth definition becomes $M, x \models \exists o \forall o' \subseteq o \in O : x \in o' \wedge \forall y \in o' : M, y, o \models \varphi$, which trivially simplifies to $M, x \models \exists o x \in o \wedge \forall y \in o : M, y, o \models \varphi$ which is precisely the definition of $\Box\varphi$.

QED

The two level language affords a nice new view on the **S4**-behavior of our original topological interpretation. E.g., consider the behavior of the **S4** axioms.

$$\Box\varphi \rightarrow \varphi \text{ becomes } \Diamond\Box\varphi \rightarrow \varphi, \quad (2.1)$$

which, in a two-sorted modal logic, expresses the fact that the accessibility relation for s is contained in the *converse* of that for p . This is a natural connection between ' $x \in A$ ' and ' $A \ni x$ '. Note that *reflexivity* vanishes!

$$\Box\varphi \rightarrow \Box\Box\varphi \text{ becomes } \Diamond\Box\varphi \rightarrow \Diamond\Box\Diamond\Box\varphi, \quad (2.2)$$

which follows from

$$\Box\varphi \rightarrow \Box\Diamond\Box\varphi$$

which is simply a minimally valid consequence of conversion ($\psi \rightarrow \Box\Diamond\psi$). The rest is an application of the valid modal base rule “from $\gamma \rightarrow \sigma$ to $\Diamond\gamma \rightarrow \Diamond\sigma$.”

$$\Box\varphi \wedge \Box\psi \rightarrow \Box(\varphi \wedge \psi) \text{ becomes } \Diamond\Box\varphi \wedge \Diamond\Box\psi \rightarrow \Diamond\Box(\varphi \wedge \psi), \quad (2.3)$$

a principle which has no obvious meaning in a two-sorted modal language. We can analyze its meaning by *frame correspondence* techniques [Blackburn et al., 2001], to obtain:

$$\forall A, B : ((x \in A \wedge x \in B) \rightarrow \exists C : (x \in C \wedge \forall y \in C : y \in A \vee y \in B)).$$

The full axiomatization of the logic is known [Dabrowski et al., 1996]. The set modality \Diamond has the **S5** axiomatization, while the point modality \Box retains the **S4** axiomatization. Depending on which models we consider there is a number of different interaction axioms that also hold. If we consider models for which the set O follows the laws of open spaces, rather than just being a family of subsets with no specific structure (cf. neighbourhood semantics), one gets:

$$\begin{aligned} \Diamond\Box\varphi &\rightarrow \Box\Diamond\varphi && \text{(Cross)} \\ \Diamond\varphi \wedge \Box\Diamond\psi &\rightarrow \Diamond(\Diamond\varphi \wedge \Box\Diamond\psi \wedge \Diamond\Diamond\Box(\varphi \vee \psi)) && \text{(Union)} \end{aligned}$$

Either way, whether by a single modality defined by a second-order existential and an universal quantifiers or by a two-sorted modal logic defined by first-order quantifications, there is a landscape of possible modal languages for topological patterns whose nature is by no means understood. For instance, one would like to understand what are natural well-chosen languages for simulations, and also, what are the *complexity jumps* between languages and their logics in this spectrum.

CHAPTER 3

THE TOPO APPROACH: AXIOMATICS

Regarding the modal box as an interior operator, one gets the feeling for why the modal logic **S4** is complete with respect to arbitrary topological spaces as modal logic axioms mimic Kuratowski's topological axioms. But there are classical results with much more mathematical content, such as McKinsey and Tarski's beautiful theorem stating that **S4** is the complete logic of the reals, and indeed of any metric separable space without isolated points. Even so, the topological interpretation has always remained something of a side-show in modal logic and intuitionistic logic, often tucked away in notes and appendices. The purpose of this chapter is to take it one step further as a first stage in a program of independent interest, viz. the modal analysis of space—showing how one can get more generality, as well as some nice new questions. In particular, this chapter contains (a) a modern analysis of the modal language **S4** as presented in Chapter 2 in terms of 'topo-bisimulation', (b) a number of connections between topological models and Kripke models, (c) a new general proof of McKinsey and Tarski's Theorem (inspired by [Mints, 1998]), (d) an analysis of special topological logics on the reals, pointing toward a landscape of spatial logics above **S4**.

3.1 Topological spaces and Kripke models

The purpose of this section is a link-up with the better-known world of 'standard' semantics for modal logic. At the same time, this comparison increases our understanding of the 'topological content' of modal logic.

3.1.1 The basic connection

The standard Kripke semantics is a particular case of its more general topological semantics. Recall that an **S4**-frame (henceforth 'frame', for short) is a pair $\langle W, R \rangle$, where W is a non-empty set and R a quasi-order (transitive and reflexive) on W . Call a set $X \subseteq W$ *upward closed* if $w \in X$ and wRv imply $v \in X$.

3.1.1. FACT. Every frame $\langle W, R \rangle$ induces a topological space $\langle W, \tau_R \rangle$, where τ_R is the set of all upward closed subsets of $\langle W, R \rangle$.

It is easy to check that τ_R is a topology on W , and that the closure and interior operators of $\langle W, \tau_R \rangle$ are respectively $R^{-1}(X)$ and $W - R^{-1}(W - X)$, where $R^{-1}(w) = \{v \in W \mid vRw\}$ and $R^{-1}(X) = \bigcup_{w \in X} R^{-1}(w)$, for $w \in W, X \subseteq W$. Indeed, τ_R is a rather special topology on W : for any family $\{X_i\}_{i \in I} \subseteq \tau_R$, we have $\bigcap_{i \in I} X_i \in \tau_R$. Such spaces are called *Alexandroff spaces*, in which every point has a least neighborhood. In frames, the least neighborhood of a point w is evidently $\{v \in W \mid wRv\}$, which is usually denoted by $R(w)$.

Conversely, every topological space $\langle W, \tau \rangle$ naturally induces a quasi-order R_τ defined by putting

$$wR_\tau v \text{ iff } w \in \overline{\{v\}} \text{ iff } w \in U \text{ implies } v \in U, \text{ for every } U \in \tau.$$

This is called the *specialization order* in the topological literature. Again it is easy to check that R_τ is transitive and reflexive, and that every open set of τ is R_τ -upward closed. Moreover, R_τ is anti-symmetric iff $\langle W, \tau \rangle$ satisfies the T_0 separation axiom (that is, any two different points are separated by an open set). Hence R_τ is a partial order iff $\langle W, \tau \rangle$ is a T_0 -space.

Combining the two mappings, $R = R_{\tau_R}, \tau \subseteq \tau_{R_\tau}$, and $\tau = \tau_{R_\tau}$ iff $\langle W, \tau \rangle$ is an Alexandroff space. Indeed, $wR_{\tau_R}v$ iff $w \in \overline{\{v\}}$ iff $w \in R^{-1}(v)$ iff wRv . Also, as every open set of τ is R_τ -upward closed, $\tau \subseteq \tau_{R_\tau}$. Finally, $\tau = \tau_{R_\tau}$ iff every R_τ -upward closed set belongs to τ iff every point of W has a least neighborhood in $\langle W, \tau \rangle$ iff $\langle W, \tau \rangle$ is an Alexandroff space.

The upshot of all this is a one-to-one correspondence between quasi-ordered sets and Alexandroff spaces, and between partially ordered sets and Alexandroff T_0 -spaces. Since every finite topological space is an Alexandroff space, this immediately gives a one-to-one correspondence between finite quasi-ordered sets and finite topological spaces, and finite partially ordered sets and finite T_0 -spaces.

There is also a one-to-one correspondence between continuous maps and order preserving maps, as well as open maps and p -morphisms. Indeed, let two topological spaces $\langle W_1, \tau_1 \rangle$ and $\langle W_2, \tau_2 \rangle$ be given. Recall that a function $f : W_1 \rightarrow W_2$ is *continuous* if $f^{-1}(V) \in \tau_1$ for every $V \in \tau_2$. Moreover, f is *open* if it is continuous and $f(U) \in \tau_2$ for every $U \in \tau_1$. It is well-known that f is continuous iff $\overline{f^{-1}(X)} \subseteq f^{-1}(\overline{X})$, and that f is open iff $\overline{f^{-1}(X)} = f^{-1}(\overline{X})$, for every $X \subseteq W_2$.

Next, for two quasi-orders $\langle W_1, R_1 \rangle$ and $\langle W_2, R_2 \rangle$, $f : W_1 \rightarrow W_2$ is said to be *order preserving* if wR_1v implies $f(w)R_2f(v)$, for $w, v \in W_1$. f is a *p -morphism* if it is order preserving, and in addition $f(w)R_2v$ implies that there exists $u \in W_1$ such that wR_1u and $f(u) = v$, for $w \in W_1$ and $v \in W_2$. It is well-known that f is order preserving iff $R_1^{-1}f^{-1}(w) \subseteq f^{-1}R_2^{-1}(w)$, and that f is a p -morphism iff $R_1^{-1}f^{-1}(w) = f^{-1}R_2^{-1}(w)$, for every $w \in W_2$.

Putting this together, one easily sees that f is monotone iff f is continuous, and that f is p -morphism iff f is open.

As an easy consequence we obtain that the category **ATop** of Alexandroff spaces and continuous maps is isomorphic to the category **Qos** of quasi ordered sets and order preserving maps, and that the category **ATop**⁺ of Alexandroff spaces and open maps is isomorphic to the category **Qos**⁺ of quasi ordered sets and p -morphisms. Similarly, the category **ATop** _{T_0} of Alexandroff T_0 -spaces and continuous maps is isomorphic to the category **Pos** of partially ordered sets and order preserving maps, and the category **ATop** _{T_0} ⁺ of Alexandroff T_0 -spaces and open maps is isomorphic to the category **Pos**⁺ of partially ordered sets and p -morphisms.

In the finite case, we get that the category **FinTop** of finite spaces and continuous maps is isomorphic to the category **FinQos** of finite quasi ordered sets and order preserving maps, and that the category **FinTop**⁺ of finite topological spaces and open maps is isomorphic to the category **FinQos**⁺ of finite quasi ordered sets and p -morphisms. Similarly, the category **FinTop** _{T_0} of finite T_0 -spaces and continuous maps is isomorphic to the category **FinPos** of finite partially ordered sets and order preserving maps, and the category **FinTop** _{T_0} ⁺ of finite T_0 -spaces and open maps is isomorphic to the category **FinPos**⁺ of finite partially ordered sets and p -morphisms.

3.1.2 Analogies

The tight connection between modal frames and topological spaces explains the earlier-mentioned analogies in their semantic development, such as *locality* and *invariance for bisimulation*. It may be extended to include other basic modal topics, such as *correspondence theory* [van Benthem, 1985]. Likewise, the modern move toward extended modal languages makes equally good sense for the topological interpretation. Many natural topological notions need extra modal power for their definition: good examples are the basic *separation axioms*. We just saw that, among the quasi orders, partial orders correspond to topological spaces satisfying the T_0 separation axiom. But this difference does not show up in our basic modal language: **S4** is complete with respect to arbitrary partial orders. Defining separation axioms requires various expressive extensions of the modal base language.

Finally, in a more technical sense, there still seems to be a vast difference. The format of the topological interpretation looks more complex than the usual one which quantifies over accessible worlds only. For, it involves a *second-order* quantification over sets of worlds, plus a first-order quantification over their members. But this difference is more apparent than real, because the quantification is over open sets only, and we may plausibly think of topological models as *two-sorted first-order models* with separate domains of ‘points’ and ‘opens’, see Section 2.3.

3.2 General completeness

The preceding section shows that standard modal models are a particular case of a more general topological semantics. Hence, the known completeness of **S4** plus the

topological soundness of its axioms immediately give us general topological completeness. Even so, we now give a direct model-theoretic proof of this result. It is closely related to the standard modal Henkin construction, but with some nice topological twists. (Compare [Chellas, 1980] for the quite analogous case of modal ‘neighborhood semantics’.)

3.2.1 The main argument

Soundness is immediate, and hence we move directly to completeness. Call a set Γ of formulas of \mathcal{L} (**S4**)-consistent if for no finite set $\{\varphi_1, \dots, \varphi_n\} \subseteq \Gamma$ we have that $\mathbf{S4} \vdash \neg(\varphi_1 \wedge \dots \wedge \varphi_n)$. A consistent set of formulas Γ is called *maximally consistent* if there is no consistent set of formulas properly containing Γ . It is well-known that Γ is maximally consistent iff, for any formula φ of \mathcal{L} , either $\varphi \in \Gamma$ or $\neg\varphi \in \Gamma$, but not both. Now we define a topological space out of maximally consistent sets of formulas.

3.2.1. DEFINITION (CANONICAL SPACE). The *canonical topological space* is the pair $S^\mathcal{L} = \langle W^\mathcal{L}, \tau^\mathcal{L} \rangle$, where:

- $W^\mathcal{L}$ is the set of all maximally consistent sets Γ_{max} ;
- $\tau^\mathcal{L}$ is the set generated by arbitrary unions of the following *basic sets* $B^\mathcal{L} = \{\widehat{\Box\varphi} \mid \varphi \text{ is any formula}\}$, where $\widehat{\varphi} =_{\text{def}} \{\Gamma_{max} \in W^\mathcal{L} \mid \varphi \in \Gamma_{max}\}$. In other words, basic sets are the families of the form: $U_\varphi = \{\Gamma_{max} \in W^\mathcal{L} \mid \Box\varphi \in \Gamma_{max}\}$.

Let us first check that $S^\mathcal{L}$ is indeed a topological space.

3.2.2. LEMMA. $B^\mathcal{L}$ forms a basis for the topology.

Proof We only need to show the following two properties:

- For any $U_\varphi, U_\psi \in B^\mathcal{L}$ and any $\Gamma_{max} \in U_\varphi \cap U_\psi$, there is $U_\chi \in B^\mathcal{L}$ such that $\Gamma_{max} \in U_\chi \subseteq U_\varphi \cap U_\psi$;
- For any $\Gamma_{max} \in W^\mathcal{L}$, there is $U_\varphi \in B^\mathcal{L}$ such that $\Gamma_{max} \in U_\varphi$.

Now, (N) implies that $\Box\top \in \Gamma_{max}$, for any Γ_{max} . Hence $W^\mathcal{L} = \widehat{\Box\top}$ and the second item is satisfied. As for the first item, thanks to (R), one can easily check that $\widehat{\Box(\varphi \wedge \psi)} = \widehat{\Box\varphi} \cap \widehat{\Box\psi}$. Hence $U_\varphi \cap U_\psi \in B^\mathcal{L}$, and so $B^\mathcal{L}$ is closed under finite intersections: whence the first item is satisfied. QED

Next we define the canonical topological model.

3.2.3. DEFINITION (CANONICAL MODEL). The *canonical topological model* is the pair $M^\mathcal{L} = \langle S^\mathcal{L}, \nu^\mathcal{L} \rangle$, where:

- $S^{\mathcal{L}}$ is the canonical topological space;
- $\nu^{\mathcal{L}}(P) = \{\Gamma_{max} \in X^{\mathcal{L}} \mid P \in \Gamma_{max}\}$.

The valuation $\nu^{\mathcal{L}}$ equates truth of a proposition letter *at* a maximally consistent set with its membership *in* that set. We now show this harmony between the two viewpoints lifts to all formulas.

3.2.4. LEMMA (TRUTH LEMMA). For all modal formulas φ ,

$$M^{\mathcal{L}}, w \models_{\mathcal{L}} \varphi \text{ iff } w \in \widehat{\varphi}.$$

Proof Induction on the complexity of φ . The base case was just described. The case of the Booleans follows from the following well-known identities for maximally consistent sets:

- $\widehat{\neg\varphi} = W^{\mathcal{L}} - \widehat{\varphi}$;
- $\widehat{\varphi \wedge \psi} = \widehat{\varphi} \cap \widehat{\psi}$.

The interesting case is that of the modal operator \Box . We do the two relevant implications separately, starting with the easy one.

\Leftarrow ‘From membership to truth.’ Suppose $w \in \widehat{\Box\varphi}$. By definition, $\widehat{\Box\varphi}$ is a basic set, hence open. Moreover, thanks to axiom (T), $\widehat{\Box\varphi} \subseteq \widehat{\varphi}$. Hence there exists an open neighborhood $U = \widehat{\Box\varphi}$ of w such that for any $v \in U$, $v \in \widehat{\varphi}$, and by the induction hypothesis, $M^{\mathcal{L}}, v \models_{\mathcal{L}} \varphi$. Thus $M^{\mathcal{L}}, w \models_{\mathcal{L}} \Box\varphi$.

\Rightarrow ‘From truth to membership.’ Suppose $M^{\mathcal{L}}, w \models_{\mathcal{L}} \Box\varphi$. Then there exists a basic set $\widehat{\Box\psi} \in B^{\mathcal{L}}$ such that $w \in \widehat{\Box\psi}$ and for all $v \in \widehat{\Box\psi}$, $M^{\mathcal{L}}, v \models_{\mathcal{L}} \varphi$. By the induction hypothesis, $\forall v \in \widehat{\Box\psi}, v \in \widehat{\varphi}$: i.e., $\widehat{\Box\psi} \subseteq \widehat{\varphi}$. But this implies that the logic **S4** can prove the implication $\Box\psi \rightarrow \varphi$. (If not, then there would be some maximally consistent set containing both $\Box\psi$ and $\neg\varphi$.) But then we can prove the implication $\Box\Box\psi \rightarrow \Box\varphi$, and hence, using the **S4** transitivity axiom, $\Box\psi \rightarrow \Box\varphi$. It follows that $\widehat{\Box\psi} \subseteq \widehat{\Box\varphi}$, and hence the world w belongs to $\widehat{\Box\varphi}$. QED

Now we can clinch the proof of our main result.

3.2.5. THEOREM (COMPLETENESS). For any set of formulas Γ ,

$$\text{if } \Gamma \models_{\mathcal{L}} \varphi \text{ then } \Gamma \vdash_{\mathbf{S4}} \varphi.$$

Proof Suppose that $\Gamma \vdash_{\mathbf{S4}} \varphi$. Then $\Gamma \cup \{\neg\varphi\}$ is consistent, and by the Lindenbaum Lemma it can be extended to a maximally consistent set Γ_{max} . By the Truth Lemma, $M^{\mathcal{L}}, \Gamma_{max} \models_{\mathcal{L}} \neg\varphi$, whence $\Gamma_{max} \not\models_{\mathcal{L}} \varphi$, and we have constructed the required counter-model. QED

3.2.2 Topological comments

Let us now look at some topological aspects of this construction. In proving the box case of Truth Lemma, we did not use the standard modal argument, which crucially invokes the *distribution axiom* of the minimal modal logic. Normally, one shows that, if a formula $\Box\phi$ does not belong to a maximally consistent set Γ , then there exists some maximally consistent successor set of Γ containing $\neg\phi$. This is not necessary in the topological version at this stage. We only need the reflexivity and transitivity axioms, plus the Lindenbaum Lemma on maximally consistent extensions. The modal distribution axiom still plays a crucial role, but that was at the earlier stage of verifying that we had really defined a topology. This different way of ‘cutting the cake’ provides an additional proof-theoretic explanation why **S4** is the weakest axiom system complete for topological semantics. Moreover, the divergence with the ‘standard’ argument explodes the prejudice that one single ‘well-known’ interpretation for a language must be the only natural one. Comparing our construction with the standard Henkin model for **S4** $\langle W^\mathcal{L}, R^\mathcal{L}, \models_\mathcal{L} \rangle$, the basic sets of our topology $S^\mathcal{L}$ are $R^\mathcal{L}$ -upward closed. Hence every open of $S^\mathcal{L}$ is $R^\mathcal{L}$ -upward closed, and $S^\mathcal{L}$ is weaker than the topology $\tau_{R^\mathcal{L}}$ corresponding to $R^\mathcal{L}$. In particular, our canonical space is *not* an Alexandroff space.

Here are some further topological aspects of the above construction. First, it is worthwhile to compare Stone’s famous construction which uses the alternative basis $\{\widehat{\varphi} \mid \varphi \text{ any formula}\}$, yielding a space which we denote by $\langle W^\mathcal{L}, \tau^\mathcal{S} \rangle$. It is well-known that $\langle W^\mathcal{L}, \tau^\mathcal{S} \rangle$ is homeomorphic to the Cantor space—and so, up to homeomorphism, $\langle W^\mathcal{L}, \tau^\mathcal{S} \rangle$ is compact, metric, 0-dimensional, and dense-in-itself. The basis of our topology, however, was the sub-family $\{\widehat{\Box\varphi} \mid \varphi \text{ any formula}\}$. Now every subtopology of one that is compact and dense-in-itself is also compact and dense-in-itself. Therefore, we get these same properties for our canonical topological space. But we can be more precise than this.

3.2.6. FACT. *The canonical topology is actually the intersection of the Kripke and Stone topologies.*

In other words, $\tau^\mathcal{L} = \tau_{R^\mathcal{L}} \cap \tau^\mathcal{S}$. Indeed, since $\tau^\mathcal{L} \subseteq \tau_{R^\mathcal{L}}$ and $\tau^\mathcal{L} \subseteq \tau^\mathcal{S}$, obviously $\tau^\mathcal{L} \subseteq \tau_{R^\mathcal{L}} \cap \tau^\mathcal{S}$. Conversely, since every base set $\widehat{\varphi}$ of Stone’s topology is $R^\mathcal{L}$ -upward closed iff $\widehat{\varphi} = \widehat{\Box\psi}$ for some ψ , $\tau_{R^\mathcal{L}} \cap \tau^\mathcal{S} \subseteq \tau^\mathcal{L}$, and $\tau^\mathcal{L} = \tau_{R^\mathcal{L}} \cap \tau^\mathcal{S}$.

One can also connect modal formulas and topological properties more directly, by giving a direct proof of the fact that $S^\mathcal{L}$ is compact and dense-in-itself. The former fact goes just as for the Stone space, but we display it for the sake of illustration.

3.2.7. LEMMA. *$S^\mathcal{L}$ is compact.*

Proof Ad absurdum, there is a family $\{\widehat{\Box\psi_i}\}_{i \in I} \subseteq B^\mathcal{L}$ such that $\bigcup_{i \in I} \widehat{\Box\psi_i} = W^\mathcal{L}$, and for no finite subfamily $\{\widehat{\Box\psi_{i_1}}, \dots, \widehat{\Box\psi_{i_n}}\}$ we have $\widehat{\Box\psi_{i_1}} \cup \dots \cup \widehat{\Box\psi_{i_n}} = W^\mathcal{L}$. Let $\Gamma = \{\neg\Box\psi_i\}_{i \in I}$.

3.2.8. CLAIM. Γ is consistent.

Proof Ad absurdum, there is a finite number of formulas $\neg\Box\psi_1, \dots, \neg\Box\psi_n \in \Gamma$ such that $\mathbf{S4} \vdash \neg(\neg\Box\psi_1 \wedge \dots \wedge \neg\Box\psi_n)$. Hence $\mathbf{S4} \vdash \Box\psi_1 \vee \dots \vee \Box\psi_n$. But then $\widehat{\Box\psi_1} \cup \dots \cup \widehat{\Box\psi_n} = W^\mathcal{L}$, which is a contradiction. QED

Since Γ is consistent, it can be extended to a maximally consistent set Γ_{max} . Obviously $\neg\Box\psi_i \in \Gamma_{max}$ for any $i \in I$. Hence $\Gamma_{max} \in \widehat{\neg\Box\psi_i}$ for any $i \in I$. Since $\widehat{\neg\Box\psi_i} = W^\mathcal{L} - \widehat{\Box\psi_i}$, $\Gamma_{max} \in W^\mathcal{L} - \widehat{\Box\psi_i}$ for any $i \in I$. Hence $\Gamma_{max} \in W^\mathcal{L} - \bigcup_{i \in I} \widehat{\Box\psi_i}$, which contradicts our assumption. Thus, $S^\mathcal{L}$ is compact. QED

3.2.9. LEMMA. $S^\mathcal{L}$ is dense-in-itself.

Proof Suppose there was an isolated point w in $S^\mathcal{L}$. Then there is a formula $\Box\varphi$ with $\widehat{\Box\varphi} = \{w\}$. This means $\Box\varphi \in w$ and for any ψ , $\psi \in w$ iff $\mathbf{S4} \vdash \Box\varphi \rightarrow \psi$, which is obviously a contradiction—since we are working in a language with infinitely many propositional letters. QED

3.2.10. COROLLARY. $\mathbf{S4}$ is the logic of the class of all topological spaces which are compact and dense-in-itself.

Still, the canonical topological space $S^\mathcal{L}$ is neither 0-dimensional nor metric (it is not even a T_0 -space). So, $S^\mathcal{L}$ is not homeomorphic to the Cantor space. In the next section, we show how to get completeness of $\mathbf{S4}$ with respect to the Cantor space by a different route.

3.2.3 Finite spaces suffice

We conclude with an observation that is important for later arguments. The whole construction in the completeness proof would also work if we restricted attention to the *finite* language consisting of the initial formula and all its subformulas. All definitions go through, and our arguments never needs to go beyond it. This means that we only get finitely many maximally consistent sets, and so non-provable formulas can be refuted on *finite models*, whose size is effectively computable from the formula itself. Note however that the obtained finite model will not necessarily be dense-in-itself.

3.2.11. COROLLARY. $\mathbf{S4}$ has the effective finite model property w.r.t. the class of topological spaces.

Incidentally, this also shows that validity in $\mathbf{S4}$ is *decidable*, but we forego such computability issues in this thesis.

The resulting models have some interesting topological extras. Consider any finite modal frame $\mathcal{F} = \langle W, R \rangle$. We define some auxiliary notions. For any $w \in W$, let $C(w) = \{v \in W \mid wRv \ \& \ vRw\}$. Call a set C a *cluster* if it is of the form $C(w)$ for some w : the *cluster generated by w* . $C(w)$ is *simple* if $C(w) = \{w\}$, and *proper* otherwise. $w \in W$ is called *minimal* if vRw implies wRv for any $v \in W$. A cluster C is *minimal* if there exists a minimal $w \in W$ such that $C = C(w)$. Next, call \mathcal{F} *rooted* if there is $w \in W$ such that wRv for any $v \in W$: w is then a *root* of \mathcal{F} . This w needs not be unique: any point from $C(w)$, the *initial cluster* of \mathcal{F} , will do.

Evidently, a finite Kripke frame \mathcal{F} is rooted iff it has only one minimal cluster. Topologically, this property is related to the earlier notion of *connectedness*. A topological space $\langle W, \tau \rangle$ is *connected* if its universe cannot be written as a union of two disjoint open sets. $\langle W, \tau \rangle$ is *well-connected* if $W = U \cup V$ implies $W = U$ or $W = V$, for any $U, V \in \tau$. Obviously well-connectedness is a stronger notion than connectedness. It corresponds to $\langle W, R_\tau \rangle$ being rooted. For this observe that, dually, well-connectedness can be stated as follows:

For any two closed subsets C, D of $\langle W, \tau \rangle$, $C \cap D = \emptyset$ implies $C = \emptyset$ or $D = \emptyset$.

3.2.12. LEMMA. *A finite Kripke frame is rooted if and only if the corresponding topological space is well-connected.*

Proof Suppose $\langle W, R \rangle$ is a rooted Kripke frame with a root w , and $\langle W, \tau_R \rangle$ the corresponding topological space. Let X_1 and X_2 be closed sets of $\langle W, \tau_R \rangle$ such that $X_1 \cap X_2 = \emptyset$. By an easy dualization of the notions of Section 3.1.1, a set $X \subseteq W$ is topologically closed iff it is *downward closed* in the ordering, that is $u \in X$ and vRu imply $v \in X$, for any $u, v \in W$. Now if both X_1 and X_2 are non-empty, then w belongs to both of them, which is a contradiction. Hence one of them should be empty, and $\langle W, \tau_R \rangle$ is well-connected.

Conversely, suppose $\langle W, R \rangle$ is not rooted. Then there are at least two different minimal clusters C_1 and C_2 in W . Since C_1 and C_2 are minimal clusters, they are downward closed, and hence closed in $\langle W, \tau_R \rangle$. Moreover, since they are different, $C_1 \cap C_2 = \emptyset$. Hence $\langle W, \tau_R \rangle$ is not well-connected. QED

This allows us to improve on Corollary 3.2.11.

3.2.13. THEOREM. **S4** *is the logic of finite well-connected topological spaces.*

Proof It suffices to observe the following. If a modal formula has a counter-example on a finite Kripke model, it fails in some point there. But then by standard ‘Locality’, it also fails in the submodel *generated* by that point and its relational successors, which is rooted—and hence transforms into a well-connected topological space. QED

Again, there is a downside to such an upgraded completeness result. What it also means is that the basic modal language cannot *define* such a nice topological property as well-connectedness. As we saw in Section 2.4, the definition of connectedness requires introduction of additional modalities. So does well-connectedness.

Finally, let us mention that for refuting non-theorems of **S4** it is enough to restrict ourselves to the class of those finite rooted models for which every cluster is proper. As we already mentioned in Section 3.1.1, having only simple clusters topologically corresponds to the T_0 separation axiom, which in finite case is equivalent to the T_D separation axiom (every point is obtained as intersection of an open and a closed sets). Consequently, having only proper clusters topologically corresponds to the fact that no point can be obtained as intersection of an open and a closed sets. Call spaces with this property *essentially non- T_D* . Then we can improve a little bit on Theorem 3.2.13:

3.2.14. THEOREM. **S4** is the logic of finite well-connected essentially non- T_D topological spaces.

Proof Suppose a modal formula φ has a counter-example on a finite rooted Kripke model $M = \langle W, R, \models \rangle$. Then replacing every cluster of W by an n -element cluster, where n is the maximum among the sizes of the clusters of W , we obtain a new frame $\langle W', R' \rangle$. Obviously $\langle W, R \rangle$ is a p -morphic image of $\langle W', R' \rangle$. This allows us to define \models' on $\langle W', R' \rangle$ so that φ has also a counter-example on $M' = \langle W', R', \models' \rangle$. Now every cluster of W' is proper, hence $\langle W', R' \rangle$ transforms into a well-connected essentially non- T_D topological space. QED

3.3 Completeness on the reals

As early as 1944, McKinsey and Tarski proved the following beautiful result, which is an expansion of a completeness theorem by Tarski for intuitionistic propositional logic from 1938:

3.3.1. THEOREM (MCKINSEY AND TARSKI). **S4** is the complete logic of any metric separable dense-in-itself space.

Most importantly, this theorem implies completeness of **S4** with respect to the real line \mathbb{R} . It also implies completeness of **S4** with respect to the Cantor space \mathcal{C} .

Our presentation does not present any startling new results improving on this theorem. It rather takes a systematic look at its proof, and what it achieves. The original algebraic proof in [McKinsey and Tarski, 1944] was very complex, the later more topological version in [Rasiowa and Sikorski, 1963] is not much more accessible. Recently, Mints [Mints, 1998] replaced these by a much more perspicuous model-theoretic construction, extending earlier ideas of Beth and Kripke to get faster completeness of **S4**

with respect to the Cantor space. We generalize its model-theoretic structure, using topo-bisimulations, and also provide a modification for completeness on the reals.

Our strategy in the following subsections starts from the standard modal completeness for **S4** involving counter-examples on finite rooted models, and then exhibits a topo-bisimulation resulting in “tree-like” topological model homeomorphic to the Cantor space \mathcal{C} . We then show how to extract completeness of **S4** with respect to the reals from the completeness of **S4** with respect to \mathcal{C} .

3.3.1 Cantorization

Our starting point is an arbitrary modal formula which is not provable in **S4**. We have already seen that such a non-theorem can be refuted on a finite rooted Kripke model. Now we show how to transform the latter into a counterexample on the Cantor space \mathcal{C} . Our technique is *selective unraveling*, a refinement of the *unraveling* technique [Blackburn et al., 2001].

Suppose $M = \langle W, R, \models \rangle$ is a finite rooted model with a root w . Our goal is to select those infinite paths of M which are in a one-to-one correspondence with infinite paths of the full infinite binary tree T_2 . In order to give an easier description of our construction, we assume that every cluster of W is proper. This can be done by Theorem 3.2.14. Now start with a root w , and announce (w) as a selective path. Then if (w_1, \dots, w_k) is already a selective path, introduce a *left* move by announcing (w_1, \dots, w_k, w_k) as a selective path; and introduce a *right* move by announcing $(w_1, \dots, w_k, w_{k+1})$ as a selective path if $w_k R w_{k+1}$ and $w_k \neq w_{k+1}$. (Since we assumed that every cluster of W is proper, such w_{k+1} will exist for every w_k .)

To make this idea precise, we need some definitions. For $u, v \in W$, call v a *strong successor* of u if uRv and $u \neq v$. Write $SSuc(u)$ for the set of all strong successors of u . Since we assumed that every cluster of W is proper, $SSuc(u) \neq \emptyset$ for every $u \in W$. Suppose v_1, \dots, v_n is a complete enumeration of $SSuc(u)$ for every $u \in W$. Now define a *selective path* of W recursively:

- 1 (w) is a selective path;
- 2 If (w_1, \dots, w_k) is a selective path of length k , then $(w_1, \dots, w_k, w_{k+1})$ is a selective path of length $k + 1$, where $w_{k+1} = w_k$;
- 3 If (w_1, \dots, w_k) is a selective path of length k , then $(w_1, \dots, w_k, w_{k+1})$ is a selective path of length $k + 1$, where $w_{k+1} = v_i$ with $i \equiv k(\text{mod } n)$;¹
- 4 That’s all!

¹In other words, w_{k+1} is the first strong successor of w_k in the complete enumeration of $SSuc(w_k)$ which has not appeared in any selective path of length k ; if all strong successors of w_k have already appeared in one of selective paths of length k , then we start over again and put w_{k+1} to be the first strong successor of w_k in the complete enumeration of $SSuc(w_k)$.

Use Σ to denote the set of all infinite selective paths of W . For a finite selective path (w_1, \dots, w_k) , let

$$B_{(w_1, \dots, w_k)} = \{\sigma \in \Sigma \mid \sigma \text{ has an initial segment } (w_1, \dots, w_k)\}.$$

Define a topology τ_Σ on Σ by introducing

$$\mathcal{B}_\Sigma = \{B_{(w_1, \dots, w_k)} \mid (w_1, \dots, w_k) \text{ is a finite selective path of } W\}$$

as a basis. To see that \mathcal{B}_Σ is a basis, observe that $B_{(w)} = \Sigma$, and that

$$B_{(w_1, \dots, w_k)} \cap B_{(v_1, \dots, v_m)} = \begin{cases} B_{(w_1, \dots, w_k)} & \text{if } (v_1, \dots, v_m) \text{ is an initial segment} \\ & \text{of } (w_1, \dots, w_k), \\ B_{(v_1, \dots, v_m)} & \text{if } (w_1, \dots, w_k) \text{ is an initial segment} \\ & \text{of } (v_1, \dots, v_m), \\ \emptyset & \text{otherwise.} \end{cases}$$

In order to define \models_Σ note that every infinite selective path σ of W either gets stable or keeps cycling. In other words, either $\sigma = (w_1, \dots, w_k, w_k, \dots)$, or $\sigma = (w_1, \dots, w_n, w_{n+1}, \dots)$, where w_i belongs to some cluster $C \subseteq W$ for $i > n$. In the former case we say that w_k *stabilizes* σ , and in the latter that σ *keeps cycling* in C . Now define \models_Σ on Σ by putting

$$\sigma \models_\Sigma P \text{ iff } \begin{cases} w_k \models P & \text{if } w_k \text{ stabilizes } \sigma, \\ \rho(C) \models P & \text{if } \sigma \text{ keeps cycling in } C \subseteq W, \text{ where } \rho(C) \text{ is some} \\ & \text{arbitrarily chosen representative of } C. \end{cases}$$

All we need to show is that $\langle \Sigma, \tau_\Sigma \rangle$ is homeomorphic to the Cantor space, and that $M_\Sigma = \langle \Sigma, \tau_\Sigma, \models_\Sigma \rangle$ is topo-bisimilar to the initial M . In order to show the first claim, let us recall that the Cantor space is homeomorphic to the countable topological product of the two element set $\mathbf{2} = \{0, 1\}$ with the discrete topology. So, $\mathcal{C} \cong \mathbf{2}^\omega$ with the subbasic sets for the topology being $U = \prod_{i \in \omega} U_i$, where all but one U_i coincide with $\mathbf{2}$, or equivalently with the basic sets for the topology being $U = \prod_{i \in \omega} U_i$, where all but finitely many U_i coincide with $\mathbf{2}$.

To picture the Cantor space, one can think of the full infinite binary tree T_2 : starting at the root, one associates 0 to every left-son of a node, and 1 with every right-son. Then the points of the Cantor space are the infinite branches of T_2 .

3.3.2. PROPOSITION. $\langle \Sigma, \tau_\Sigma \rangle$ is homeomorphic to \mathcal{C} .

Proof Suppose $\sigma = (w_1, w_2, w_3, \dots, w_k, \dots) \in \Sigma$, where $w_1 = w$ is a root of W . With each w_k ($k > 1$) associate 0 if $w_{k-1} = w_k$, and associate 1 if w_k is a strong

successor of w_{k-1} . Denote an element of $\mathbf{2}$ associated with w_k by $g(w_k)$ and define $G : \Sigma \rightarrow \mathbf{2}^\omega$ by putting

$$G(w_1, w_2, w_3, \dots, w_k, \dots) = (g(w_2), g(w_3), \dots, g(w_k), \dots).$$

It should be clear from the definition that G is a bijection. In order to prove that it is a homeomorphism, we need to check that G is open. So, suppose $B_{(w_1, \dots, w_k)}$ is a basic open set of τ_Σ . Then

$$G(B_{(w_1, \dots, w_k)}) = \{g(w_2)\} \times \dots \times \{g(w_k)\} \times \mathbf{2}^\omega$$

is a basic open of \mathcal{C} , G preserves basic opens, hence preserves opens. Conversely, suppose $U = \mathbf{2}^{k-1} \times \{c_k\} \times \mathbf{2}^\omega$, where $c_k = 0$ or 1 , is a subbasic open of \mathcal{C} . Then

$$G^{-1}(U) = \bigcup_{g(w_k)=c_k} B_{(w_1, \dots, w_k)},$$

which obviously belongs to τ_Σ . Thus, G is open, hence a homeomorphism. QED

It is left to be shown that M_Σ is topo-bisimilar to M . Define $F : \Sigma \rightarrow W$ by putting

$$F(\sigma) = \begin{cases} w_k & \text{if } w_k \text{ stabilizes } \sigma, \\ \rho(C) & \text{if } \sigma \text{ keeps cycling in } C. \end{cases}$$

F is well-defined, and surjective. (For any $w_k \in W$, $F(\sigma_0, w_k, w_k, \dots) = w_k$, where σ_0 is a (finite) selective path from w_1 to w_k .)

3.3.3. PROPOSITION. *F is a total topo-bisimulation between $M_\Sigma = \langle \Sigma, \tau_\Sigma, \models_\Sigma \rangle$ and $M = \langle W, R, \models \rangle$.*

Proof Recall from the previous section that a finite topological space $\langle W, \tau_R \rangle$ is associated with $\langle W, R \rangle$ (since $\langle W, R \rangle$ is rooted, $\langle W, \tau_R \rangle$ is actually well-connected). Let us check that $F : \langle \Sigma, \tau_\Sigma \rangle \rightarrow \langle W, \tau_R \rangle$ is open. Recall that $R(v)$, for $v \in W$, are basic opens of τ_R . So, in order to check that F is continuous, we need to show that the F inverse image of every $R(v)$ is open in τ_Σ . Observe that for any $v \in W$,

$$F^{-1}(R(v)) = \bigcup_{k \in \omega, vRw_k} B_{(w_1, \dots, w_k)},$$

which is an element of τ_Σ . Indeed, suppose $\sigma \in \bigcup_{k \in \omega, vRw_k} B_{(w_1, \dots, w_k)}$. Then σ belongs to one of $B_{(w_1, \dots, w_k)}$ with vRw_k . But then $w_k RF(\sigma)$, which together with vRw_k and transitivity of R imply that $vRF(\sigma)$. So, $F(\sigma) \in R(v)$, and $\sigma \in F^{-1}(R(v))$. Conversely, suppose $\sigma \in F^{-1}(R(v))$. Then $F(\sigma) \in R(v)$, and $vRF(\sigma)$. Now either w_k stabilizes σ , or σ keeps cycling in a cluster C . In the former case, $\sigma = (w_1, \dots, w_k, w_k, \dots)$, where $w_k = F(\sigma)$. Hence, $\sigma \in B_{(w_1, \dots, w_k)}$ with vRw_k . In the latter case, $\sigma = (w_1, \dots, w_n, w_{n+1}, \dots)$, where $w_i \in C$ for $i > n$, and $F(\sigma) =$

$\rho(C)$. Hence, $\sigma \in B_{(w_1, \dots, w_n, w_{n+1})}$ with vRw_{n+1} . In either case, $F^{-1}(R(v)) \subseteq \bigcup_{k \in \omega, vRw_k} B_{(w_1, \dots, w_k)}$. Thus, $F^{-1}(R(v)) = \bigcup_{k \in \omega, vRw_k} B_{(w_1, \dots, w_k)}$, and F is continuous.

In order to show that F preserves opens, consider any basic set $B_{(w_1, \dots, w_k)}$ of τ_Σ and show that $F(B_{(w_1, \dots, w_k)})$ is open in τ_R . For this we show that

$$F(B_{(w_1, \dots, w_k)}) = R(w_k).$$

Suppose $v \in F(B_{(w_1, \dots, w_k)})$. Then there exists $\sigma = (w_1, \dots, w_k, \dots) \in B_{(w_1, \dots, w_k)}$ such that $F(\sigma) = v$. Hence, we have that w_kRv . Conversely, suppose w_kRv . Consider a (finite) selective path σ_0 from w_1 to v containing (w_1, \dots, w_k) as an initial segment. Then $\sigma = (\sigma_0, v, v, v, \dots) \in B_{(w_1, \dots, w_k)}$ and $F(\sigma) = v$. Hence $F(B_{(w_1, \dots, w_k)}) = R(w_k)$, which is a basic open of τ_R . So, F is open.

Moreover, as follows from the definition of \models_Σ ,

$$\sigma \models_\Sigma P \text{ iff } F(\sigma) \models P.$$

Now, since every continuous and open map satisfying this condition is a topo-bisimulation (cf. Theorem 2.1.5), so is our F . QED

3.3.4. THEOREM. **S4** is complete with respect to the Cantor space.

Proof Suppose **S4** $\not\models \varphi$. Then by Theorem 3.2.13 there is a finite rooted Kripke model M refuting φ . By Theorem 3.2.14 we can assume that every cluster of M is proper. By Propositions 3.3.2 and 3.3.3 there exists a valuation $\models_{\mathcal{C}}$ on the Cantor set \mathcal{C} such that $\langle \mathcal{C}, \models_{\mathcal{C}} \rangle$ is topo-bisimilar to M . Hence, φ is refuted on \mathcal{C} . QED

3.3.2 Counterexamples on the reals

In the previous subsection, we described how selective unraveling transforms counterexamples on a finite rooted Kripke model M into counterexamples on the Cantor space \mathcal{C} . In this subsection we show how to transfer counterexamples from M to $(0, 1)$. As a result, we obtain a new proof of completeness of **S4** with respect to the real line.

Our strategy is similar to that in Section 3.3.1: we start with a non-theorem of **S4** having a counterexample on a finite rooted Kripke model $M = \langle W, R, \models \rangle$ whose every cluster is proper. Then we construct the set Σ of all selective paths of W , and subtract a proper subset Λ of Σ , which is in a one-to-one correspondence with $(0, 1)$. After that we define a topology τ_Λ on Λ so that $\langle \Lambda, \tau_\Lambda \rangle$ is homeomorphic to $(0, 1)$ with its natural topology. Finally, we define a valuation \models_Λ on Λ , and show that $\langle \Lambda, \tau_\Lambda, \models_\Lambda \rangle$ is topo-bisimilar to M . Note that since τ_Λ is pretty different from τ_Σ , the topo-bisimulation between $\langle \Lambda, \tau_\Lambda, \models_\Lambda \rangle$ and M is not simply the restriction of the topo-bisimulation between $\langle \Sigma, \tau_\Sigma, \models_\Sigma \rangle$ and M constructed in Section 3.3.1, but rather its appropriate modification.

Recall from Section 3.3.1 that in selective unraveling we had three different types of selective branches: going infinitely to the left, infinitely to the right, or infinitely zigzagging. Also recall that a selective branch σ is going infinitely to the left if $\sigma = (w_1, \dots, w_k, w_k, \dots)$; σ is going infinitely to the right if $\sigma = (w_1, \dots, w_n, w_{n+1}, \dots)$, where w_{k+1} is a strong successor of w_k for any $k \geq n$; and finally, σ is zigzagging if $\sigma = (w_1, \dots, w_n, w_{n+1}, \dots)$, where there are infinitely many $k \geq n$ with $w_{k+1} = w_k$, and there are also infinitely many $k \geq n$ with w_{k+1} being a strong successor of w_k .

In order to transfer counterexamples from M to $(0, 1)$, in the definition of selective unraveling we need to restrict ourselves only to those branches which are either going infinitely to the left or are infinitely zigzagging. That is, we define a *real path* of W to be a selective path of W either going infinitely to the left or infinitely zigzagging.

Denote by Π the set of all real infinite paths of W . So, Π is the subset of the set Σ of all selective infinite paths of W consisting of all selective paths going infinitely to the left or infinitely zigzagging. Therefore, Π is in a one-to-one correspondence with the set of those infinite branches of the infinite binary tree T_2 which either have 0 from some node on or are infinitely zigzagging.

This correspondence sets up the desired connection between Π and $(0, 1)$. To see this recall the dyadic representation of a number from $[0, 1]$. Let $x \in [0, 1]$. To construct an infinite branch $\alpha = (a_n)_{n \in \omega}$ of T_2 representing x observe that either $x \in [0, \frac{1}{2}]$ or $x \in [\frac{1}{2}, 1]$. In the former case put $a_1 = 0$ and in the latter case put $a_1 = 1$. Assume $x \in [0, \frac{1}{2}]$. Then either $x \in [0, \frac{1}{4}]$ or $x \in [\frac{1}{4}, \frac{1}{2}]$. Again in the former case put $a_2 = 0$ and in the latter case put $a_2 = 1$. Continuing this process, we get an infinite branch $\alpha = (a_n)_{n \in \omega}$ of T_2 which in turn represents x .

Note that there are two ways for the dyadic representation of $\frac{1}{2}$: $(0, 1, 1, 1, \dots)$ or $(1, 0, 0, 0, \dots)$. In general, there are two ways for the dyadic representation of any number $\frac{m}{2^n} \in [0, 1]$ ($m, n \in \omega$, $0 < m < 2^n$): either as $(a_1, \dots, a_k, 1, 0, 0, 0, \dots)$ or as $(a_1, \dots, a_k, 0, 1, 1, 1, \dots)$. Therefore, if we throw away all infinite branches of T_2 having 1 from some node on plus $(0, 0, 0, \dots)$, we obtain a one-to-one correspondence between $(0, 1)$ and the remaining infinite branches of T_2 . Hence, there exists a one-to-one correspondence between $(0, 1)$ and $\Lambda = \Pi - \{(w, w, w, \dots)\}$.

Suppose $(w_1, \dots, w_{k-1}, w_k, w_k, \dots) \in \Lambda$ ($w_{k-1} \neq w_k$) represents $\frac{m}{2^n} \in (0, 1)$. Also suppose

$$C_{(w_1, \dots, w_k)} = \{\lambda \in \Lambda \mid \text{the initial segment of } \lambda \text{ is } (w_1, \dots, w_k)\}.$$

(Observe that $C_{(w_1, \dots, w_k)} = B_{(w_1, \dots, w_k)} \cap \Lambda$.)

In order to transfer topological structure of $(0, 1)$ to Λ observe that the family $\{(\frac{m}{2^n}, \frac{m+1}{2^n}) \mid m, n \in \omega, 0 < m+1 < 2^n\}$ forms a basis for the topology on $(0, 1)$, and that the subset of Λ representing $(\frac{m}{2^n}, \frac{m+1}{2^n})$ is $D_{(w_1, \dots, w_k)} = C_{(w_1, \dots, w_k)} - \{(w_1, \dots, w_{k-1}, w_k, w_k, \dots)\}$. Hence, if we define a topology τ_Λ on Λ by introducing

$$\{D_{(w_1, \dots, w_k)} \mid (w_1, \dots, w_k) \text{ is a finite selective path of } \Lambda\}$$

as a basis, the following obvious fact holds:

3.3.5. FACT. (Λ, τ_Λ) is homeomorphic to $(0, 1)$.

Now we define \models_Λ on Λ , and show that there exists a topo-bisimulation between $(\Lambda, \tau_\Lambda, \models_\Lambda)$ and M .

In order to define \models_Λ observe that either $\lambda \in \Lambda$ gets stable or it keeps cycling. In other words, either $\lambda = (w_1, \dots, w_{k-1}, w_k, w_k, \dots)$, or $\lambda = (w_1, \dots, w_n, w_{n+1}, \dots)$, where w_i belongs to some cluster $C \subseteq W$, for $i > n$. In the former case we say that w_k stabilizes λ , and in the latter—that λ keeps cycling in C . Now define \models_Λ on Λ by putting

$$\lambda \models_\Lambda P \text{ iff } \begin{cases} w_{k-1} \models P & \text{if } w_k \text{ stabilizes } \lambda, \\ \rho(C) \models P & \text{if } \lambda \text{ keeps cycling in } C \subseteq W, \text{ where } \rho(C) \text{ is} \\ & \text{some arbitrarily chosen representative of } C. \end{cases}$$

Finally, define a function $F : \Lambda \rightarrow W$ by putting

$$F(\lambda) = \begin{cases} w_{k-1} & \text{if } w_k \text{ stabilizes } \lambda, \\ \rho(C) & \text{if } \lambda \text{ keeps cycling in } C. \end{cases}$$

3.3.6. PROPOSITION. F is a total topo-bisimulation between $M_\Lambda = \langle \Lambda, \tau_\Lambda, \models_\Lambda \rangle$ and $M = \langle W, R, \models \rangle$.

Proof Obviously F is well-defined, and is actually surjective. (For any $w_k \in W$, $F(w_1, \dots, w_k, w_{k+1}, w_{k+1}, \dots) = w_k$, where (w_1, \dots, w_k) is a finite selective path from w_1 to w_k , and w_{k+1} is a strong successor w_k . Note that w_{k+1} exists, since every cluster of W is proper.) Let us check that $F : \langle \Lambda, \tau_\Lambda \rangle \rightarrow \langle W, \tau_R \rangle$ is open. Recall that $R(v)$, for $v \in W$, are basic opens of τ_R . So, in order to check that F is continuous, we need to show that the F inverse image of every $R(v)$ is open in τ_Λ . Observe that for any $v \in W$,

$$F^{-1}(R(v)) = \bigcup_{k \in \omega, vRw_k} D_{(w_1, \dots, w_k)},$$

which is an element of τ_Λ . Indeed, suppose $\lambda \in \bigcup_{k \in \omega, vRw_k} D_{(w_1, \dots, w_k)}$. Then λ belongs to one of $D_{(w_1, \dots, w_k)}$ with vRw_k . Now $\lambda \in D_{(w_1, \dots, w_k)}$ implies $w_k RF(\lambda)$, which together with vRw_k and transitivity of R yield $vRF(\lambda)$. Hence, $F(\lambda) \in R(v)$, and $\lambda \in F^{-1}(R(v))$. Conversely, suppose $\lambda \in F^{-1}(R(v))$. Then $F(\lambda) \in R(v)$, and $vRF(\lambda)$. Now either λ is going infinitely to the left or is infinitely zigzagging. In the former case, $\lambda = (w_1, \dots, w_k, w_{k+1}, w_{k+1}, \dots)$, where $w_k = F(\lambda)$. Hence, $\lambda \in D_{(w_1, \dots, w_k)}$ with vRw_k . In the latter case, $\lambda = (w_1, \dots, w_n, w_{n+1}, w_{n+2}, \dots)$, where $F(\lambda) \in C(w_{n+1})$. Hence, $\lambda \in D_{(w_1, \dots, w_n, w_{n+1})}$ with vRw_{n+1} . In either case, $\lambda \in \bigcup_{k \in \omega, vRw_k} D_{(w_1, \dots, w_k)}$, and $F^{-1}(R(v)) = \bigcup_{k \in \omega, vRw_k} D_{(w_1, \dots, w_k)}$. Hence, F is continuous.

In order to show that F preserves opens, consider any basic set $D_{(w_1, \dots, w_k)}$ of τ_Λ and show that $F(D_{(w_1, \dots, w_k)})$ is open in τ_R . For this we show that

$$F(D_{(w_1, \dots, w_k)}) = R(w_k).$$

Suppose $v \in F(D_{(w_1, \dots, w_k)})$. Then there exists $\lambda = (w_1, \dots, w_k, \dots) \in D_{(w_1, \dots, w_k)}$ such that $F(\lambda) = v$. Now either λ is going infinitely to the left or is infinitely zigzagging. In the former case, $\lambda = (w_1, \dots, w_k, \dots, w_{k+l}, w_{k+l+1}, w_{k+l+1}, \dots)$, where $w_{k+l} = v$. In the latter case, v is a representative of a cluster C where λ keeps cycling. In either case, $w_k R v$. Hence, $v \in R(w_k)$. Conversely, suppose $v \in R(w_k)$. Then $w_k R v$. Consider $\lambda = (w_1, \dots, w_k, \dots, v, u, u, \dots)$, where $(w_1, \dots, w_k, \dots, v)$ is a finite selective path of W from w_1 to v containing (w_1, \dots, w_k) as an initial segment, and u is a strong successor of v . (u exists, since every cluster of W is proper.) Then $\lambda \in D_{(w_1, \dots, w_k)}$ and $F(\lambda) = v$. Hence $F(D_{(w_1, \dots, w_k)}) = R(w_k)$, which is a basic open of τ_R . So, F is open.

Moreover, as follows from the definition of \models_Λ ,

$$\lambda \models_\Lambda P \text{ iff } F(\lambda) \models P.$$

Now since every continuous and open map satisfying this condition is a topobisimulation (cf. Theorem 2.1.5), so is our F . QED

3.3.7. COROLLARY. **S4** is complete with respect to $(0, 1)$.

Proof Suppose **S4** \dashv φ . Then by Theorem 3.2.13 there is a finite rooted Kripke model M refuting φ . By Theorem 3.2.14 we can assume that every cluster of M is proper. By Proposition 3.3.6, M is topo-bisimilar to $M_\Lambda = \langle \Lambda, \tau_\Lambda, \models_\Lambda \rangle$. Hence, M_Λ is refuting φ . Now since $\langle \Lambda, \tau_\Lambda \rangle$ is homeomorphic to $(0, 1)$, φ is refuted on $(0, 1)$. QED

3.3.8. THEOREM. **S4** is complete with respect to the real line \mathbb{R} .

Proof Suppose **S4** \dashv φ . Then by Corollary 3.3.7 there exists a valuation $\models_{(0,1)}$ on $(0, 1)$ refuting φ . Now since $(0, 1)$ is homeomorphic to \mathbb{R} , φ is refuted on \mathbb{R} . QED

This provides an alternative proof of McKinsey and Tarski's original proof. It should be noted that we can improve a little bit on their result. Indeed, McKinsey and Tarski proved that for any non-theorem φ of **S4** there exists a valuation ν on \mathbb{R} falsifying φ .

3.3.9. COROLLARY. There exists a single valuation ν on \mathbb{R} falsifying all the non-theorems of **S4**.

Proof Enumerate all the non-theorems of **S4**. This can be done since the language of **S4** is countable. Let this enumeration be $\{\varphi_1, \varphi_2, \dots\}$. Since the interval $(n, n+1)$ is homeomorphic to \mathbb{R} , from Theorem 3.3.8 it follows that there exists a valuation ν_n on $(n, n+1)$ such that $\langle (n, n+1), \nu_n \rangle$ falsifies φ_n . (Note that we need not know anything about the shape of $\nu_n(\varphi_n)$.) Now take $\bigcup_{n \in \omega} (n, n+1)$. For any propositional letter P let $\nu(P) = \bigcup_{n \in \omega} \nu_n(P)$ be the valuation of P on \mathbb{R} . Note that each $\langle (n, n+1), \nu_n \rangle$ is an open submodel of $\langle \mathbb{R}, \nu \rangle$, where the ‘identity embedding’ is a topo-bisimulation. Hence, the truth values of modal formulas do not change moving from each $\langle (n, n+1), \nu_n \rangle$ to $\langle \mathbb{R}, \nu \rangle$. Therefore, φ_n is still falsified on the whole \mathbb{R} for each n . Thus, we have constructed a single valuation ν on \mathbb{R} falsifying all the non-theorems of **S4**. QED

This also shows that though very different from the standard canonical Kripke model of **S4**, \mathbb{R} shares some of its universal properties.

3.3.3 Logical non-finiteness on the reals

Recall that two formulas φ and ψ are said to be **S4**-equivalent if $\mathbf{S4} \vdash \varphi \leftrightarrow \psi$. It is well known that there exist infinitely many formulas of one-variable which are not **S4**-equivalent. E.g., consider the following list of formulas:

$$\varphi_0 = P;$$

$$\varphi_n = \varphi_{n-1} \wedge \diamond(\diamond\varphi_{n-1} \wedge \neg\varphi_{n-1}).$$

We can easily construct a Kripke model on which all φ_n have different interpretations. Let $M = \langle \omega, R, \models \rangle$, where ω denotes the set of all natural numbers, nRm iff $m \leq n$, and $n \models P$ iff n is odd. Then one can readily check that φ_n is true at all odd points $> n$. Hence every φ_n has a different interpretation on M . It implies that the φ_n are not **S4**-equivalent. Now we give a topological flavor to this result by showing that interpreting a propositional variable as a certain subset of \mathbb{R} allows us to construct infinitely many **S4**-non-equivalent formulas of one variable. Corollary 3.3.9 already told us such a uniform choice must exist, but the proof does not construct $\nu(P)$ explicitly. The following argument does, and thereby also highlights the topological content of our modal completeness theorem.

We use \diamond and \square instead of the standard notations $\overline{(\)}$ and $Int(\)$ for the closure and interior operators of topology. This modal notation shows its basic use in topology because it allows us to write topological formulas in a more perspicuous fashion.

To proceed further we need to recall the definition of Hausdorff’s residue of a given set. Suppose a topological space $\langle W, \tau \rangle$ and $X \subseteq W$ are given. $\varrho(X) = X \cap \diamond(\diamond X - X)$ is called the *Hausdorff residue* of X . Let $\varrho^0(X) = X$, $\varrho^1(X) = \varrho(X)$ and $\varrho^{n+1}(X) = \varrho\varrho^n(X)$.

X is said to be of *rank* n , written $r(X) = n$, if n is the least natural number such that $\varrho^n(X) = \emptyset$. X is said to be of *finite rank* if there exists a natural n such that X is of rank n . X is said to be of *infinite rank* if it is not of finite rank.

The point $x \in X$ is said to be of *rank* n if $x \in \varrho^n(X)$, but $x \notin \varrho^{n+1}(X)$. $x \in X$ is said to be of *finite rank* if there exists a natural n such that x is of rank n . x is said to be of *infinite rank* if it is not of finite rank.

Obviously X is of rank n iff the rank of every element of X is strictly less than n , and there is at least one element of X of rank $n - 1$; X is of finite rank iff there is a natural n such that the rank of every element of X is strictly less than n ; and X is of infinite rank iff there is no finite bound on the ranks of elements of X .

If we interpret P as a subset X of \mathbb{R} , then φ_n will be interpreted as $\varrho^n(X)$. So, in order to show that different φ_n are **S4**-non-equivalent, it is sufficient to show that there is $X \subset \mathbb{R}$ such that $\varrho(X) \supset \varrho^2(X) \supset \cdots \supset \varrho^n(X) \supset \cdots$. We have the following

3.3.10. PROPOSITION. *There exists a subset X of \mathbb{R} such that $\varrho(X) \supset \varrho^2(X) \supset \cdots \supset \varrho^n(X) \supset \cdots$.*

Proof We construct X inductively. Fix a natural number k .

Step 1: Consider a sequence $\{x_{i_1}\}_{i_1=1}^\infty$ from $(k-1, k)$ converging to $k-1$, and put

$$X_1 = \{k-1\} \cup \bigcup_{i_1=1}^\infty \{y_{i_2}^{i_1}\}_{i_2=1}^\infty,$$

where $\{y_{i_2}^{i_1}\}_{i_2=1}^\infty$ is a sequence from (x_{i_1+1}, x_{i_1}) converging to x_{i_1+1} . Note that

$$\begin{aligned} \diamond X_1 &= X_1 \cup \{x_{i_1}\}_{i_1=1}^\infty, \\ \diamond X_1 - X_1 &= \{x_{i_1}\}_{i_1=1}^\infty, \\ \diamond(\diamond X_1 - X_1) &= \{k-1\} \cup \{x_{i_1}\}_{i_1=1}^\infty, \text{ and} \\ \varrho(X_1) &= \{k-1\}. \end{aligned}$$

So, $k-1$ is the only point of X_1 of rank 1, and $r(X_1) = 2$.

Step 2: Consider a sequence $\{x_{i_3}^{i_1, i_2}\}_{i_3=1}^\infty$ from $(y_{i_2+1}^{i_1}, y_{i_2}^{i_1})$ converging to $y_{i_2+1}^{i_1}$, and put

$$X_2 = \{k-1\} \cup \bigcup_{i_1=1}^\infty \{y_{i_2}^{i_1}\}_{i_2=1}^\infty \cup \bigcup_{i_1=1}^\infty \bigcup_{i_2=1}^\infty \bigcup_{i_3=1}^\infty \{y_{i_4}^{i_1, i_2, i_3}\}_{i_4=1}^\infty,$$

where $\{y_{i_4}^{i_1, i_2, i_3}\}_{i_4=1}^\infty$ is a sequence from $(x_{i_3+1}^{i_1, i_2}, x_{i_3}^{i_1, i_2})$ converging to $x_{i_3+1}^{i_1, i_2}$. Note that $X_2 \supset X_1$, and

$$\begin{aligned} \diamond X_2 &= X_2 \cup \{x_{i_1}\}_{i_1=1}^\infty \cup \bigcup_{i_1=1}^\infty \bigcup_{i_2=1}^\infty \{x_{i_3}^{i_1, i_2}\}_{i_3=1}^\infty, \\ \diamond X_2 - X_2 &= \{x_{i_1}\}_{i_1=1}^\infty \cup \bigcup_{i_1=1}^\infty \bigcup_{i_2=1}^\infty \{x_{i_3}^{i_1, i_2}\}_{i_3=1}^\infty, \\ \diamond(\diamond X_2 - X_2) &= \{k-1\} \cup \bigcup_{i_1=1}^\infty \{y_{i_2}^{i_1}\}_{i_2=1}^\infty \cup \{x_{i_1}\}_{i_1=1}^\infty \cup \bigcup_{i_1=1}^\infty \bigcup_{i_2=1}^\infty \{x_{i_3}^{i_1, i_2}\}_{i_3=1}^\infty, \end{aligned}$$

$$\varrho(X_2) = \{k-1\} \cup \bigcup_{i_1=1}^{\infty} \{y_{i_2}^{i_1}\}_{i_2=1}^{\infty}, \text{ and}$$

$$\varrho^2(X_2) = \{k-1\}.$$

So, the points of X_2 of rank 1 are $y_{i_2}^{i_1}$, for arbitrary i_1 and i_2 , $k-1$ is the only point of X_2 of rank 2, and $r(X_2) = 3$.

Step n: For $n \geq 1$ consider a sequence $\{x_{i_{2n-1}}^{i_1, \dots, i_{2n-2}}\}_{i_{2n-1}=1}^{\infty}$ from $(y_{i_{2n-2}+1}^{i_1, \dots, i_{2n-3}}, y_{i_{2n-2}}^{i_1, \dots, i_{2n-3}})$ converging to $y_{i_{2n-2}+1}^{i_1, \dots, i_{2n-3}}$, and put

$$X_n = \{k-1\} \cup \bigcup_{i_1=1}^{\infty} \{y_{i_2}^{i_1}\}_{i_2=1}^{\infty} \cup \dots \cup \bigcup_{i_1=1}^{\infty} \dots \bigcup_{i_{2n-1}=1}^{\infty} \{y_{i_{2n}}^{i_1, \dots, i_{2n-1}}\}_{i_{2n}=1}^{\infty},$$

where $\{y_{i_{2n}}^{i_1, \dots, i_{2n-1}}\}_{i_{2n}=1}^{\infty}$ is a sequence from $(x_{i_{2n-1}+1}^{i_1, \dots, i_{2n-2}}, x_{i_{2n-1}}^{i_1, \dots, i_{2n-2}})$ to $x_{i_{2n-1}+1}^{i_1, \dots, i_{2n-2}}$. Also let

$$A = \{x_{i_1}\}_{i_1=1}^{\infty} \cup \dots \cup \bigcup_{i_1=1}^{\infty} \dots \bigcup_{i_{2n-2}=1}^{\infty} \{x_{i_{2n-1}}^{i_1, \dots, i_{2n-2}}\}_{i_{2n-1}=1}^{\infty}.$$

Then note that $X_n \supset X_{n-1} \supset \dots \supset X_2 \supset X_1$, and

$$\begin{aligned} \diamond X_n &= X_n \cup A, \\ \diamond X_n - X_n &= A, \\ \diamond (\diamond X_n - X_n) &= A \cup (X_n - [\bigcup_{i_1=1}^{\infty} \dots \bigcup_{i_{2n-1}=1}^{\infty} \{y_{i_{2n}}^{i_1, \dots, i_{2n-1}}\}_{i_{2n}=1}^{\infty}]), \\ \varrho(X_n) &= X_n - [\bigcup_{i_1=1}^{\infty} \dots \bigcup_{i_{2n-1}=1}^{\infty} \{y_{i_{2n}}^{i_1, \dots, i_{2n-1}}\}_{i_{2n}=1}^{\infty}], \\ \varrho^2(X_n) &= \rho(X_n) - [\bigcup_{i_1=1}^{\infty} \dots \bigcup_{i_{2n-3}=1}^{\infty} \{y_{i_{2n-2}}^{i_1, \dots, i_{2n-3}}\}_{i_{2n-2}=1}^{\infty}], \\ &\dots \\ \varrho^n(X_n) &= \{k-1\}. \end{aligned}$$

So, the points of X_n of rank 1 are

$$X_n - [\bigcup_{i_1=1}^{\infty} \dots \bigcup_{i_{2n-1}=1}^{\infty} \{y_{i_{2n}}^{i_1, \dots, i_{2n-1}}\}_{i_{2n}=1}^{\infty}],$$

the points of X_n of rank 2 are

$$X_n - [\bigcup_{i_1=1}^{\infty} \dots \bigcup_{i_{2n-3}=1}^{\infty} \{y_{i_{2n-2}}^{i_1, \dots, i_{2n-3}}\}_{i_{2n-2}=1}^{\infty} \cup \bigcup_{i_1=1}^{\infty} \dots \bigcup_{i_{2n-1}=1}^{\infty} \{y_{i_{2n}}^{i_1, \dots, i_{2n-1}}\}_{i_{2n}=1}^{\infty}],$$

and so on; finally, $k - 1$ is the only point of X_n of rank n , and $r(X_n) = n + 1$.

Now let X_1 be constructed in $(0, 1)$, X_2 in $(1, 2)$, X_n in $(n - 1, n)$, and so on. We put

$$X = \bigcup_{n=1}^{\infty} X_n.$$

Then $n - 1 \in \varrho^n(X)$ and $n - 1 \notin \varrho^{n+1}(X)$, for any natural n . So, $\varrho(X) \supset \varrho^2(X) \supset \cdots \supset \varrho^n(X) \supset \cdots$, and X contains points of every finite rank. QED

3.3.1. REMARK (INFINITE RANK). The X constructed above does not contain elements of infinite rank. However, a little adjustment of the above construction allow us to construct a subset of \mathbb{R} with an element of infinite rank. Actually, it is possible to construct a subset of \mathbb{R} containing elements of rank α , for any ordinal $\alpha < \aleph_1$.

Returning to our list of formulas, with P as the just constructed X , the interpretation of every φ_n in \mathbb{R} will be different, in terms of some topologically significant phenomenon. In the next section, we show that if we restrict ourselves to only “good” subsets of \mathbb{R} , then the situation drastically changes.

3.4 Axiomatizing special kinds of regions

By interpreting propositional variables as certain subsets of the real line \mathbb{R} , we can refute every non-theorem of **S4** on \mathbb{R} . Certainly not all subsets of \mathbb{R} are required for refuting the non-theorems of **S4**. In this section, we analyze the complexity of the subsets of \mathbb{R} required for refuting the non-theorems of **S4**. We prefer to use \diamond and \square to denote the closure and interior operators of a topological space. For consistency we also use \wedge, \vee and \neg to denote set-theoretical intersection, union and complement.

3.4.1 Serial sets on the real line

To start with, consider subsets of \mathbb{R} with the simplest intuitive structure. Call $X \subseteq \mathbb{R}$ *convex* if all points lying in between any two points of X belong to X . In other words, X is convex if $x, y \in X$ and $x \leq y$ imply $[x, y] \subseteq X$. Every convex subset of \mathbb{R} has one of the following forms:

$$\emptyset, (x, y), [x, y], [x, y), (x, y], (-\infty, x), (-\infty, x], (x, +\infty), [x, +\infty), \mathbb{R}.$$

3.4.1. DEFINITION. Call a subset of \mathbb{R} *serial* if it is a finite union of convex subsets of \mathbb{R} . Denote the set of all serial subsets of \mathbb{R} by $\mathcal{S}(\mathbb{R})$. So,

$$\mathcal{S}(\mathbb{R}) = \{X \in \mathcal{P}(\mathbb{R}) \mid X \text{ is a serial subset of } \mathbb{R}\}.$$

Obviously the X constructed in Proposition 3.3.10 is not serial, and actually this was absolutely crucial in showing that X has points of any finite rank. Indeed, we have the following

3.4.2. LEMMA. $r(X) = 0$ for any $X \in \mathcal{S}(\mathbb{R})$.

Proof First, $r(Y) = 0$ for any convex subset Y of \mathbb{R} . For, if Y is convex, then $\diamond Y \wedge \neg Y$ consists of at most two points, $\diamond(\diamond Y \wedge \neg Y) = \diamond Y \wedge \neg Y$, and $\varrho(Y) = Y \wedge \diamond(\diamond Y \wedge \neg Y) = Y \wedge (\diamond Y \wedge \neg Y) = \emptyset$. Hence $r(Y) = 0$.

Now let X be a serial subset of \mathbb{R} . Then $X = \bigvee_{i=1}^n X_i$, where every X_i is a convex subset of \mathbb{R} , and actually we can assume that all X_i are disjoint. But then $\varrho(X) = \bigvee_{i=1}^n \varrho(X_i) = \emptyset$, and hence $r(X) = 0$. QED

It follows that if we interpret P as a serial subset of \mathbb{R} , then no two φ_n ($n \geq 1$) from the previous section define sets equal to each other.

Call a valuation ν of our language \mathcal{L} to subsets of \mathbb{R} *serial* if $\nu(P) \in \mathcal{S}(\mathbb{R})$ for any propositional variable P . Since $\mathcal{S}(\mathbb{R})$ is closed with respect to \neg , \wedge and \diamond , we have that $\nu(\varphi) \in \mathcal{S}(\mathbb{R})$ for any serial valuation ν . Call a formula φ *S-true* if it is true in \mathbb{R} under a serial valuation. Call φ *S-valid* if φ is S-true for any serial valuation on \mathbb{R} . Let $L(\mathcal{S}) = \{\varphi \mid \varphi \text{ is } \mathcal{S}\text{-valid}\}$.

3.4.3. FACT. $L(\mathcal{S})$ is a normal modal logic over **S4**.

Obviously all φ_n ($n \geq 1$) from the previous section are $L(\mathcal{S})$ -equivalent. So, it is natural to expect that there are only finitely many formulas in one variable which are $L(\mathcal{S})$ -non-equivalent, and indeed that $L(\mathcal{S})$ is a much stronger logic than **S4**.

As a first step in this direction, we show that the Grzegorzczuk axiom

$$\mathbf{Grz} = \Box(\Box(P \rightarrow \Box P) \rightarrow P) \rightarrow P$$

belongs to $L(\mathcal{S})$.

3.4.4. FACT. **Grz** is *S-valid*.

Proof **Grz** is S-valid iff $X \subseteq \diamond(X \wedge \neg \diamond(\diamond X \wedge \neg X))$ for any $X \in \mathcal{S}(\mathbb{R})$. Suppose $X \in \mathcal{S}(\mathbb{R})$. Since $\diamond X \wedge \neg X$ is finite, $\diamond(\diamond X \wedge \neg X) = \diamond X \wedge \neg X$. Hence $\diamond(X \wedge \neg \diamond(\diamond X \wedge \neg X)) = \diamond(X \wedge \neg(\diamond X \wedge \neg X)) = \diamond(X \wedge (\neg \diamond X \vee X)) = \diamond X$, which clearly contains X . So, $X \subseteq \diamond(X \wedge \neg \diamond(\diamond X \wedge \neg X))$. QED

As a next step, we show that the axioms

$$\mathbf{BD}_2 = (\neg P \wedge \diamond P) \rightarrow \diamond \Box P, \text{ and}$$

$$\mathbf{BW}_2 = \neg(P \wedge Q \wedge \diamond(P \wedge \neg Q)) \wedge \diamond(\neg P \wedge Q) \wedge \diamond(\neg P \wedge \neg Q),$$

bounding the depth and the width of a Kripke model to 2, are *S-valid*.

3.4.5. FACT. \mathbf{BD}_2 and \mathbf{BW}_2 are \mathcal{S} -valid.

Proof Note that \mathbf{BD}_2 is \mathcal{S} -valid iff $\diamond X \wedge \neg X \subseteq \diamond \Box X$ for any $X \in \mathcal{S}(\mathbb{R})$, and that \mathbf{BW}_2 is \mathcal{S} -valid iff $X \wedge Y \wedge \diamond(X \wedge \neg Y) \wedge \diamond(Y \wedge \neg X) \wedge \diamond(\neg X \wedge \neg Y) = \emptyset$ for any $X, Y \in \mathcal{S}(\mathbb{R})$.

To show that $\diamond X \wedge \neg X \subseteq \diamond \Box X$ for any $X \in \mathcal{S}(\mathbb{R})$, suppose $x \in \diamond X \wedge \neg X$. Then x is a limit point of X not belonging to X . Since X is serial, there is $y \in \mathbb{R}$ such that either $y < x$ and $(y, x) \subseteq X$, or $x < y$ and $(x, y) \subseteq X$. In both cases $x \in \diamond \Box X$. So, $\diamond X \wedge \neg X \subseteq \diamond \Box X$.

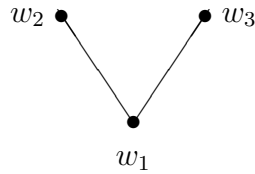
To show that $X \wedge Y \wedge \diamond(X \wedge \neg Y) \wedge \diamond(Y \wedge \neg X) \wedge \diamond(\neg X \wedge \neg Y) = \emptyset$ for any $X, Y \in \mathcal{S}(\mathbb{R})$, suppose $x \in X \wedge Y \wedge \diamond(X \wedge \neg Y) \wedge \diamond(Y \wedge \neg X)$. Then $x \notin \Box X$ and $x \notin \Box Y$. Hence there exist $y, z \in \mathbb{R}$ such that $y < x < z$ and $(y, z) \cap (\neg X \wedge \neg Y) = \emptyset$, which means that $x \notin \diamond(\neg X \wedge \neg Y)$. So, $X \wedge Y \wedge \diamond(X \wedge \neg Y) \wedge \diamond(Y \wedge \neg X) \wedge \diamond(\neg X \wedge \neg Y) = \emptyset$.

QED

The following is an immediate consequence of our observations.

3.4.6. COROLLARY. $\mathbf{S4} + \mathbf{Grz} + \mathbf{BD}_2 + \mathbf{BW}_2 \subseteq L(\mathcal{S})$.

In order to prove the converse, and hence complete our axiomatization of the logic of serial subsets of \mathbb{R} , observe that $\mathbf{S4} + \mathbf{Grz} + \mathbf{BD}_2 + \mathbf{BW}_2$ is actually the complete modal logic of the following ‘2-fork’ Kripke frame $\langle W, R \rangle$, where $W = \{w_1, w_2, w_3\}$ and $w_1 R w_2, w_1 R w_3, w_2 R w_3$:



Indeed, it is well known that \mathbf{Grz} is valid on a Kripke frame iff it is a Noetherian partial order, that \mathbf{BD}_2 is valid on a partially ordered Kripke frame iff its depth is bounded by 2, and that \mathbf{BW}_2 is valid on a partially ordered Kripke frame of a depth ≤ 2 iff its width is bounded by 2. Now, denoting the logic of $\langle W, R \rangle$ by $L(\langle W, R \rangle)$, we have the following:

3.4.7. THEOREM. $\mathbf{S4} + \mathbf{Grz} + \mathbf{BD}_2 + \mathbf{BW}_2 = L(\langle W, R \rangle)$.

Proof Denote $\mathbf{S4} + \mathbf{Grz} + \mathbf{BD}_2 + \mathbf{BW}_2$ by L . Then, $\langle W, R \rangle \models \mathbf{Grz}, \mathbf{BD}_2, \mathbf{BW}_2$. Hence $L \subseteq L(\langle W, R \rangle)$. Conversely, since \mathbf{Grz} is a theorem of L , every L -frame is a Noetherian partial order. Since \mathbf{BD}_2 is a theorem of L , every L -frame is of depth

≤ 2 , hence L has the finite model property, and thus is complete with respect to finite rooted partially ordered Kripke frames of depth ≤ 2 . Since \mathbf{BW}_2 is a theorem of L , then the width of finite rooted L -frames is also ≤ 2 . But then it is routine to check that every such frame is a p -morphic image of $\langle W, R \rangle$. Hence $L(\langle W, R \rangle) \subseteq L$, and $L = L(\langle W, R \rangle)$. QED

As a final move, we show that $\langle W, \tau_R \rangle$ is an open and *serial* image of \mathbb{R} , meaning that there is an open map $f : \mathbb{R} \rightarrow W$ such that $f^{-1}(X) \in \mathcal{S}(\mathbb{R})$ for any subset X of W .

Recall that τ_R consists of the upward closed subsets of W , which obviously are \emptyset , $\{w_2\}$, $\{w_3\}$, $\{w_2, w_3\}$, and W . Fix any $x \in \mathbb{R}$ and define $f : \mathbb{R} \rightarrow W$ by putting

$$f(y) = \begin{cases} w_1 & \text{for } y = x, \\ w_2 & \text{for } y < x, \\ w_3 & \text{for } y > x. \end{cases}$$

Then it is routine to check that $f^{-1}(\emptyset) = \emptyset$, $f^{-1}(\{w_2\}) = (-\infty, x)$, $f^{-1}(\{w_3\}) = (x, +\infty)$, $f^{-1}(\{w_2, w_3\}) = (-\infty, x) \cup (x, +\infty)$, and $f^{-1}(W) = \mathbb{R}$. So, f is continuous. Moreover, for any open subset U of \mathbb{R} , if $x \in U$, then $f(U) = W$; and if $x \notin U$, then $f(U) \subseteq \{w_2, w_3\}$, which is always open. Hence, f is open. Furthermore, from the definition of f it follows that the f -inverse image of any subset of W is a serial subset of \mathbb{R} . So, $\langle W, \tau_R \rangle$ is an open and serial image of \mathbb{R} .

As a trivial consequence of this observation, we obtain that for every valuation \models on $\langle W, R \rangle$ there is a serial valuation \models_S on \mathbb{R} such that $\langle W, R, \models \rangle$ is topo-bisimilar to $\langle \mathbb{R}, \models_S \rangle$. Hence, every non-theorem of $L(\langle W, R \rangle)$ is a non-theorem of $L(S)$, and we have the following:

3.4.8. COROLLARY. $L(S) \subseteq L(\langle W, R \rangle)$.

Combining Corollaries 3.4.6 and 3.4.8 and Theorem 3.4.7 one obtains:

3.4.9. THEOREM. $L(S) = L(\langle W, R \rangle) = \mathbf{S4} + \mathbf{Grz} + \mathbf{BD}_2 + \mathbf{BW}_2$.

3.4.2 Formulas in one variable over the serial sets

This section provides some more concrete information on ‘serial sets’. As $L(S)$ is the logic of the finite ‘2-fork’ frame, for every natural number $n \geq 0$, there are only finitely many $L(S)$ -non-equivalent formulas built from the variables P_1, \dots, P_n . In this subsection we show that there are exactly 64 $L(S)$ -non-equivalent formulas in one variable, and describe them all.

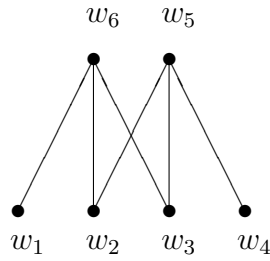
3.4.10. THEOREM. *Every formula in one variable is $L(S)$ -equivalent to a disjunction of the following six formulas:*

$$\begin{aligned}
 & \Box P, \\
 & \Box \neg P, \\
 & P \wedge \Box \Diamond \neg P, \\
 & \neg P \wedge \Box \Diamond P, \\
 & P \wedge \Diamond \Box \neg P \wedge \Diamond \Box P, \text{ and} \\
 & \neg P \wedge \Diamond \Box P \wedge \Diamond \Box \neg P.
 \end{aligned}$$

Hence, there are exactly 64 $L(\mathcal{S})$ -non-equivalent formulas in one variable.

Proof In line with our interest in tying up ‘modal’ and ‘topological’ ways of thinking, we give two different proofs of this result. One proceeds by constructing the 1-universal Kripke model of $L(\mathcal{S})$, which is a standard technique in modal logic, the other is purely topological, using some basic observations on serial subsets of \mathbb{R} .

First Proof Since $L(\mathcal{S})$ is the logic of the ‘2-fork’ frame, we can easily construct the 1-universal Kripke model $\langle W(1), \models_{(1)} \rangle$ of $L(\mathcal{S})$:



Here $w_n \models P$ iff n is even. Now one can readily check that each point of $W(1)$ corresponds to one of the six formulas in the condition of the theorem. Hence every formula in one variable is $L(\mathcal{S})$ -equivalent to a disjunction of the above six formulas. Since there are exactly 2^6 different subsets of $W(1)$, we obtain that there are exactly 64 $L(\mathcal{S})$ -non-equivalent formulas in one variable.

Second Proof Observe that there exists a serial subset X of \mathbb{R} such that $\Box X \neq \Box \neg X \neq X \wedge \Box \Diamond \neg X \neq \neg X \wedge \Box \Diamond X \neq X \wedge \Diamond \Box \neg X \wedge \Diamond \Box X \neq \neg X \wedge \Diamond \Box X \wedge \Diamond \Box \neg X$. For example, let $x < y < z < u$, and take $X = [x, y) \cup (y, z) \cup \{u\}$. Then one can readily check that

$$\begin{aligned}
 \Box X &= (x, y) \cup (y, z), \\
 \Box \neg X &= (-\infty, x) \cup (z, u) \cup (u, +\infty), \\
 X \wedge \Box \Diamond \neg X &= \{u\}, \\
 \neg X \wedge \Box \Diamond X &= \{y\}, \\
 X \wedge \Diamond \Box \neg X \wedge \Diamond \Box X &= \{x\}, \text{ and} \\
 \neg X \wedge \Diamond \Box X \wedge \Diamond \Box \neg X &= \{z\}.
 \end{aligned}$$

Hence, we can always interpret P as a serial subset of \mathcal{R} such that all the six formulas of the theorem correspond to different serial subsets of \mathcal{R} .

Now, we prove that every subset of \mathcal{R} obtained by repeatedly applying \neg, \wedge, \Box to a serial set X is a finite (including the empty) union of the following serial subsets:

$$\begin{aligned} T_1 &= \Box X, \\ T_2 &= \Box \neg X, \\ T_3 &= X \wedge \Box \Diamond \neg X, \\ T_4 &= \neg X \wedge \Box \Diamond X, \\ T_5 &= X \wedge \Diamond \Box \neg X \wedge \Diamond \Box X, \text{ and} \\ T_6 &= \neg X \wedge \Diamond \Box X \wedge \Diamond \Box \neg X. \end{aligned}$$

For this, first observe that $T_i \wedge T_j = \emptyset$ if $i \neq j$, and that $\bigvee_{i=1}^6 T_i = \mathcal{R}$. So, these six serial subsets of \mathcal{R} are mutually disjoint and jointly exhaustive. Next observe that $\neg T_i = T_j \vee T_k \vee T_l \vee T_m \vee T_n$, where $i, j, k, l, m, n \in \{1, 2, 3, 4, 5, 6\}$ are different from each other. Finally, $\Box T_1 = T_1$, $\Box T_2 = T_2$, and $\Box T_3 = \Box T_4 = \Box T_5 = \Box T_6 = \emptyset$.

Hence every subset of \mathcal{R} obtained by repeatedly applying \neg, \wedge, \Box to $\{T_1, \dots, T_6\}$ is a finite (including the empty) union of $\{T_1, \dots, T_6\}$.

Now suppose $Y \subseteq \mathcal{R}$ is obtained by repeatedly applying \neg, \wedge, \Box to X . We prove by induction on the complexity of Y that Y is equal to a finite (including the empty) union of $\{T_1, \dots, T_6\}$.

Base case. Since $X = T_1 \vee T_3 \vee T_5$ (and $\neg X = T_2 \vee T_4 \vee T_6$), the base case (that is when $Y = X$) is obvious.

Complement. Suppose $Y = \neg Z$ and $Z = T_{i_1} \vee \dots \vee T_{i_k}$, where $i_1, \dots, i_k \in \{1, \dots, 6\}$. Then $Y = \neg(T_{i_1} \vee \dots \vee T_{i_k}) = \neg T_{i_1} \wedge \dots \wedge \neg T_{i_k}$. Since every $\neg T_{i_j}$ is equal to $\bigvee_{i_s \neq i_j} T_{i_s}$, using the distributivity law we obtain that $Y = \bigvee_{i_s, i_t \in \{1, \dots, 6\}} (T_{i_s} \wedge T_{i_t})$. Since for different i_s and i_t , $T_{i_s} \wedge T_{i_t} = \emptyset$, which is the empty union of T_i s, we finally obtain that Y is a finite union of $\{T_1, \dots, T_6\}$.

Intersection. Suppose $Y = Z_1 \wedge Z_2$, $Z_1 = T_{i_1} \vee \dots \vee T_{i_k}$ and $Z_2 = T_{j_1} \vee \dots \vee T_{j_m}$, where $i_1, \dots, i_k, j_1, \dots, j_m \in \{1, \dots, 6\}$. Similarly to the above case, using the distributivity law we obtain that Y is a finite union of $\{T_1, \dots, T_6\}$.

Interior. Suppose $Y = \Box Z$ and $Z = T_{i_1} \vee \dots \vee T_{i_k}$, where $i_1, \dots, i_k \in \{1, \dots, 6\}$. Since T_i s are mutually disjoint, $Y = \Box T_{i_1} \vee \dots \vee \Box T_{i_k}$. Now since $\{T_1, \dots, T_6\}$ is closed with respect to \Box , we obtain that Y is a finite union of $\{T_1, \dots, T_6\}$.

Hence, every subset of \mathcal{R} obtained by repeatedly applying \neg, \wedge, \Box to a serial set X is equal to a finite (including the empty) union of $\{T_1, \dots, T_6\}$. Since there are exactly 2^6 different subsets obtained as a union of $\{T_1, \dots, T_6\}$, we obtain that there are exactly 64 different subsets of \mathcal{R} obtained by repeatedly applying \neg, \wedge, \Box to a serial set X . This implies that there are exactly 64 $L(\mathcal{S})$ -non-equivalent formulas in one variable.

QED

The same technique can also be used to prove the normal form theorem over $L(\mathcal{S})$ for every formula with more than one proposition variable.

3.4.3 Countable unions of convex sets on the real line

Let us now be a bit more systematic. By Theorem 3.3.8, **S4** is the complete logic of \mathbb{R} , and hence sets of reals suffice as values $\nu(P)$ in refuting non-theorems. But how complex must these sets be? In first-order logic, e.g., we know that completeness requires atomic predicates over the integers which are at least Δ_2^0 . With only simpler predicates in the arithmetic hierarchy, the logic gets richer. In a topological space like \mathbb{R} , it seems reasonable to look at the Borel Hierarchy \mathcal{G} . How high up do we have to go for our **S4**-counterexamples? One could analyze our construction in Section 3.3.3 to have an upper bound. But here we state some more direct information.

Consider the set τ of all open subsets of \mathbb{R} . Let $\mathcal{B}(\tau)$ denote the Boolean closure of τ . Since $\mathcal{B}(\tau)$ contains all closed subsets of \mathbb{R} , then $\mathcal{B}(\tau)$ is closed with respect to \diamond . Obviously $\mathcal{S}(\mathbb{R})$ is properly contained in $\mathcal{B}(\tau)$. It is natural to ask whether the elements of $\mathcal{B}(\tau)$ are enough for refuting all the non-theorems of **S4**. The answer is negative: the modal logic is still richer.

3.4.11. FACT (BEZHANISHVILI AND GEHRKE). The complete logic of $\mathcal{B}(\tau)$ is **Grz**.

Hence, we need to seek something bigger than $\mathcal{B}(\tau)$. Let $\mathcal{C}^\infty(\mathbb{R})$ denote the set of countable unions of convex subsets of \mathbb{R} . Since every open subset of \mathbb{R} is a countable union of open intervals, then $\tau \subseteq \mathcal{C}^\infty(\mathbb{R})$. Let $\mathcal{B}(\mathcal{C}^\infty(\mathbb{R}))$ denote the Boolean closure of $\mathcal{C}^\infty(\mathbb{R})$. Since $\tau \subseteq \mathcal{C}^\infty(\mathbb{R})$, we also have $\mathcal{B}(\tau) \subseteq \mathcal{B}(\mathcal{C}^\infty(\mathbb{R}))$. It follows that $\mathcal{B}(\mathcal{C}^\infty(\mathbb{R}))$ is also closed with respect to \diamond . Moreover, $\mathcal{B}(\tau)$ is properly contained in $\mathcal{B}(\mathcal{C}^\infty(\mathbb{R}))$, since the set \mathbf{Q} of rationals belongs to $\mathcal{B}(\mathcal{C}^\infty(\mathbb{R}))$ but does not belong to $\mathcal{B}(\tau)$.

3.4.12. THEOREM (BEZHANISHVILI AND GEHRKE). The logic **S4** is complete with respect to $\mathcal{B}(\mathcal{C}^\infty(\mathbb{R}))$.

So, the Boolean combinations of countable unions of convex subsets of \mathbb{R} are exactly what we need for refuting the non-theorems of **S4**. Since every countable union of convex subsets of \mathbb{R} belongs to the Borel hierarchy \mathcal{G}_2 over the opens of \mathbb{R} , a very low level of the Borel hierarchy suffices for refuting the non-theorems of **S4**. So, \mathcal{G} itself is more than sufficient for refuting the non-theorems of **S4**.

Summarizing, we constructed five Boolean algebras of subsets of \mathbb{R} forming a chain under inclusion: $\mathcal{S}(\mathbb{R}) \subset \mathcal{B}(\tau) \subset \mathcal{B}(\mathcal{C}^\infty(\mathbb{R})) \subset \mathcal{G} \subset \mathcal{P}(\mathbb{R})$, where $\mathcal{S}(\mathbb{R})$ is the Boolean algebra of all serial subsets of \mathbb{R} , $\mathcal{B}(\tau)$ the Boolean closure of the set of all open subsets of \mathbb{R} , $\mathcal{B}(\mathcal{C}^\infty(\mathbb{R}))$ the Boolean closure of the set of all countable unions of convex subsets of \mathbb{R} , \mathcal{G} the Boolean algebra of all Borel subsets of \mathbb{R} , and $\mathcal{P}(\mathbb{R})$ the power-set of \mathbb{R} . All of these Boolean algebras are closed with respect to \diamond . The modal logic of the last three algebras is **S4**, that of the second one is **Grz**, and the modal logic of the first is the logic of the ‘2-fork’ Kripke frame.

3.4.4 Generalization to \mathbb{R}^2

In this section, we shift aim in a different direction. We generalize our results on the serial subsets of \mathbb{R} to the chequered subsets of \mathbb{R}^2 , and indicate further generalizations to any Euclidean space \mathbb{R}^n .

A set $X \subseteq \mathbb{R}^2$ is *convex* if all points laying in between any two points of X belong to X . It is said to be *serial* if X is a finite union of convex subsets of \mathbb{R}^2 . Denote the set of all serial subsets of \mathbb{R}^2 by $\mathcal{S}(\mathbb{R}^2)$.

Here is a real difference between \mathbb{R} and \mathbb{R}^2 . Unlike $\mathcal{S}(\mathbb{R})$, $\mathcal{S}(\mathbb{R}^2)$ is not closed with respect to complement. For instance, a full circle is obviously a convex subset of \mathbb{R}^2 . However, its complement is not serial.

One natural way of overcoming this difficulty is to work with a smaller family of chequered subsets of \mathbb{R}^2 , which also has a reasonable claim to being ‘the two-dimensional generalization of the one-dimensional serial sets’.

A set $X \subseteq \mathbb{R}^2$ is a *rectangular convex* if $X = X_1 \times X_2$, where both X_1 and X_2 are convex subsets of \mathbb{R} [van Benthem, 1983b]. Every rectangular convex is a convex set in the usual sense, but not vice versa: a circle is not a rectangular convex.

A set $X \subseteq \mathbb{R}^2$ is said to be *chequered* if it is a finite union of rectangular convex subsets of \mathbb{R}^2 . Denote the set of all chequered subsets of \mathbb{R}^2 by $\mathcal{CH}(\mathbb{R}^2)$. Obviously $\mathcal{CH}(\mathbb{R}^2) \subset \mathcal{S}(\mathbb{R}^2)$. Note that unlike $\mathcal{S}(\mathbb{R}^2)$, $\mathcal{CH}(\mathbb{R}^2)$ does form a Boolean algebra. Moreover, $\Box X, \Diamond X \in \mathcal{CH}(\mathbb{R}^2)$ for any $X \in \mathcal{CH}(\mathbb{R}^2)$.

3.4.13. FACT. $\mathcal{CH}(\mathbb{R}^2)$ forms a Boolean algebra closed with respect to \Box and \Diamond .

Proof In order to show that $\mathcal{CH}(\mathbb{R}^2)$ forms a Boolean algebra it is sufficient to show that $\mathcal{CH}(\mathbb{R}^2)$ is closed with respect to \neg . For this, observe that the complement of a rectangular convex is a union of at most four rectangular convexes, and that the finite intersection of rectangular convexes is again a rectangular convex. Now, suppose $A \in \mathcal{CH}(\mathbb{R}^2)$. Then there exist rectangular convexes A_1, \dots, A_n such that $A = \bigcup_{i=1}^n A_i$. But $\neg A = \bigcap_{i=1}^n \neg A_i$, which by the above observation and distributivity is chequered.

Since $\mathcal{CH}(\mathbb{R}^2)$ forms a Boolean algebra, in order to show that $\mathcal{CH}(\mathbb{R}^2)$ is closed with respect to \Box and \Diamond , it is sufficient to check that $\mathcal{CH}(\mathbb{R}^2)$ is closed with respect to \Diamond . For the latter observe that the closure of a rectangular convex is again a rectangular convex, and that the closure commutes with finite unions. Now suppose $A \in \mathcal{CH}(\mathbb{R}^2)$. Then there exist rectangular convexes A_1, \dots, A_n such that $A = \bigcup_{i=1}^n A_i$. But then $\Diamond A = \bigcup_{i=1}^n \Diamond A_i$, which is a chequered set by the above observation. QED

Hence, interpreting propositional variables as chequered subsets of \mathbb{R}^2 , every formula of our language will be also interpreted as a chequered subset of \mathbb{R}^2 .

This approach leads to a logic, which we just sketch here. Call a valuation ν of \mathcal{L} to subsets of \mathbb{R}^2 *chequered* if $\nu(P) \in \mathcal{CH}(\mathbb{R}^2)$ for any propositional variable P . Since $\mathcal{CH}(\mathbb{R}^2)$ is closed with respect to \neg, \wedge and \Diamond , we have that $\nu(\varphi) \in \mathcal{CH}(\mathbb{R}^2)$ for any chequered interpretation ν . Call a formula φ *\mathcal{CH} -true* if it is true in \mathbb{R}^2 under a

chequered valuation. Call φ \mathcal{CH} -valid if φ is \mathcal{CH} -true for any chequered valuation on \mathbb{R}^2 . Let $L(\mathcal{CH}) = \{\varphi \mid \varphi \text{ is } \mathcal{CH}\text{-valid}\}$.

3.4.14. FACT. $L(\mathcal{CH})$ is a normal modal logic over **S4**.

Similarly to $L(\mathcal{S})$, the Grzegorzcyk axiom **Grz** is provable in $L(\mathcal{CH})$. For this it is sufficient to show that **Grz** is \mathcal{CH} -valid.

3.4.15. FACT. **Grz** is \mathcal{CH} -valid.

Proof **Grz** is \mathcal{CH} -valid iff $X \subseteq \diamond(X \wedge \neg\diamond(\diamond X \wedge \neg X))$ for any $X \in \mathcal{CH}(\mathbb{R}^2)$. Suppose $X \in \mathcal{CH}(\mathbb{R}^2)$. Observe that, unlike $\mathcal{S}(\mathbb{R})$, $\diamond X \wedge \neg X$ is not finite. However, in this case the set $\diamond(\diamond X \wedge \neg X) - (\diamond X \wedge \neg X)$ is finite. Denote it by F . Then $\diamond(X \wedge \neg\diamond(\diamond X \wedge \neg X)) = \diamond(X \wedge \neg[(\diamond X \wedge \neg X) \vee F]) = \diamond(X \wedge (\neg\diamond X \vee X) \wedge \neg F) = \diamond(X - F)$. Now since F is finite, $\diamond(X - F) = \diamond X$. Therefore, $\diamond(X \wedge \neg\diamond(\diamond X \wedge \neg X)) = \diamond X$, which obviously contains X . So, $X \subseteq \diamond(X \wedge \neg\diamond(\diamond X \wedge \neg X))$. QED

Now we show that the axioms

$$\begin{aligned} \mathbf{BD}_3 &= \diamond(\Box P_3 \wedge \diamond(\Box P_2 \wedge \diamond\Box P_1 \wedge \neg P_1) \wedge \neg P_2) \rightarrow P_3, \text{ and} \\ \mathbf{BW}_4 &= \bigwedge_{i=0}^4 \diamond P_i \rightarrow \bigvee_{0 \leq i \neq j \leq 4} \diamond(P_i \wedge \diamond P_j), \end{aligned}$$

which bound the depth and the width of a Kripke model to 3 and 4, respectively, are also provable in $L(\mathcal{CH})$. For this, we show that both \mathbf{BD}_3 and \mathbf{BW}_4 are \mathcal{CH} -valid.

3.4.16. FACT. (1) \mathbf{BD}_3 is \mathcal{CH} -valid.

(2) \mathbf{BW}_4 is \mathcal{CH} -valid.

Proof (1) \mathbf{BD}_3 is \mathcal{CH} -valid iff $\diamond(\Box X_3 \wedge \diamond(\Box X_2 \wedge \diamond\Box X_1 \wedge \neg X_1) \wedge \neg X_2) \subseteq X_3$ for any $X_1, X_2, X_3 \in \mathcal{CH}(\mathbb{R}^2)$. Observe that $\diamond\Box X_1 \wedge \neg X_1$ is a subset of the frontier $Fr(X_1) = \diamond X_1 \wedge \neg\Box X_1$ of X_1 . Hence, $\diamond(\Box X_3 \wedge \diamond(\Box X_2 \wedge \diamond\Box X_1 \wedge \neg X_1) \wedge \neg X_2) \subseteq \diamond(\Box X_3 \wedge \diamond(\Box X_2 \wedge Fr(X_1)) \wedge \neg X_2)$. Let $X_2^* = \Box X_2 \wedge Fr(X_1)$ and $X_3^* = \Box X_3 \wedge Fr(X_1)$. Also let \neg^* , \diamond^* and \Box^* denote the corresponding operations of a closed subspace $Fr(X_1)$ of \mathbb{R}^2 . Then $\diamond(\Box X_3 \wedge \diamond(\Box X_2 \wedge Fr(X_1)) \wedge \neg X_2) = \diamond(\Box X_3 \wedge \diamond X_2^* \wedge \neg X_2) = \diamond(\Box X_3 \wedge \diamond^* X_2^* \wedge \neg X_2) \subseteq \diamond(\Box X_3 \wedge \diamond^* X_2^* \wedge \neg\Box X_2) = \diamond(\Box X_3 \wedge \diamond^* X_2^* \wedge \neg^* X_2^*) = \diamond(X_3^* \wedge \diamond^* X_2^* \wedge \neg^* X_2^*) = \diamond^*(X_3^* \wedge \diamond^* X_2^* \wedge \neg^* X_2^*)$. Since $Fr(X_1)$ is of dimension 1, $Fr(X_1)$ is homeomorphic to a closed serial subspace of \mathbb{R} . Since \mathbf{BD}_2 is \mathcal{S} -valid in \mathbb{R} , $\diamond^*(X \wedge \diamond^* Y \wedge \neg^* Y) \subseteq X$ for any open subsets X, Y of $Fr(X_1)$. Hence, $\diamond^*(X_3^* \wedge \diamond^* X_2^* \wedge \neg^* X_2^*) \subseteq X_3^*$. Thus, $\diamond(\Box X_3 \wedge \diamond(\Box X_2 \wedge \diamond\Box X_1 \wedge \neg X_1) \wedge \neg X_2) \subseteq X_3$, and \mathbf{BD}_3 is \mathcal{CH} -valid.

(2) \mathbf{BW}_2 is \mathcal{CH} -valid iff $\bigwedge_{i=0}^4 \diamond X_i \subseteq \bigvee_{0 \leq i \neq j \leq 4} \diamond(X_i \wedge \diamond X_j)$ for any $X_0, \dots, X_4 \in \mathcal{CH}(\mathbb{R}^2)$. Suppose $x \in \bigwedge_{i=0}^4 \diamond X_i$. Then x is a limit point of all X_i . Since there are

five X_i , and every X_i belongs to $\mathcal{CH}(\mathbb{R}^2)$, there should exist X_i and X_j such that x is a limit point of $X_i \wedge X_j$. So, $x \in \bigvee_{0 \leq i \neq j \leq 4} \diamond(X_i \wedge \diamond X_j)$. QED

As an immediate consequence we obtain that $L(\mathcal{CH}) \vdash \text{Grz}, \text{BD}_3, \text{BW}_4$. Hence, like $L(\mathcal{S})$, $L(\mathcal{CH})$ is also a tabular logic. In a similar fashion, by induction on the dimension of \mathbb{R}^n , we can prove that the logic of chequered subsets of \mathbb{R}^n is also tabular. In particular, it validates BD_{n+1} and BW_{2^n} . Hence, we are capable of capturing the dimension of Euclidean spaces.

3.5 A general picture

3.5.1 The deductive landscape

The logics that we have studied in this chapter fit into a more general environment. Typical for modal logic is its lattice of deductive systems such as **K**, **S4**, **S5** or **GL**. These form a large family describing different classes of relational frames, with often very different motivations (cf. the series of books “Advances in Modal Logic²”). Among the uncountably many modal logics, a small number are distinguished for one of two reasons. Logics like **S4** or **S5** were originally proposed as syntactic proof theories for notions of modality, and then turned out to be semantically complete with respect to natural frame classes, such as (for **S4**) transitive reflexive orders. Other modal logics, however, were discovered as the complete theories of important frames, such as the natural numbers with their standard ordering. What about a similar landscape of modal logics on the topological interpretation?

Some well-known modal logics extending **S4** indeed correspond to natural classes of topological spaces. E.g., it is easy to see that the ‘identity logic’ with axiom $\varphi \rightarrow \Box\varphi$ axiomatizes the complete logic of all *discrete spaces*. And it also defines them semantically through the usual notion of *frame correspondence*—which can be lifted to the topological semantics in a straight-forward manner. But already **S5** corresponds to a less standard condition, viz. that every point has an open neighborhood all of whose points have x in all their open neighborhoods. (Alternatively, this says that every open set is closed.) Also, even rich topological spaces do not seem to validate very spectacular modal logics, witness the fact that \mathbb{R} has just **S4** for its modal theory. We did find stronger logics with ‘general frames’ though, i.e., frames with a designated interior algebra of subsets, such as \mathbb{R} with the serial sets. The latter turned out to be a well-known modal ‘frame logic’, and we have not been able so far to find really new modal logics arising on the topological interpretation.

A related question is what becomes of the known general results on completeness and correspondence for modal logic in the topological setting. There appear to be some obstacles here. E.g., the substitution method for Sahlqvist correspondence (cf.

²<http://www.aiml.net>

[Blackburn et al., 2001]) has only a limited range. It does work for axioms like the above $\varphi \rightarrow \Box\varphi$, where it automatically generates the first-order condition

$$(\forall x)(\exists U \in \tau)(x \in U \ \& \ (\forall y \in U)(y = x)),$$

i.e., discreteness. Likewise, it works for the **S5** symmetry axiom $P \rightarrow \Box\Diamond P$, where it produces the above-mentioned

$$(\forall x)(\exists U \in \tau)(x \in U \ \& \ (\forall y \in U)(\forall V \in \tau)(y \in V \rightarrow x \in V)).$$

The method also works for antecedents of the form $\Box P$ —but things stop with antecedents like $\Diamond P$ or $\Box\Box P$. The reason is that, on the topological semantics, one modality \Box expresses a *two-quantifier* combination

$$\exists U \in \tau \text{ such that } \forall x \in U,$$

so that syntactic complexity builds up more rapidly than in standard modal logic, where each modality is one quantifier over relational successors of the current world. General correspondence or completeness results for topological modal logics therefore seem harder to obtain—and we may need different syntactic notions for them (see [Gabelaia, 2002] for recent results in that direction).

Modal logics are, most notably, languages for describing relational structures. One considers these formalisms, in contrast with first or second-order theories, because of the nice balance between expressive power and computational properties. The logic **S4** introduced in Chapter 2 is the minimal normal modal logic with the topological interpretation, as shown in Chapter 3. It is a general formalism with respect to topological structures as it is complete for all topological spaces. Such a high abstraction is a beauty, but also a handicap. The language is not expressive enough and cannot capture specifics of some interesting topological spaces.

An extremely useful technique in modal logics to gain expressive power without leaving the guarded area of decidable languages is to add a modal operator. For instance, if one needs to express notions connected to equality of states in Kripke semantics, one may add a difference operator $D\varphi$ which reads “there is a state different from the current one that satisfies φ .” This is exactly what we do in this chapter. We consider important topological relations not captured by **S4** alone which can be safely expressed by ‘adding’ appropriate new modal operators. We have entered the realm of extended modal languages, see [de Rijke, 1993, van Benthem, 1991b].

The first limitation to overcome is **S4**’s locality. The formulas are evaluated at points and provide local information, e.g., the point x is in the open set given by the intersection of the interior of φ and ψ ($M, x \models \Box\varphi \wedge \Box\psi$). By this information we know a lot about the point x , but very little about the set denoted by $\Box\varphi \wedge \Box\psi$, we merely know that there is one point satisfying it, the point x . Introducing an universal (or global) modality is the solutions to this problem. For instance, with **S4+**(the universal modality) one is able to express whether a topological space is connected or not, which is clearly a global property of the space and not a local one of some points of the space. We shall explain such behavior in Section 4.1. Extending with different modal operators enables different gains in expressive power, we present alternative extensions in Section 4.2. These extensions can be viewed as a fragment of higher order languages. We give a higher order formalism in Section 4.3 to give a general perspective.

4.1 Universal reference

Even though ideas related to the universal modality have been around for a while, [Prior, 1967], it is safe to say that it was ‘officially introduced’ as a modal logic extension tool in [Goranko and Passy, 1992]. In [Bennett, 1995], Bennett introduced the universal language topologically interpreted to identify tractable fragments of a language of topological relations over regions.

The truth definition of **S4**, Definition 2.1.1, is extended with the following:

$$\begin{aligned} M, x \models E\varphi & \text{ iff } \exists y \in X, M, y \models \varphi \\ M, x \models U\varphi & \text{ iff } \forall y \in X : M, y \models \varphi \end{aligned}$$

The definition reads, for $E\varphi$, “there exists a point in the model satisfying φ ,” and dually for $U\varphi$, “all the points in the model satisfy φ .” The U and E modalities follow the axiomatization of **S5**:

$$\begin{aligned} U(\varphi \rightarrow \psi) & \rightarrow (U\varphi \rightarrow U\psi) & \text{(K)} \\ U\varphi & \rightarrow \varphi & \text{(T)} \\ U\varphi & \rightarrow UU\varphi & \text{(4)} \\ \varphi & \rightarrow UE\varphi & \text{(B)} \end{aligned}$$

In addition, the following ‘connecting’ principle is part of the axioms:

$$\diamond\varphi \rightarrow E\varphi \quad \text{(Con)}$$

The axiomatization suggests to search for a normal form. The nesting of universal modal operators is redundant, as the next proposition shows.

4.1.1. PROPOSITION. *Every formula of $\mathbf{S4}_u$ is equivalent to one without nested occurrences of E, U .*

Proof Here is one way of seeing this. The following well formed formula is valid in the semantics of $\mathbf{S4}_u$. Let $\varphi[E\psi]$ be any formula containing a subformula $E\psi$. Then we have

$$\varphi[E\psi] \leftrightarrow (E\psi \wedge \varphi[\top]) \vee (\neg E\psi \wedge \varphi[\perp])$$

The reason is that subformulas $E\psi$ are globally true or false, across modalities \Box, \Diamond, E, U . This observation also produces an effective algorithm for finding the normal form. E.g.

$$\begin{aligned} \Box(Ep \wedge \neg\Box Eq) & \leftrightarrow \\ (Ep \wedge \Box(\top \wedge \neg\Box Eq)) \vee (\neg Ep \wedge \Box(\perp \wedge \neg\Box Eq)) & \leftrightarrow \\ (Ep \wedge \Box\neg\Box Eq) \vee (\neg Ep \wedge \Box\perp) & \leftrightarrow \\ (Ep \wedge ((Eq \wedge \Box\neg\Box\top) \vee (\neg Eq \wedge \Box\neg\Box\perp))) \vee (\neg Ep \wedge \perp) & \leftrightarrow \\ (Ep \wedge Eq \wedge \Box\neg\top) \vee (Ep \wedge \neg Eq \wedge \Box\neg\perp) \vee \perp & \leftrightarrow \\ (Ep \wedge Eq \wedge \perp) \vee (Ep \wedge \neg Eq) & \leftrightarrow \\ Ep \wedge \neg Eq & \end{aligned}$$

QED

Another way of seeing this is by proving some more familiar reduction principles (either in the semantics, or from the given axioms), such as

$$\diamond E\varphi \leftrightarrow E\varphi, \quad \square E\varphi \leftrightarrow E\varphi$$

Note that we do not get, e.g., $E\square\varphi \leftrightarrow \square\varphi$ or $E\square\varphi \leftrightarrow E\varphi$. The normal forms that we obtain may be described as follows,

$$\bigvee \bigwedge [U|E]\varphi$$

where $[U|E]$ is U or E or nothing, and φ is a formula of our original language $\mathbf{S4}$.

One extends, together with the truth definition of the language, all of the tools in the topo-approach (Chapter 2): first and foremost, topo-bisimulations. Definition 2.1.2 straightforwardly extends. In fact it is exactly the same, except for the constraint that the relation has to be defined for all points of the spaces, in the ‘universal’ spirit of the extended language.

4.1.2. DEFINITION (TOPOLOGICAL BISIMULATION). Given two topological models $\langle X, O, \nu \rangle, \langle X', O', \nu' \rangle$, a *total topological bisimulation* is a non-empty relation $\rightleftharpoons \subseteq X \times X'$ defined for all $x \in X$ and for all $x' \in X'$ such that if $x \rightleftharpoons x'$:

- (base): $x \in \nu(p)$ iff $x' \in \nu'(p)$ (for any proposition p)
- (forth condition): if $x \in o \in O$ then
 $\exists o' \in O' : x' \in o' \text{ and } \forall y' \in o' : \exists y \in o : y \rightleftharpoons y'$
- (back condition): if $x' \in o' \in O'$ then
 $\exists o \in O : x \in o \text{ and } \forall y \in o : \exists y' \in o' : y \rightleftharpoons y'$

If only conditions (i) and (ii) hold, the second model *simulates* the first one.

One must show that the above definition is adequate.

4.1.3. THEOREM. Let $M = \langle X, O, \nu \rangle, M' = \langle X', O', \nu' \rangle$ be two models, $x \in X$, and $x' \in X'$ bisimilar points. Then, for any modal formula φ in $\mathbf{S4}_w$, $M, x \models \varphi$ iff $M', x' \models \varphi$.

4.1.4. THEOREM. Let $M = \langle X, O, \nu \rangle, M' = \langle X', O', \nu' \rangle$ be two models with finite $O, O', x \in X$, and $x' \in X'$ such that for every φ in $\mathbf{S4}_w$, $M, x \models \varphi$ iff $M', x' \models \varphi$. Then there exists a total bisimulation between M and M' connecting x and x' .

In words, extended modal formulas are invariant under total bisimulations, while finite modally equivalent models are totally bisimilar.

The other fundamental tool of the topo-approach is the definition of model comparison games. Here is the extension of Definition 2.2.1.

4.1.5. DEFINITION (TOPO-GAME). Consider two topological models $\langle X, O, \nu \rangle$, $\langle X', O', \nu' \rangle$ and a natural number n . A *topo-game* of length n , notation $TG(X, X', n)$, consists of n rounds between two players, Spoiler and Duplicator, who move alternatively. Spoiler is granted the first move and always chooses which type of round to engage. The two sorts of rounds are as follows:

- | | | |
|---------------|---|---|
| global | { | <ul style="list-style-type: none"> (i) Spoiler chooses a model X_s and picks a point \bar{x}_s anywhere in X_s (ii) Duplicator chooses a point \bar{x}_d anywhere in the other model X_d |
| local | { | <ul style="list-style-type: none"> (i) Spoiler chooses a model X_s and an open o_s containing the current point x_s of that model (ii) Duplicator chooses an open o_d in the other model X_d containing its current point x_d (iii) Spoiler picks a point \bar{x}_d in Duplicator's open o_d in the X_d model (iv) Duplicator replies by picking a point \bar{x}_s in Spoiler's open o_s in X_s |

The points \bar{x}_s and \bar{x}_d become the new current points. A game always starts by a global round. By this succession of actions, two sequences are built: $\{x_1, x_2, \dots, x_n\}$ and $\{x'_1, x'_2, \dots, x'_n\}$. After n rounds, if x_i and x'_i (with $i \in [1, n]$) satisfy the same propositional atoms, Duplicator *wins*, otherwise, Spoiler wins. A *winning strategy (w.s.)* for Duplicator is a function from any sequence of moves by Spoiler to appropriate responses which always ends in a win. Spoiler's winning strategy is defined dually.

The *multi-modal rank* of a $\mathbf{S4}_u$ formula is the maximum number of nested modal operators appearing in it (i.e., \square , \diamond , U and E modalities). The following adequacy of the games with respect to the mereotopological language holds.

4.1.6. THEOREM (ADEQUACY). Duplicator has a winning strategy for n rounds in $TG(X, X', n)$ iff X, X' satisfy the same formulas of multi-modal rank at most n .

The interesting result is that of having a game theoretic tool to compare topological models. Given any two models, they can be played upon. If Spoiler has a winning strategy in a certain number of rounds, then the two models are different up to a certain degree. The degree is exactly the minimal number of rounds needed by Spoiler to win. On the other hand, one knows that if Spoiler has no w.s. in any number of rounds, and therefore Duplicator has in all games, including the infinite round game, then the two models are bisimilar.

4.1.1. EXAMPLE (COMPARING CUTLERY). As we did in Section 2.2, we can play on 'table items', i.e., regions in topological spaces. Differently from the local games, one

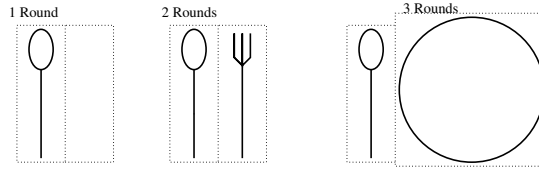


Figure 4.1: Plays of topological games for the universal language $\mathbf{S4}_u$. Above the two models is the number of rounds needed by Spoiler to win.

may notice that there is no starting points in the two models. Spoiler can decide where to play, by means of a global move. By this added freedom, Spoiler can win games in which the players compare spoons and forks, spoons and plates or even spoons with an empty table cloth.

Similar to before, we have a way of tying Spoiler’s winning strategies with formulas (of $\mathbf{S4}_u$) true in the models. Note that the formulas can be true in the entire model, not in only two particular starting points, as before. This reflects our earlier observation that E or U formulas are really true across a model.

Referring to Figure 4.1, we can write down a distinguishing formula of the appropriate multi-modal rank that is true in one model but not in the other. In the case of the 1 round game, Spoiler can win in one round since on the right model the formula Ep is true, while its negation is true in the other model. Think of it as the empty table which should be set, so there is no region p yet: $U\neg p$.

By a similar reasoning we can write the formula $E\Box p$ (the interior of p is non empty) for the 2 round game. This formula is only true in the left model. For the 3 round game, a distinguishing formula is $U(p \leftrightarrow \Diamond\Box p)$. This formula encodes closed regularity of regions, i.e., coincidence with the closure of its interior. This formula is true for the plate on the right but not for the spoon on the left. The negation of the regularity formula can be written as $E(p \wedge \Box\Diamond\neg p) \vee E(\neg p \wedge \Diamond\Box p)$. The first half of this accounts for external lower-dimensional spikes in the region p , the second for lower dimensional cracks. For the spoon the handle is a lower dimensional spike.

4.1.2. REMARK (INFINITE GAMES). The definition can be easily extended to infinite games. Just let $n \rightarrow \infty$ and hence the sequences x_n, x'_n be infinite. The Adequacy Theorem is still valid. Duplicator has a winning strategy in the infinite round game iff the models are bisimilar in our extended sense.

4.1.3. REMARK (STRATEGIES AND NORMAL FORMS). From the practical perspective of playing topological games, Spoiler should bear in mind that identifying formulas that differentiate the models is not enough. Spoiler may consume too many turns if he is using a long formula (in terms of multi-modal depth) which has a shorter logical equivalent. Similarly Duplicator may have the illusion of a win, if he makes the same mistake. Once ‘difference formulas’ are identified in the models they should be reduced to logically equivalent ones with the lowest multi-modal depth. Normal forms

are of great help for this purpose. E.g., here is the game-theoretic content of our earlier normal form for $\mathbf{S4}_u$. Having only one ‘outermost’ existential or universal modality means that Spoiler needs to engage only once in a global round. Furthermore, since such a modality is the first to appear, that is the first type of move Spoiler should play. This can also be seen directly in the game. If Spoiler engages in more than one global round, it is like jumping around the space, not having understood where the difference between the models resides.

One might try to extend this line of reasoning to the inner $\mathbf{S4}$ part. After all, $\mathbf{S4}$ validates reduction laws like $\Box\Box\varphi \leftrightarrow \Box\varphi$, or $\Box\Diamond\Box\Diamond\varphi \leftrightarrow \Box\Diamond\varphi$. Can this be used to simplify Spoiler’s strategies? We have not been able to find a general principle here that would be of much use.

The use of normal forms can lead to a redefinition of the rules. The new game would have always one starting global round and thereafter only local rounds.

Finally, after having presented all the tools of the topo-approach for the extended language, it is important to remark what $\mathbf{S4}_u$ captures of the topological structure.

The relation between $\mathbf{S4}_u$ and connected spaces has recent origins: [Shehtman, 1999] and [Aiello and van Benthem, 1999]. A topological space is defined to be *connected* if the only two sets that are both open and closed are the empty set and whole space itself. The definition is expressible in $\mathbf{S4}_u$ in the following way:

$$U(\Diamond p \rightarrow \Box p) \rightarrow Up \vee U\neg p \quad (4.1)$$

In topology, an alternative definition of connected space (cf. page 30) states that a space is connected if there do not exist two open sets whose union covers the whole space and that are disjoint. Again we can express the phrasing of the theorem in $\mathbf{S4}_u$:

$$U(\Box p \vee \Box q) \wedge Ep \wedge Eq \rightarrow E(p \wedge q) \quad (4.2)$$

Here is the purely logical version of the well-known topological fact.

4.1.7. FACT. $\vdash_{\mathbf{S4}_u} (4.1)$ *implies* $\vdash_{\mathbf{S4}_u} (4.2)$.

Proof Ad absurdum, suppose that not (4.2):

$$\neg(U(\Box p \vee \Box q) \wedge Ep \wedge Eq \rightarrow E(p \wedge q))$$

Substituting the propositional variable q by $\neg p$, one obtains

$$\begin{aligned} &\neg(U(\Box p \vee \Box\neg p) \wedge Ep \wedge E\neg p \rightarrow E(p \wedge \neg p)) \\ &\neg(U(\Box p \vee \neg\Diamond p) \wedge \neg(\neg Ep \vee \neg E\neg p) \rightarrow E\perp) \\ &\quad U(\neg\Diamond p \vee \Box p) \wedge \neg(\neg Ep \vee \neg E\neg p) \\ &\quad U(\Diamond p \rightarrow \Box p) \wedge \neg(U\neg p \vee Up) \\ &\quad \neg(\neg U(\Diamond p \rightarrow \Box p) \vee (U\neg p \vee Up)) \\ &\quad \neg(U(\Diamond p \rightarrow \Box p) \rightarrow Up \vee U\neg p) \end{aligned}$$

Thus, contradicting the hypothesis $\vdash_{\mathbf{S4}_u}$ (4.1). QED

Results like this can be used for a systematic analysis of well-known topological preservation phenomena. But a more striking example, one that builds on topobisimulations and gives a semantic proof of the topological fact, is the following corollary of Theorem 4.1.3.

4.1.8. COROLLARY (CONNECTEDNESS). Consider $\langle X, O \rangle$ and $\langle X', O' \rangle$, and a continuous surjective map $f : X \rightarrow X'$. If the topological space $\langle X, O \rangle$ is connected, then the space $\langle X', O' \rangle$ is connected.

Proof Our first observation is a modal definition for connectedness, in the extended modal language $\mathbf{S4}_u$. We say that a topological space $\langle X, O \rangle$ *validates* a modal formula φ if φ is true at every point under every valuation. Now we have that the following two statements are equivalent:

- (i) $\langle X, O \rangle$ is connected
- (ii) $\langle X, O \rangle \models U(\diamond p \rightarrow \Box p) \rightarrow Up \vee U\neg p$

To see this, note that the antecedent of this extended formula holds if the denomination of p is both open and closed, while the consequent says that either $p = X$ or $p = \emptyset$.

Now, we return to the statement of the Corollary. We must show that $\langle X', O' \rangle$ is connected. Suppose that it is not. Then there exists a valuation ν' and a point x' such that $\langle X', O', \nu' \rangle, x' \models \neg(\text{ii})$. Next, we use the given continuous map f to define a simulation \leftarrow from M' to M (note the reversal in direction here):

$$x \leftarrow x' \quad \text{iff} \quad x' = f(x)$$

In particular, the definition of continuous map gives the forward simulation clause. Moreover, the surjectiveness of f guarantees that \leftarrow is surjective and total on M' . Next, we define a valuation ν on M by ‘copying ν' along f ’:

$$\nu(p) = f^{-1}(\nu'(p))$$

The result is a simulation \leftarrow from $\langle X', O', \nu' \rangle$ onto $\langle X, O, \nu \rangle$ such that $x \leftarrow x'$ for some point $x \in X$.

Finally, we note that the negated formulas $\neg(\text{ii})$ is logically equivalent (by some syntactic manipulation) to the $\mathbf{S4}_u$ formula without \diamond

$$U(\Box\neg p \vee \Box p) \wedge E\neg p \vee Ep$$

By Corollary 4.1.3, this formula also holds for X in M , and hence $\langle X, O \rangle$ is not connected. A contradiction. QED

The above is another piece of evidence for the claimed usefulness of bisimulations.

4.2 Alternative extensions

The universal extension is not the possibility to enhance the logical power. Here, we present some alternatives.

4.2.1 Hybrid reference

Another way to enhance the logical power of the basic topological language is to consider hybrid modal references, cf. [Arecas, 2000]. The idea of hybrid logics is that of naming worlds in Kripke structures enabling explicit reference and naming power at the language level. The resulting languages are very expressive as one can jump ‘quite freely’ from one world to another remembering names of places to visit or visited. A similar approach is also feasible for spatial logics interpreted on topological spaces. One simply gives a name, say r , to a region and directly refers to it at the level of the language with an appropriate modal operator “at region r .”

4.2.1. DEFINITION (SYNTAX OF $\mathbf{S4}_@$). Let $\langle X, O \rangle$ be a topological space, and let $P = \{p_1, p_2, \dots\}$ be a countable set of proposition letters (i.e., region names). The well-formed formulas of the language $\mathbf{S4}_@$ in the signature $\langle X, O, P \rangle$ are

$$F = \top \mid p \mid \neg\varphi \mid \varphi \wedge \psi \mid \Box\varphi \mid @_A\varphi$$

where $p, A \in P$ and $\varphi, \psi \in F$.

4.2.2. DEFINITION (TOPOLOGICAL SEMANTICS OF $\mathbf{S4}_@$). A topological model $M = \langle X, O, \nu \rangle$ is defined as for $\mathbf{S4}$. The interpretation is as for $\mathbf{S4}$ with the addition of the following definitions:

$$\begin{aligned} M, x \models @_A\varphi & \text{ iff } \forall y \in \nu(A) M, y \models \varphi \\ M, x \models @_a\varphi & \text{ iff } \exists y \in \nu(A) M, y \models \varphi \end{aligned}$$

One can look at the @ operator in two ways. On the one hand, it is a restricted version of the universal modality.

$$@_A\varphi \leftrightarrow U(A \rightarrow \varphi)$$

On the other hand, the operator resembles closely that of hybrid logics, though in $\mathbf{S4}_@$ there is no use of different sorts for propositional variables, nominals and states. Different sorts could be used, for instance, if considering special points with unique names. One would end up with the full topology of spatial regions and with names for some particular witnessing points. (Think for example of the topology of Europe, giving unique names to a certain number of distinguished points: the capitals of European countries.) Exploiting nominals would provide for extra expressive power. Most notably irreflexivity, which is not expressible in ordinary modal logics, can be expressed in hybrid systems, [Gabbay, 1981].

What one gets for the logic $\mathbf{S4}_@$ is a behavior that is a mixture of the universal modality and of the hybrid @ operator.

$$@_A\varphi \leftrightarrow \neg@_a\neg\varphi \quad (\text{Dual}_@)$$

Purely topological behaviors are:

$$\begin{aligned} @_aB &\leftrightarrow @_bA && (\text{Intersection}) \\ @_AB \wedge @_BC &\rightarrow @_AC && (4_@) \end{aligned}$$

Some hybrid behaviors are retained:

$$\begin{aligned} @_A(\varphi \rightarrow \psi) &\rightarrow (@_A\varphi \rightarrow @_A\psi) && (\mathbf{K}_@) \\ A \wedge @_A\varphi &\rightarrow \varphi && (\mathbf{T}_@) \\ @_AA &&& (\text{Label}) \\ @_A@_B\varphi &\rightarrow @_B\varphi && (\text{Scope}) \end{aligned}$$

As with $\mathbf{S4}_u$ there is also some purely topological power in the language. For instance, one can express the regularity of a region:

$$@_A\Diamond\Box A \quad \text{regularity of the region } A$$

Though, other global topological properties fall beyond the power of $\mathbf{S4}_@$. For instance, the property of a topological space to be connected or not, which is expressible in $\mathbf{S4}_u$ by (1), is not expressible in terms of $\mathbf{S4}_@$. To show this fact we need the basic tool of our topo-approach: topo-bisimulations. Here is the adequate notion for the hybrid language.

4.2.3. DEFINITION (TOPOLOGICAL BISIMULATION). Consider the language $\mathbf{S4}$ and two topological models $\langle X, O, \nu \rangle, \langle X', O', \nu' \rangle$. A *topological bisimulation* is a non-empty relation $\rightleftharpoons \subseteq X \times X'$ such that:

- (i) $\forall p \in P \forall x \in \nu(p) \exists x' \in \nu'(p)$ such that $x \rightleftharpoons x'$
- (ii) $\forall p \in P \forall x' \in \nu'(p) \exists x \in \nu(p)$ such that $x \rightleftharpoons x'$
- (iii) (*forth condition*): $x \in o \in O \Rightarrow \exists o' \in O' : x' \in o'$ and $\forall y' \in o' : \exists y \in o : y \rightleftharpoons y'$
- (iv) (*back condition*): $x' \in o' \in O' \Rightarrow \exists o \in O : x \in o$ and $\forall y \in o : \exists y' \in o' : y \rightleftharpoons y'$

Now consider the two topological models formed one by the real interval $(0, 1)$, the other by the interval $(0, 2) - \{1\}$, both with the following valuation function:

$$\nu(x) = x \bmod 2$$

The topological space underlying the first model is connected, while the second is not as it is the union of two open sets. Now consider the following bisimulation defined for all $x \in (0, 2)$:

$$x \rightleftharpoons x' \quad \text{iff} \quad \nu(x) = x \bmod 2$$

The $\mathbf{S4}_u$ formula of Equation (4.1) is true on the first model, but not on the second one. Therefore a total topo-bisimulation cannot be established, which implies that $\mathbf{S4}_u$ can distinguish between connected and non-connected spaces. If now one considers the relation for all $x \in (0, 2) - \{1\}$:

$$x \rightleftharpoons x' \quad \text{iff} \quad \nu(x) = x \bmod 2$$

it is easy to see that it is a bisimulation as defined in Definition 4.2.3. Therefore, connectedness is not expressible by means of $\mathbf{S4}_@$.

4.2.2 Until a boundary

Another source of inspiration for extension of the expressive power of the basic language of topology comes from temporal formalisms. Consider the Since and Until logic of [Kamp, 1968]. If one abstracts from the temporal behavior and interprets the modality in spaces with dimensionality greater than one, one gets an operator expressing something to be valid *up to a certain boundary* region, a sort of fence surrounding the current region. Here is a natural notion of spatial ‘Until’ in topological models:

$$M, x \models \varphi \mathcal{U} \psi \quad \text{iff} \quad \exists A : O(A) \wedge x \in A \wedge \forall y \in A. \varphi(y) \wedge \forall z (z \text{ is on the boundary of } A \wedge \psi(z))$$

Defining the dual modality $\varphi \mathcal{U}^D \psi$ as usual is $\neg(\neg\varphi \mathcal{U} \neg\psi)$ we get:

$$M, x \models \varphi \mathcal{U}^D \psi \quad \text{iff} \quad \forall A : O(A) \wedge x \in A \rightarrow (\exists y \in A. \varphi(y) \vee \exists z (z \text{ is on the boundary of } A \wedge \psi(z)))$$

Using the notation of the basic modal language, we recall the topological definition of boundary of a set A :

$$\text{boundary}(A) = \diamond A \wedge \diamond \neg A$$

A graphical representation of the Until operator is presented in Figure 4.2. Its expressiveness is richer than that of the basic modal language of space. E.g., one can express global properties inside connected components:

$$\mathcal{U}\varphi \perp \text{ iff some open component around the current point is all } \varphi$$

In connected spaces, this is equivalent to the universal modality U .

Which temporal principles valid in \mathbb{R} survive the move to more than one dimension? We do not provide a full axiomatization, but rather look at how temporal axioms

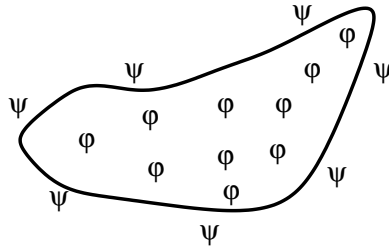


Figure 4.2: The region involved in $\varphi\mathcal{U}\psi$.

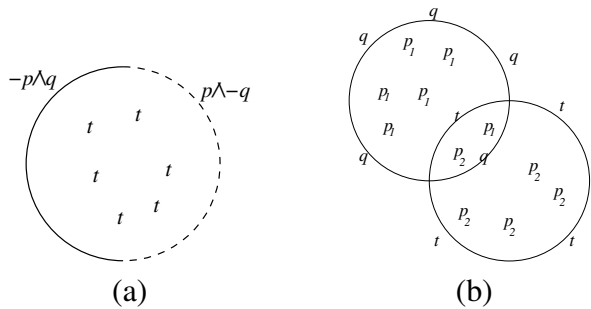


Figure 4.3: Examples of Until models.

behave in space and which new ones may arise. Two useful equivalences for obtaining normal forms in the one dimensional case are

$$\begin{aligned} t\mathcal{U}(p \vee q) &\leftrightarrow (t\mathcal{U}p) \vee (t\mathcal{U}q) \\ (p \wedge q)\mathcal{U}t &\leftrightarrow (p\mathcal{U}t) \wedge (q\mathcal{U}t) \end{aligned}$$

In our spatial setting, the first equivalence fails: Figure 4.3.a refutes the \rightarrow implication. But the other direction remains a valid principle of monotonicity. As for the second equivalence, its direction \rightarrow is a general monotonicity principle again. Conversely, we get even have a stronger valid law:

$$p_1\mathcal{U}q \wedge p_2\mathcal{U}t \rightarrow (p_1 \wedge p_2)\mathcal{U}(q \vee t)$$

Proof Let O, O' be the two open sets such that p_1 is true everywhere inside O and p_2 everywhere in O' , q is true on the boundary of O and t on the boundary of O' . Now consider the set $O \cap O'$. In such a set $p_1 \wedge p_2$ is true everywhere. In addition, every boundary point x of $O \cap O'$ is either a boundary point of O or of O' . In fact, consider a boundary point x of $(O \cap O')$, then $x \in \diamond(O \cap O')$ and $x \notin \square(O \cap O')$. Since $x \notin \square(O \cap O')$, $x \notin (O \cap O')$, as $O \cap O'$ is open. Say $x \notin O$. Then $x \notin \square O$, while also $x \in \diamond O$ (as $x \in \diamond(O \cap O')$), that is, x is a boundary point of O . See Figure 4.3.b for an illustration. Thus, our x must satisfy $q \vee t$. QED

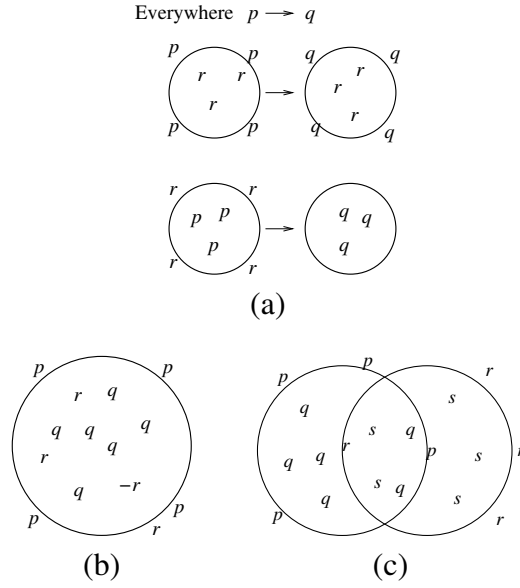


Figure 4.4: More examples of Until models.

Burgess [1984] reviews basic tense logic providing, among other things, an axiomatization of the Since and Until logic for total dense orders. Departing from these axioms, we consider their spatial validity. First, we let us define an abbreviation G :

$$Gp \leftrightarrow p\mathcal{U}\perp$$

Here is the set of axioms:

$$G(p \rightarrow q) \rightarrow ((r\mathcal{U}p) \rightarrow (r\mathcal{U}q)) \wedge ((p\mathcal{U}r) \rightarrow (q\mathcal{U}r)) \quad (4.3)$$

$$p \wedge (r\mathcal{U}q) \rightarrow (r\mathcal{U}(q \wedge (r\mathcal{S}p))) \quad (4.4)$$

$$(q\mathcal{U}p) \leftrightarrow ((q \wedge (q\mathcal{U}p))\mathcal{U}p) \leftrightarrow q\mathcal{U}(q \wedge (q\mathcal{U}p)) \quad (4.5)$$

$$((q\mathcal{U}p) \wedge \neg(r\mathcal{U}p)) \rightarrow q\mathcal{U}(p \wedge \neg r) \quad (4.6)$$

$$((q\mathcal{U}p) \wedge (s\mathcal{U}r)) \rightarrow (((q \wedge s)\mathcal{U}(p \wedge r)) \vee ((q \wedge s)\mathcal{U}(p \wedge s)) \vee ((q \wedge s)\mathcal{U}(q \wedge r))) \quad (4.7)$$

For now, this serves as an illustration of ‘transfer’ of temporal logic principles to spatial settings. Finally, as for topo-bisimulations for this richer language, we would need an extension of the proposals in [Kurtonina and de Rijke, 1997] for dealing with the $\exists\forall$ -complexity of the truth condition for the spatial Until.

Axiom 4.3 is valid for the spatial Since and Until. If everywhere $G p$ implies q , then it must be the case that if the region r has a p boundary then it also has a q boundary. Similarly, if a p region has a r boundary, so does the q region defined by the same p points, cf. Figure 4.4.a. Axiom 4.4 does not make sense in the spatial setting where there is no notion of past and, therefore, no Since operator. Axiom 4.5 expresses

some kind of density and is valid in the spatial version if the model is dense, as it is trivial to show by contraposition. Axiom 4.6 and Axiom 4.7 do *not* hold in spaces with more than one dimensions. Here are two simple counter-examples, respectively for Axiom 4.6 and Axiom 4.7, cf. Figure 4.4.b,c. Consider an open set A made of q points with a boundary of p points. Inside A consider a number of isolated $\neg r$ points, while outside the open A there are only r points. It is easy to see that inside A the left hand side of the implication of Axiom 4.6 is satisfied, while the right hand side is not. In fact it is impossible to find an open set all made of q points with a continuous boundary of $\neg r$ points. A counter-example to Axiom 4.7 is also easy to build. Consider two open circles of the same radius but different centers. Circle A is made of q points, its circumference is made of p points, everywhere else it is $\neg p$. Circle B is built similarly by replacing q by s and p by r . The circle A and B overlap. It is easy to check that there does not exist an open set made of $q \wedge s$ points whose boundary is made of exclusively $p \wedge r$ points nor only of $p \wedge s$ points and also not only of $p \wedge r$ points. At most one can hope for a weaker version of the axiom valid in the temporal case (Axiom 4.7):

$$((q\mathcal{U}p) \wedge (s\mathcal{U}r)) \rightarrow ((q \wedge s)\mathcal{U}((p \wedge r) \vee (p \wedge s) \vee (q \wedge r))).$$

Proving soundness for the spatial version of Since and Until has shown a fundamental difference with the temporal version. The reasoning does not involve trees, but full fledged topological spaces.

4.3 Standard logical analysis

The modal hierarchy of topological languages has a common root. All operators given have truth conditions in a second-order language quantifying over both points and sets of points. E.g., $\Box p$ says that $\exists A : O(A) \wedge x \in A \wedge \forall y : y \in A \rightarrow P(y)$. This language has the following vocabulary:

$\forall x$	quantification over points
$\forall A$	quantification over sets of points
$x = y$	identity
$x \in A$	membership of points in sets
$O(A)$	predicate of openness of sets

All fundamental topological notions are definable in this formalism. Here are two relevant observations.

4.3.1. FACT. *Formulas of the second-order language without free predicate variables are preserved under topological homeomorphisms.*

The proof is a simple induction.

4.3.2. FACT. *All topological separation axioms T_i (with $0 \leq i \leq 4$) are expressible in the second-order language.*

For example, one can express the T_2 axiom (defining the Hausdorff spaces) in the following way:

$$\forall x, y : (x \neq y \rightarrow \exists A, B : O(A) \wedge O(B) \wedge \neg \exists z (z \in A \wedge z \in B) \wedge x \in A \wedge y \in B)$$

Similarly we can write the definition of the axiom for T_4 spaces:

$$\begin{aligned} & \forall C, D : (O(\neg C) \wedge O(\neg D) \wedge \neg \exists z (z \in C \wedge z \in D)) \\ & \exists A, B : O(A) \wedge O(B) \wedge \neg \exists z (z \in A \wedge z \in B) \wedge \forall x \in C \ x \in A \wedge \forall x \in D \ x \in B \end{aligned}$$

Of course, this strong language has various much more tractable fragments, and the goal in ‘modal topology’ is finding these. But the second-orderness in this analysis maybe somewhat spurious. One can see this by the ‘deconstruction’ of Section 2.3.

5.1 Affine Geometry

Extending the expressive power of a modal logic of space may go beyond mere logical power, cf. Chapter 4. One can also enrich *geometrical power* by endowing spaces with more structure. A first elementary example is the property of a point's being in the convex closure of a set of points. That is, there exists a segment containing the points whose end-points are in the set. The notion of convexity is very important in many fields related to space (e.g., computational geometry [Preparata and Shamos, 1985]), but also in abstract cognitive settings (e.g., conceptual spaces [Gärdenfors, 2000]). Capturing convexity modally involves a standard similarity type, that of frames of points with a ternary relation of betweenness:

$$M, x \models C\varphi \text{ iff } \exists y, z : M, y \models \varphi \wedge M, z \models \varphi \wedge x \text{ lies in between } y \text{ and } z \quad (5.1)$$

This definition is slightly different from the usual notion of convex closure. It is a *one-step convexity* operator whose countable iteration yields the standard convex closure. The difference between the two definitions is visible in Figure 5.9. On the left are three points denoting a region. The standard convex closure operator gives the full triangle depicted on the right. The one-step convexity, on the other hand, gives the frame of the triangle and only when applied twice yields the full triangle. Another illustration is presented in Figure 5.1. One-step convexity exhibits a modal pattern for an existential binary modality:

$$\exists yz : \beta(yxz) \wedge \varphi(y) \wedge \varphi(z)$$

From now on, we shall use the term convexity operator to refer to the one-step convexity operator defined in (5.1).

5.1.1 Basic geometry

Geometrical modal logic starts from standard bits of mathematics, viz. affine geometry, [Blumenthal, 1961]. For later reference, here are the affine base axioms in a language

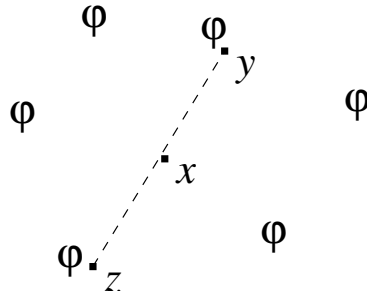


Figure 5.1: The point x is in the one-step convex closure φ .

with two sorts for points and lines, and an incidence relation as presented by Goldblatt [Goldblatt, 1987]:

A1 Any two distinct points lie on exactly one line.

A2 There exist at least three non-collinear points.

A3 Given a point a and a line L , there is exactly one line M that passes through a and is parallel to L .

There are also some properties that further classify affine planes. In particular, an affine plane is *Pappian* if every pair of its lines has the Pappus property:

A pair L, M of lines in an affine plane has the Pappus property if whenever a, b, c is a triple of points on L , and a', b', c' is a triple on M such that ab' is parallel to $a'b$ and ac' is parallel to $a'c$, then $b'c$ is parallel to bc' .

Affine spaces have a strong modal flavor, as shown by [Balbiani et al., 1997, Balbiani, 1998, Venema, 1999, Stebletsova, 2000], where two roads are taken. One merges points and lines into one sort of pairs $\langle \text{point}, \text{line} \rangle$ equipped with two incidence relations. The other has two sorts for points and lines, and a matching modal operator.

But there are more expressive classical approaches to affine structure. Tarski [1959] gave a full first-order axiomatization of elementary geometry in terms of a ternary betweenness predicate β and quaternary equidistance δ . We display it as a kind of ‘upper limit’:

A1 $\forall xy(\beta(xyx) \rightarrow (x = y))$, identity axiom for betweenness.

A2 $\forall xyzu(\beta((xyu) \wedge \beta(yzu)) \rightarrow \beta(xyz))$, transitivity axiom for betweenness,

A3 $\forall xyzu(\beta(xyz) \wedge \beta(xyu) \wedge (x \neq y) \rightarrow \beta(xzu) \vee \beta(xuz))$ connectivity axiom for betweenness,

A4 $\forall xy(\delta(xyyx))$, reflexivity axiom for equidistance,

- A5** $\forall xyz(\delta(xyz) \rightarrow (x = y))$, identity axiom for equidistance,
- A6** $\forall xyzuvw(\delta(xyzu) \wedge \delta(xyvw) \rightarrow \delta(zuvw))$, transitivity axiom for equidistance,
- A7** $\forall txyzu\exists v(\beta(xtu) \wedge \beta(yuz) \rightarrow \beta(xvy) \wedge \beta(ztv))$, Pasch’s axiom,
- A8** $\forall txyzu\exists vw(\beta(xut) \wedge \beta(yuz) \wedge (x \neq u) \rightarrow \beta(xzv) \wedge \beta(xyw) \wedge \beta(vtw))$, Euclid’s axiom,
- A9** $\forall xx'y'y'zz'u'u'(\delta(xyx'y') \wedge \delta(yzy'z') \wedge \delta(xux'u') \wedge \delta(yuy'u') \wedge \beta(xyz) \wedge \beta(x'y'z') \wedge (x \neq y) \rightarrow \delta(zuz'u'))$, five-segment axiom,
- A10** $\forall xyuv\exists z(\beta(xyz) \wedge \delta(yzuv))$, axiom of segment construction,
- A11** $\forall xyz(\neg\beta(xyz) \wedge \neg\beta(yzx) \wedge \neg\beta(zxy))$, lower dimension axiom,
- A12** $\forall xyzuv(\delta(xuxv) \wedge \delta(yuyv) \wedge \delta(zuzv) \wedge (u \neq v) \rightarrow \beta(xyz) \vee \beta(yzx) \vee \beta(zxy))$, upper dimension axiom,
- A13** All sentences of the form $\forall vw \dots (\exists z\forall xy(\psi \wedge \varphi \rightarrow \beta(zxy)) \rightarrow \exists u\forall xy(\psi \wedge \varphi \rightarrow \beta(xuy)))$, elementary continuity axioms.

Why is this beautiful complete and decidable axiomatization not all one wants to know? From a modal standpoint, there are two infelicities in this system. The axioms are too powerful, and one wants to look at more tractable fragments. But also, the axioms mix betweenness and equidistance—whereas one first wants to understand affine and metric structure separately.

5.1.2 The general logic of betweenness

Our choice of primitives for affine space is again betweenness, where $\beta(xyz)$ means that point y lies in between x and z , allowing y to be one of these end-points. Line structure is immediately available by defining *collinearity* in terms of betweenness:

$$xyz \text{ are collinear iff } \beta(xyz) \vee \beta(yzx) \vee \beta(zxy)$$

‘Geometrical extensions’ of this sort can even define ‘extended modalities’, i.e., ‘logical extensions’ in our earlier terminology. Here is the *existential* “at some point:”

$$E\varphi \text{ iff } \langle B \rangle(\varphi, \top) \tag{5.2}$$

This will work provided we require betweenness to satisfy:

$$\forall x\forall y\beta(xxy).$$

Without this, the defined modality will just range over the connected component of the current point of evaluation.

Natural specific structures on which to interpret our modal language include the \mathbb{R}^n for any n . But affine spaces really form a much more general class of structures. What are natural general frame conditions constraining these? As one does for temporal logics, the *universal first-order theory* of ordinary real space suggests good candidates. Consider just the betweenness part of Tarski's elementary geometry. Axioms **A1-A3** for identity, transitivity, and linearity are all plausible as general affine properties. They are not sufficient, though, as one also wants some obvious variants of transitivity and linearity with points in other positions stated explicitly. With Tarski, the latter are theorems, but their proofs go through other axioms involving equidistance. Further universal first-order assertions that hold in real space would express *dimensionality* of the space, which does not seem a plausible constraint in general.

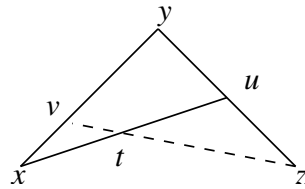


Figure 5.2: Pasch's property.

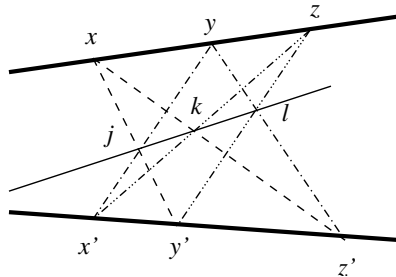


Figure 5.3: Pappus property.

At the next level of syntactic complexity, one then finds existential axioms and universal-existential ones, which require the space to have a certain richness in points. The latter expresses typical geometrical behavior, witness *Pasch's axiom A7* (see Figure 5.2) and the earlier Pappus property (see Figure 5.3):

$$\forall x x' y y' z z' \exists j k l \beta(x y z) \wedge \beta(x' y' z') \wedge \beta(x j y') \wedge \beta(y j x') \wedge \beta(x k z') \wedge \beta(z k x') \wedge \beta(y l z') \wedge \beta(z l y') \rightarrow \beta(j k l)$$

Moving to the opposite extreme of geometrical structure, consider the real line \mathbb{R} . Its universal first-order theory includes the strong dimensionality principle

$$\forall x y z, \beta(x y z) \vee \beta(y x z) \vee \beta(x z y) \tag{5.3}$$

The complete affine first-order theory here can be axiomatized very simply, by translating $\beta(xyz)$ as $y = x \vee y = z \vee x < y < z$. This reduces the one-dimensional geometry to the decidable theory of discrete unbounded linear orders. But it would be of interest to also axiomatize the universal first-order betweenness theories of the spaces \mathbb{R}^n explicitly.

5.1.3 Modal languages of betweenness

Let us now turn to modal logic over affine spaces.

5.1.3.1 The basic language

Ternary betweenness models a binary betweenness modality $\langle B \rangle$:

$$M, x \models \langle B \rangle(\varphi, \psi) \quad \text{iff} \quad \exists y, z : \beta(yxz) \wedge M, y \models \varphi \wedge M, z \models \psi$$

Note that this is a more standard modal notion than the earlier topological modality: we are working on frames, and there are no two-step quantifiers hidden in the semantics. $\langle B \rangle$ is expressive. For instance, it defines one-step convex closure as follows:

$$\text{convex}(\varphi) \text{ iff } \langle B \rangle(\varphi, \varphi) \tag{5.4}$$

Passing to points ‘in between’ two others yields the convex closure only after repeated applications of this operator, as shown in Figure 5.9. In a more elaborate set-up, we could take a leaf from dynamic logic, and add an operation of *Kleene iteration* of the betweenness predicate—much as ternary ‘composition’ is iterated in dynamic Arrow Logic (cf. Chapter 8 in [van Benthem, 1996]). Next, the existential modality has a dual universal version: $[B](\varphi, \psi) \leftrightarrow \neg \langle B \rangle(\neg\varphi, \neg\psi)$, which works out to

$$M, x \models [B](\varphi, \psi) \quad \text{iff} \quad \forall y, z : \beta(yxz) \rightarrow M, y \models \varphi \vee M, z \models \psi$$

An implicational variant of this definition is also helpful sometimes:

$$M, x \models [B](\neg\varphi, \psi) \quad \text{iff} \quad \forall y, z : \beta(yxz) \wedge M, y \models \varphi \rightarrow M, z \models \psi$$

One might think that there should be an independent conjunctive variant, saying that both end-points have their property. But this is already definable—another sign of the strength of the language:

$$[B](\varphi, \perp) \wedge [B](\perp, \psi)$$

5.1.3.2 Versatile extensions

Betweenness is natural, but biased toward 'interior positions' of a segment. But given two points x and y , one can also consider all points z such that x lies in between y and z , or all w such that y lies in between x and w . In this way, two points identify a direction and a weak notion of orientation. There are two obvious further existential modalities corresponding to this. Together with $\langle B \rangle$, they form a 'versatile' triple in the sense of [Venema, 1992]. Such triples are often easier to axiomatize together than in isolation. As an illustration, consider the table of Figure 5.4, which we have been

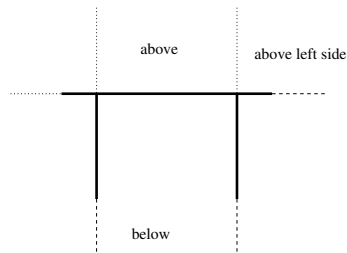


Figure 5.4: A table and the regions for versatile betweenness modalities.

setting in earlier sections. Using versatile modalities, the legs of the table and its top identify important zones of visual scenes, which also have names in natural language, such as everything 'above the table'.

5.1.3.3 Affine transformations

Affine transformations are the invariant maps for affine geometry. Their modal counterpart are *affine bisimulations* which are mappings relating points verifying the same proposition letters, and maintaining betweenness. We only display the definition for our original 'interior' betweenness—since the versatile extensions are straightforward:

5.1.1. DEFINITION (AFFINE BISIMULATION). Given two affine models $\langle X, O, \beta, \nu \rangle$, and $\langle X', O', \beta', \nu' \rangle$, an *affine bisimulation* is a non-empty relation $\rightleftharpoons \subseteq X \times X'$ such that, if $x \rightleftharpoons x'$:

- (i) x and x' satisfy the same proposition letters,
- (ii) (*forth condition*): $\beta(yxz) \Rightarrow \exists y'z' : \beta'(y'x'z')$ and $y \rightleftharpoons y'$ and $z \rightleftharpoons z'$
- (iii) (*back condition*): $\beta'(y'x'z') \Rightarrow \exists yz : \beta(yxz)$ and $y \rightleftharpoons y'$ and $z \rightleftharpoons z'$

where $x, y, z \in X$ and $x', y', z' \in X'$.

In [Goldblatt, 1987], isomorphisms are considered the only interesting maps across affine models. But in fact, just as with topological bisimulations versus homeomorphisms (Theorem 2.1.5), affine bisimulations are interesting coarser ways of comparing spatial situations. In the true modal spirit, they only consider the behavior of points

inside their local line environments. Consider the two models consisting of 6 and 4 points, respectively, on and inside two triangles, with some atomic properties indicated, Figure 5.5. The models are evidently not isomorphic, but there is an affine bisimula-

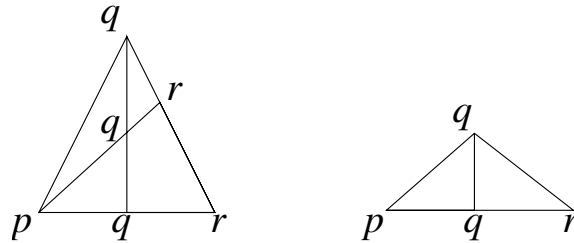


Figure 5.5: Affine bisimilar models.

tion. Simply relate the two r points on the left with the single r point on the right. Then relate the top q point on the left with the top one on the right, the remaining two q points on the left with the one on the right, and, finally, the p point on the left with the one of the right. This affine bisimulation can be regarded as a sort of ‘modal contraction’ to a smallest bisimilar model, as we did in Section 2.1.3. The models in Figure 5.6 are not bisimilar though. One can check that no relation does the job—or, more simply, note

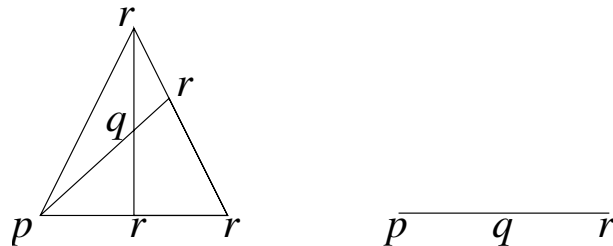


Figure 5.6: Affine bisimilar reduction.

that the modal formula $q \wedge \langle B \rangle (r, r)$ holds on the q point of the left model and nowhere on the right. Affine bisimulations preserve truth of modal formulas in an obvious way, and hence they are a coarser map than isomorphisms still giving meaningful geometrical invariances. This is exactly as we found with topological bisimulations versus homeomorphisms.

Incidentally, notice that there *is* a smaller bisimulation contraction for the left-hand triangle. The reason is that not all its points are uniquely definable in our modal language. The p and q points are uniquely definable, but all r points on the boundary satisfy the same modal statements. The contraction will look like the picture to the right, but with the middle point ‘in between’ the right point and the right point itself. (This is not a standard 2D ‘picture’, and duplicating points cannot always be contracted

if we insist on those.) This situation would change with a modality for *proper betweenness*. Then the two middle r points become uniquely distinguishable as being properly in between different pairs of points. But the top and right-bottom point remain indistinguishable, unless we add versatile operators. It is a nice exercise to show that the triangle does have every point uniquely definable in the original language when we change the atomic proposition in the top vertex and the one center bottom to q and that in the middle of the right edge to p . Consider the new valuation in Figure 5.7. In

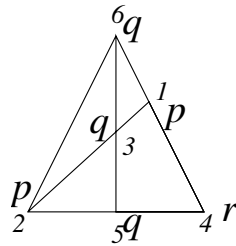


Figure 5.7: An irreducible affine model.

this case there does *not* exist a bisimilar contraction. Every point of the triangle is distinguishable by a formula which is not true on any other point, see Figure 5.8. This

Point	Formula
1	$\varphi_1 = p \wedge \langle B \rangle(q, r)$
2	$\varphi_2 = p \wedge \neg\varphi_1$
3	$\varphi_3 = q \wedge \langle B \rangle(\varphi_1, \varphi_2)$
4	$\varphi_4 = r$
5	$\varphi_5 = q \wedge \langle B \rangle(\varphi_2, \varphi_4)$
6	$\varphi_6 = q \wedge \neg\varphi_3 \wedge \neg\varphi_5$

Figure 5.8: Formulas true at points of the model in Figure 5.7.

suggests a theory of unique patterns, depending on how points are labeled in geometrical pictures.

5.1.4 Modal logics of betweenness

The preceding language has a minimal logic as usual, which does not yet have much geometrical content. Its key axioms are two distribution laws:

$$\begin{aligned} \langle B \rangle(\varphi_1 \vee \varphi_2, \psi) &\leftrightarrow \langle B \rangle(\varphi_1, \psi) \vee \langle B \rangle(\varphi_2, \psi) \\ \langle B \rangle(\psi, \varphi_1 \vee \varphi_2) &\leftrightarrow \langle B \rangle(\psi, \varphi_1) \vee \langle B \rangle(\psi, \varphi_2) \end{aligned}$$

This minimal logic by itself has all the usual modal properties, decidability among them. Other basic principles express basic universal relational conditions, such as betweenness being symmetric in end-points, and all points lying ‘in between themselves’:

$$\begin{aligned} \langle B \rangle(\varphi, \psi) &\rightarrow \langle B \rangle(\psi, \varphi) \\ \varphi &\rightarrow \langle B \rangle(\varphi, \varphi) \end{aligned}$$

These facts are simple *frame correspondences* in the usual modal sense. A slightly more tricky example is the earlier-mentioned relational condition $\forall x \forall y \beta(xxy)$. This is not definable as it stands, but the modal axiom

$$(\varphi \wedge \langle B \rangle(\top, \psi)) \rightarrow \langle B \rangle(\varphi, \psi)$$

corresponds to the related principle

$$\forall x \forall y \forall z : \beta(zxy) \rightarrow \beta(xxy)$$

More generally, special modal axioms may correspond to more complex properties of geometric interest. For example, consider *associativity* of the betweenness modality:

$$\langle B \rangle(\varphi, \langle B \rangle(\psi, \xi)) \leftrightarrow \langle B \rangle(\langle B \rangle(\varphi, \psi), \xi)$$

5.1.2. FACT. *Associativity corresponds to the Pasch Axiom.*

Proof Consider the Pasch Axiom **A7** in Tarski’s list (Figure 5.2). Suppose that

$$\forall txyz \exists v(\beta(xtu) \wedge \beta(yuz) \rightarrow \beta(xvy) \wedge \beta(ztv))$$

holds in a frame. Assume that a point t satisfies $\langle B \rangle(\varphi, \langle B \rangle(\psi, \xi))$. Then there exist points x, u with $\beta(xtu)$ such that $x \models \varphi$, $u \models \langle B \rangle(\psi, \xi)$, and hence also points y, z with $\beta(yuz)$ such that $y \models \psi$ and $z \models \xi$. Now by Pasch’s Axiom, there must be a point v with $\beta(xvy)$ and $\beta(vtz)$. Now, $v \models \langle B \rangle(\varphi, \psi)$ and hence $t \models \langle B \rangle(\langle B \rangle(\varphi, \psi), \xi)$. The other direction is similar.

Conversely, assume that $\beta(xtu)$ and $\beta(yuz)$. Define a valuation on the space by setting $\nu(p) = \{x\}$, $\nu(q) = \{y\}$, and $\nu(r) = \{z\}$. Thus, $u \models \langle B \rangle(q, r)$ and

$$t \models \langle B \rangle(p, \langle B \rangle(q, r)).$$

By the validity of modal associativity, then

$$t \models \langle B \rangle(\langle B \rangle(p, q), r)$$

So there must be points v, w with $\beta(vtw)$ such that $v \models \langle B \rangle(p, q)$ and $w \models r$. By the definition of ν , the latter means that $w=z$, the former that $\beta(xvy)$. So indeed, u is the required point. QED

The preceding correspondence may be *computed automatically*, as the associativity has ‘Sahlqvist form’. Thus, more general substitution methods apply for finding geometrical correspondents: cf. [Blackburn et al., 2001].

5.1.5 Special logics

For the affine modal logic of special models, additional considerations may apply. One example is the real line \mathbb{R} , which was also conspicuous in the topological setting. This time, the task is easy, as one can take advantage of the binary ordering $<$, defining

$$M, x \models \langle B \rangle(\varphi, \psi) \quad \text{iff} \quad \exists y, z : M, y \models \varphi \wedge M, z \models \psi \wedge z \leq x \leq y$$

Given this notion, we can use shorthand for the modalities of temporal logic: Future and Past (here, both including the present).

$$F\varphi := \langle B \rangle(\text{true}, \varphi)$$

$$P\varphi := \langle B \rangle(\varphi, \text{true})$$

Conversely, on \mathbb{R} , these two unary modalities suffice for defining $\langle B \rangle$:

$$\langle B \rangle(\varphi, \psi) \leftrightarrow P\varphi \wedge F\psi$$

Thus, a complete and decidable axiomatization for our $\langle B \rangle$ -language can be found using the well-known tense logic of future and past on \mathbb{R} [Seegerberg, 1970].

Special models also raise special issues. We have already seen the universal axiom Equation (5.3) defining one-dimensionality. What would be good versions for higher dimensions? We will address this issue once more in our next section.

5.1.6 Logics of convexity

A binary modality for a ternary frame relation gives maximal flexibility. Nevertheless, given the geometrical importance of convexity per se, here is a unary modal operator for one-step convex closure:

$$M, x \models C\varphi \quad \text{iff} \quad \exists y, z : M, y \models \varphi \wedge M, z \models \varphi \wedge x \in y-z$$

This is a fragment of the preceding modal language:

$$C\varphi \leftrightarrow \langle B \rangle(\varphi, \varphi).$$

The axiomatic behavior is different though: distributivity fails. Of the axiom

$$C(\varphi \vee \psi) \leftrightarrow C\varphi \vee C\psi$$

only the right-to-left *monotonicity* implication is valid. But the one-step convex closure of a set of two distinct points is their whole interval, while the union of their separate one-step closures is just these points themselves.

Earlier on, we already noted that one-step convex closure needs finite iteration to yield the usual convex closure of geometry. This could be brought out again in a

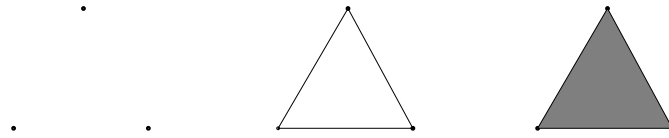


Figure 5.9: In a two dimensional space, the sequential application of the convexity operator to three non aligned points results in two different regions: a triangle (only the sides and corners of it) and the filled triangle.

language with an additional modality C^* , where the $*$ denotes Kleene iteration. This interesting spatial use of dynamic logic is not pursued here, for a reason to be explained below. First, note that the non-idempotence of C gives additional expressive power by itself. In fact, it helps us distinguish dimensions! Here is how. The principle

$$CC\varphi \leftrightarrow C\varphi$$

holds on \mathbb{R} , but not on \mathbb{R}^2 . A counter-example on \mathbb{R}^2 is shown in Figure 5.9. The region p is given by three non-collinear points. Cp is then the bare triangle: convexity has added the edges. Applying convexity again, CCp defines a different region, namely the whole triangle with its interior. One may be inclined to rush to the conclusion that principles of the form

$$C^{n+1}\varphi \leftrightarrow C^n\varphi \tag{5.5}$$

determine the dimensionality of the spaces \mathbb{R}^n for all n . But here is a surprise.

5.1.3. THEOREM. *The principle $CCC\varphi \leftrightarrow CC\varphi$ holds in \mathbb{R}^3 .*

Proof Here is a sketch. It will help the reader to visualize the situation using the tetrahedron example in Figure 5.11.

$C\varphi$ consists of all points in between two φ -points. $CC\varphi$ consists of all points in between the latter, and the implication $CC\varphi \rightarrow C\varphi$ corresponds (in the literal modal frame-theoretic sense) to the betweenness property that

$$(\beta(yxz) \wedge \beta(uyv) \wedge \beta(szt)) \rightarrow \bigwedge \{\beta(ixj) \mid i, j \in \{u, v, s, t\}\}$$

This is true in one dimension, though not in higher ones.

On the plane, $C\varphi$ consists of the same points. But we can give another description of $CC\varphi$. If x lies in between two $C\varphi$ -points, say on intervals $y-z$ and $u-v$, respectively, then x lies in/on one of the triangles yzu or yzv . Therefore, $CC\varphi$ -points lie on triangles of φ -points. Now consider any point r in $CCC\varphi$, i.e., between points s, t in/on such $CC\varphi$ triangles. Intersecting the segment $s-t$ with the two triangle boundaries, we get that r lies in a four sided polygon of φ -points, and hence, bisecting, r is already in/on a triangle of φ -points: i.e., r is in $CC\varphi$ already.

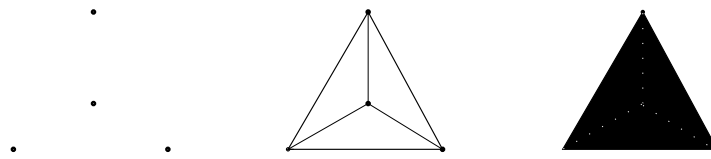


Figure 5.10: In a three dimensional space, the sequential application of the betweenness operator to four non coplanar points results in two distinct regions: the wire-frame of a tetrahedron and the filled tetrahedron.

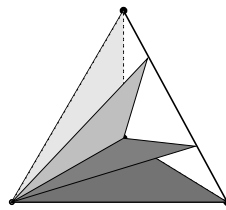


Figure 5.11: Applying convexity from the wire-frame to the full tetrahedron.

In 3D, the description for $CC\varphi$ is different, because the two segments for the $C\varphi$ -points need not lie in the same plane. The outcome is that these points lie in/on a 4-hedron of φ -points. Now consider a generic point r in $CCC\varphi$. It will lie in between points in such 4-hedra. This situation is easier to picture: take the segment on which it lies, and intersect that with the relevant faces of the 4-hedra. Then it is easy to see that the point r lies inside a 6-hedron whose vertices are φ -points. But then, cutting this up a number of times now, there is again a 4-hedron of φ -points in/on which we find r , hence, it is in $CC\varphi$ already. QED

As a corollary, for real spaces, we can then define convex closure in our language after all, using CC combinations. Hence, a full dynamic language, no matter how interesting, is not strictly needed. But for the moment we note the following fact.

5.1.4. FACT. *For any formula φ , C^n is a convex set in \mathbb{R}^n .*

But there are dimension highlighters in our language after all. An old theorem from almost a century ago [Helly, 1923] comes to the rescue:

5.1.5. THEOREM (HELLY). *If K_1, K_2, \dots, K_m are convex sets in n -dimensional Euclidean space E^n , in which $m > n + 1$, and if for every choice of $n + 1$ of the sets K_i there exists a point that belongs to all the chosen sets, then there exists a point that belongs to all the sets K_1, K_2, \dots, K_m .*

This theorem does have a modal version;

$$\bigwedge_{f:\{1,\dots,n+1\}\rightarrow\{1,\dots,m\}} E\left(\bigwedge_{i=1}^{n+1} (C^n \varphi_{f(i)})\right) \rightarrow E\left(\bigwedge_{i=1}^m C^n \varphi_i\right)$$

where E is the existential modality (defined in terms of betweenness in Equation 5.2), C^n is convexity applied n times (Fact 5.1.4), and f is a function from $\{1, \dots, n+1\}$ to $\{1, \dots, m\}$.

5.1.6.1 Digression: A proof in the projective plane

A convenient form of representing geometrical spaces is by homogeneous coordinates of the projective plane. In homogeneous coordinates (see for instance [Foley et al., 1990]), a point in an n -dimensional space is represented by $n+1$ elements of a vector. For example, the origin of the plane is represented by

$$P = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Representing points and sets of points in this form we look at convexity with two goals: one, to see the relation between bisimulations and convexity; two, to show formally that (5.5) does not hold.

The convexity operator applied to the regions formed by two generic points $P_1 = (x_1, y_1, 1)$, $P_2 = (x_2, y_2, 1)$ of the plane denotes the segment, with a slight mix of notation between the language level and model level,

$$C(P_1 \cup P_2) = \begin{pmatrix} x_1(1-c) + x_2c \\ y_1(1-c) + y_2c \\ 1 \end{pmatrix} \quad (5.6)$$

with $c \in [0, 1]$. In the following, we may abbreviate $C(P_1 \cup P_2)$ by $P_1(1-c) + P_2c$, even though it is not precise from an algebraic point of view.

5.1.6. LEMMA (AFFINE BISIMULATIONS). Affine transformations imply affine bisimulations.

Proof We sketch the proof for the two dimensional case. First a geometrical fact, affinely transforming a point P is represented, in homogeneous coordinates, by the pre-multiplication of a square matrix, $\bar{P} = T \cdot P$,

$$T = \begin{pmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ 0 & 0 & 1 \end{pmatrix} \quad (5.7)$$

where the upper 2×2 matrix is orthonormal, i.e., $T^{-1} = T^t$ and $|T| = 1$. For example, if $r_{11} = 1, r_{12} = 0, r_{21} = 0, r_{22} = 1$ one gets a translation, while if $r_{12} = 0, t_x = 0, r_{21} = 0, t_y = 0$ one gets a scaling.

First, we show that affine transformations imply bisimulations. Disregard the valuation function. A generic point $P_g \in P_1 - P_2$ is related, via an affine transformation,

to the point $T \cdot P_g$. We need to show that $T \cdot P_g \in T \cdot P_1 - T \cdot P_2$. We rewrite the last membership relation as $T \cdot P_g = T \cdot P_1(1 - c) + T \cdot P_2c$ with $c \in [0, 1]$. It is now a matter of simple matrix manipulation and substitution with Equation (5.6) and (5.7) to show that the latter equation holds. QED

Actually, there are reasons to suspect that the implication in the opposite direction holds in a vast number of cases. For instance, both affine bisimulations and affine transformations preserve convexity in a very similar manner.

5.1.7. FACT. $C\varphi$ does not necessarily denote a convex set in two or more dimensional spaces.

Proof We give a counter-example, see also Figure 5.12. Consider 3 points P_1, P_2, P_3 .

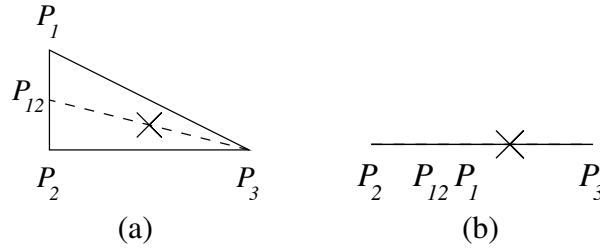


Figure 5.12: Convexity of a region made is not necessarily a convex region. In (a) it is not, while in (b) it is.

A point of $C\varphi$ is, for instance, $P_{12} = P_1(1 - \frac{1}{2}) + P_2\frac{1}{2}$. If we consider all the points between P_{12} and P_3 we see that they are in $C\varphi$ iff the three points are collinear:

$$\begin{pmatrix} \frac{1}{4}(x_1 + x_2 + 2x_3) \\ \frac{1}{4}(y_1 + y_2 + 2y_3) \\ \frac{1}{4}(z_1 + z_2 + 2z_3) \\ 1 \end{pmatrix} = \begin{pmatrix} x_1(1 - c) + x_2c \\ y_1(1 - c) + y_2c \\ z_1(1 - c) + z_2c \\ 1 \end{pmatrix}$$

has solutions in c iff the values of the $x_{\{1,2\}}, y_{\{1,2\}}, z_{\{1,2\}}$ are pairwise linearly dependent, i.e., iff the points P_1, P_2, P_3 are collinear (Figure 5.12.(b)). QED

Finally, we consider what happens applying the convexity operator one more time.

5.1.8. FACT. $CC\varphi$ denotes a convex set in a three-dimensional space.

Proof If we apply convexity twice, we obtain

$$CC\varphi = CC \bigcup_i \begin{pmatrix} x_i \\ y_i \\ z_i \\ 1 \end{pmatrix} = \{P_{cc} | \forall j, k, l, m, P_{cc} = (P_j(1 - c_1) + P_kc_1)(1 - c_3) + (P_l(1 - c_2) + P_mc_2)c_3\}$$

The expression of a point between any two generic points of the previous set P_{cc}, P'_{cc} is

$$P_g = ((P_j(1 - c_1) + P_k c_1)(1 - c_3) + (P_l(1 - c_2) + P_m c_2)c_3)(1 - c_7) + ((P_j(1 - c_4) + P_k c_4)(1 - c_6) + (P_l(1 - c_5) + P_m c_5)c_6)c_7 \quad (5.8)$$

if P_g belonged to the set of P_{cc} points, then it would have the form

$$P_g = ((P_j(1 - d_1) + P_k d_1)(1 - d_3) + (P_l(1 - d_2) + P_m d_2)d_3)(1 - d_7) \quad (5.9)$$

Substituting P_g of Equation 5.9 into Equation 5.8 yields an over-determined system of equations, which in turn is an identity. Therefore the set of P_{cc} points is a convex set.

QED

We believe this proof lifts to higher dimensions.

5.1.7 First-order affine geometry

The above modal language is again a fragment of a first-order one, under the standard translation. The relevant first-order language is not quite that of Tarski's elementary geometry for \mathbb{R}^2 , as we also get unary predicate letters denoting regions. In fact, one open question which we have not been able to resolve is this. A formula $\varphi(\beta, P, Q, \dots)$ is valid, say in the real plane, if it holds for any interpretation of the regions P, Q, \dots . Thus, we would be looking at a universal fragment of a *monadic second-order logic*:

What is the complete monadic Π_1^1 theory of the affine real plane?

We suspect it is recursively axiomatizable and decidable—perhaps using the Ehrenfeucht game methods of [Doets, 1987]. This is an extension of the affine part of Tarski's logic. But our previous discussion has also identified interesting *fragments*:

What is the universal first-order theory of the affine real plane?

As in our discussion of topology, the affine first-order language of regions is a natural limit towards which modal affine languages can strive via various logical extensions. From a geometrical viewpoint, one might also hope that 'layering' the usual language in this modal way will bring to light interesting new geometrical facts.

Another major feature of standard geometry is the *equal status of points and lines*. This would suggest a reorganization of the modal logic to a *two-sorted* one stating properties of both points and segments, viewed as independent semantic objects. There are several ways of doing this. One would be a *two-dimensional* modal language in the spirit of [Marx and Venema, 1997], handling both points and pairs of points, with various cross-sortal modalities. Another would treat both objects as primitives, and then have cross-sortal modalities for "at an end-point," "at an intermediate point," "at some surrounding segment." We think the latter is the best way to go eventually,

as it has the useful feature of replacing talk in terms of ternary relations, which are hard to visualize, by binary ones, which are easier to represent. (This is of course the key advantage of the geometer’s habit of working with points and lines.) Moreover, the two-sorted move would be in line with other modal trends such as *Arrow Logic* [van Benthem, 1996, Venema, 1996], where transitions between points become semantical objects in their own right. This gives more control over semantic structures and the complexity of reasoning. It would also help reflect geometrical duality principles of the sort that led from affine to projective geometry.

5.2 Metric geometry

There is more structure to geometry than just affine point and line patterns. Tarski’s equidistance also captures *metric* information. There are various primitives for this. Tarski used quaternary equidistance—while ternary equidistance would do just as well (x , y and z lie at equal distances). Our choice in this section is a different one, stressing the comparative character of metric structure.

5.2.1 The geometry of relative nearness

Relative nearness was introduced in [van Benthem, 1983b] (see Figure 5.13):

$$N(x, y, z) \text{ iff } y \text{ is closer to } x \text{ than } z \text{ is, i.e., } d(x, y) < d(x, z) \\ \text{where } d(x, y) \text{ is any distance function.}$$

This is meant very generally. The function d can be a geometrical metric, or some

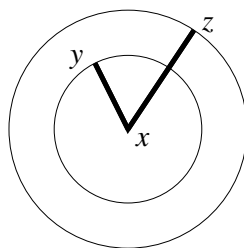


Figure 5.13: From point x , y is closer than point z .

more cognitive notion of visual closeness (van Benthem’s original interest; cf. also Gärdenfors ‘Conceptual Spaces’), or some utility function with metric behavior. Randedell et al. [2001] develop the theory of comparative nearness for the purpose of robot navigation, related to the earlier-mentioned calculus of regions RCC.

Relative nearness defines equidistance:

$$Eqd(x, y, z) : \neg N(x, y, z) \wedge \neg N(x, z, y)$$

Tarski's quaternary equidistance is expressible in terms of N as well. Details are postponed until Section 5.2.3 on first-order metric geometry.

Affine betweenness is also definable in terms of N , at least in the real spaces \mathbb{R}^n :

$$\beta(xyz) \text{ iff } \forall x' \neg(N(y, x, x') \wedge N(z, x', x))$$

Finally, note that even identity of points $x = y$ is expressible in terms of N

$$x = y \text{ iff } \neg N(x, x, y)$$

At the end of the XVIII century the mathematician Lorenzo Mascheroni proved in his tractate *The Geometry of Compasses* that everything that can be done with compass and ruler can be done with the compass alone. One can generate all of Mascheroni's constructions with the first-order logic of N and thereby achieve geometry, as we illustrate in Section 5.2.1.1.

The further analysis of this structure can proceed along much the same lines as the earlier one for affine geometry. In particular, as a source of basic constraints, one is interested in the *universal first-order theory* of relative nearness. Its complete description is an open question right now, but here are some examples showing its interest. First, comparative nearness induces a standard comparative ordering. Once a point x is fixed, the binary order $N(x, y, z)$ is irreflexive, transitive and almost-connected:

$$\begin{aligned} \forall x \forall y \forall z \forall u : (N(x, y, z) \wedge N(x, z, u)) &\rightarrow N(x, y, u) && \text{(transitivity)} \\ \forall x \forall y : \neg N(x, y, y) &&& \text{(irreflexivity)} \\ \forall x \forall y \forall z \forall u : N(x, y, z) &\rightarrow (N(x, y, u) \vee N(x, u, z)) && \text{(almost-connectedness)} \end{aligned}$$

These are like the principles of comparative order in logical semantics for counterfactuals [Lewis, 1973]. But additional valid principles are more truly geometrical, relating distances from different standpoints. These are the following *triangle inequalities*

$$\begin{aligned} \forall x \forall y \forall z \forall u : N(x, y, z) \wedge N(z, x, y) &\rightarrow N(y, x, z) \\ \forall x \forall y \forall z \forall u : \neg N(x, y, z) \wedge \neg N(z, x, y) &\rightarrow \neg N(y, x, z) \end{aligned}$$

These seem pretty universal constraints on comparative nearness in general. Further universal first-order properties of N reflect the two-dimensionality of the plane. Just inscribe 6 equilateral triangles in a circle, and see that

on a circle with radius r , the largest polygon that can be inscribed of points at distance r has 6 vertices.

This upper bound can be expressed in universal first-order form, because we can express equidistance in terms of N . Other principles of this form concern the arrangement of points on circles:

on a circle C , any point has at most two points at each of its ‘equidistance levels’ on C

and

circles with the same radius but different centers intersect in at most two points.

To obtain the complete universal first-order theory of comparative nearness in the Euclidean plane \mathbb{R}^2 , one would have to guarantee a planar embedding. Do our general axioms, including the triangle inequalities, suffice for axiomatizing the complete universal theory of *all* Euclidean spaces \mathbb{R}^n ?

5.2.1.1 Geometrical excursus: Mascheroni, Voronoi, and Delaunay

Having the predicate of comparative nearness is like having a compass. One is able to draw circles, but does not know their radius. (Think of having a map with no scale on it and to start comparing distances by using the compass alone.) If $N(x, y, z)$ holds, we know that the points y are those contained in the circle centered in x , whose radius is given by the distance $d(x, z)$. Via the defined notion of equidistance we can also refer to the circumference of the regions. This allows us to look at modal and first-order nearness geometry through some classical theorems in geometry.

Mascheroni’s geometry of compass. If one can define circumferences via equidistance and one can ‘do’ all basic geometrical constructions with the compass alone via Mascheroni constructions, then the logic of comparative nearness must be able to express all basic geometrical constructions.

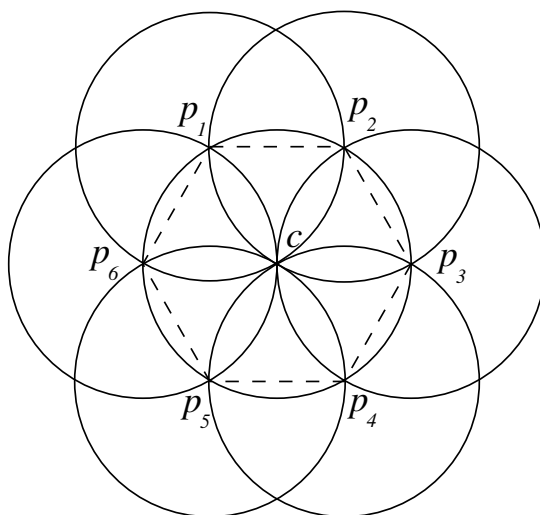


Figure 5.14: The construction of a regular hexagon via Mascheroni’s construction.

Here is an example. Let us construct a regular hexagon using the N relation alone, see Figure 5.14.

$$\begin{aligned}
 & p_1 \neq p_2 \neq p_3 \neq p_4 \neq p_5 \neq p_6 \wedge \\
 & E(c, p_1, p_2) \wedge E(c, p_2, p_3) \wedge E(c, p_3, p_4) \wedge E(c, p_4, p_5) \wedge E(c, p_5, p_6) \wedge \\
 & E(p_1, p_2, c) \wedge E(p_2, p_3, c) \wedge E(p_3, p_4, c) \wedge E(p_4, p_5, c) \wedge E(p_5, p_6, c) \rightarrow \\
 & E(p_6, p_1, c) \tag{5.10}
 \end{aligned}$$

In the first line of Equation 5.10, we identify six disjoint points $p_1 \dots p_6$. In the second line, we constrain the six points to lie on the same circumference centered in c . Finally, we build circles of the same radius as that centered in c that connect the points pairwise. As a result, the six segments $p_1-p_2, p_2-p_3, p_3-p_4, p_4-p_5, p_5-p_6, p_6-p_1$ define a regular hexagon. The technique generalizes for the other geometrical constructions.

Voronoi diagrams. Imagine having a set of marked points scattered in space. Then consider the partitioning of the space in regions, one for every marked point. A region is defined as the set of points that are closest to the marked point than to any other one. The Voronoi diagram of the marked points is the union of all the boundaries of such regions, [Voronoi, 1908]. An example of the Voronoi diagram of four points p_1, p_2, p_3, p_4 on the plane is depicted in Figure 5.15.(a). The definition of a Voronoi

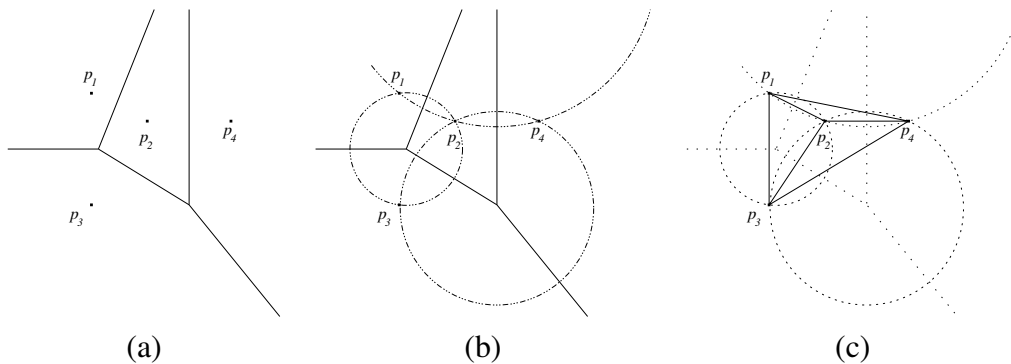


Figure 5.15: (a) The Voronoi diagram of four points on the plane. (b) The circles connecting neighboring points. (c) Delaunay triangulation.

region, also called *cell*, is, at a closer look, given in terms of comparative nearness. Following this intuition, we define a cell of the points $P = \{p_1, p_2, \dots, p_k\}$ in terms of the nearness relation:

$$\text{cell}(p, p_l) \leftrightarrow \bigwedge_{\{i=1, i \neq l\}}^k N(p_l, p, p_i)$$

The interpretation is that p is in the cell of the point p_l if it is closer to p_l than to any

other of the points in P . The whole Voronoi diagram is then:

$$\text{diagram}(p, P) \leftrightarrow \bigvee_{l=1}^k \bigvee_{\{m=1, m \neq l\}}^k E(p, p_l, p_m) \wedge \bigwedge_{\{j=1, m \neq j \neq l\}}^k N(p_l, p, p_j)$$

A point p is in the Voronoi diagram of the set of points P if it is equidistant from two given points and it is closer to these two points than to any other one of the set P .

Delaunay triangulation. By connecting with segments the points of P which share an edge in the Voronoi diagram, one obtains a graph. The operation is called Delaunay triangulation, [Delaunay, 1934]. If one thinks of the points of P as the vertices of a polyhedron, then the Delaunay triangulation gives a procedure to partition the polygon into tetrahedra. In Figure 5.15.(c) the Delaunay triangulation of four points on the plane. To express this in terms of the comparative nearness operator, we use the property that there are no points of P inside a circle circumscribing the three vertex of a Delaunay triangle.

$$\begin{aligned} \text{Delaunay}(p, P) \leftrightarrow & \\ & \bigvee_{l=1}^k \bigvee_{\{m=1, m \neq l\}}^k \bigvee_{\{n=1, n \neq m\}}^k E(p_l, p_m, c) \wedge E(p_m, p_n, c) \wedge E(p_n, p_l, c) \wedge \\ & \bigwedge_{\{j=1, j \neq l, m, n\}}^k N(p_l, c, p_j) \wedge \\ & (B(p, p_l, p_m) \vee B(p, p_m, p_n) \vee B(p, p_n, p_l)) \end{aligned}$$

The construction is a bit laborious. One begins by constructing the circle passing for any three given points of P . The center of such circle is c . If the circle contains no point of P different from p_l, p_m, p_n , then the three points form a Delaunay triangle. The triangle is then defined by its three sides, which we denote via the betweenness operator. Figure 5.15 shows the construction of the Voronoi and Delaunay triangulation stepwise for four points of the plane. We remark that the definitions given here are completely general and apply also to more than two dimensions. In three dimensions for instance, via Delaunay triangulation, one partitions the space between the points of P into a number of disjoint tetrahedra.

5.2.2 Modal logic of nearness

The ternary relation of comparative nearness lends itself to modal description, just like ternary betweenness. We will just briefly sketch the resulting logic, which is like our affine system in its broad outline. But the intuitive meaning of N also adds some new issues of its own.

5.2.2.1 Modal languages of nearness

First, one sets

$$M, x \models \langle N \rangle \phi, \psi \text{ iff } \exists y, z : M, y \models \psi \wedge M, z \models \phi \wedge N(x, y, z)$$

The universal dual is also interesting in its spatial behavior:

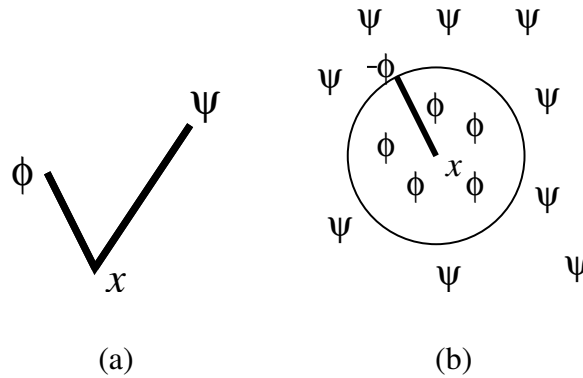


Figure 5.16: Interpreting a modal operator of nearness and its dual.

$$M, x \models [N] \phi, \psi \text{ iff } \forall y, z : N(x, y, z) \wedge M, y \models \neg\phi \rightarrow M, z \models \psi$$

Dropping the negation, one gets an interchangeable version with the following intuitive content:

if any point y around the current point x satisfies ϕ , then all points z further out must satisfy ψ .

Moreover, there are obvious *versatile* versions of these modal operators, which look at the same situation in a different way. For instance, using one of these in its universal version, we can also express the appealing statement that

if any point y around the current point x satisfies ϕ , then all points z closer to x must satisfy ψ .

See Figure 5.17 for an illustration. Finally, note that this language defines an existential modality (assuming the mild condition that $\forall y : N(x, x, y) \vee x = y$):

$$E\phi \text{ iff } \phi \vee \langle N \rangle (\top, \phi)$$

Without the stated condition, this existential modality will only range over connected components.

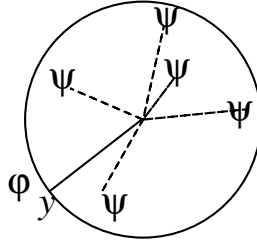


Figure 5.17: Versatile interpretation of the dual of a modal operator of nearness.

5.2.2.2 Modal logics of nearness

Modal logics of nearness again start with universally valid principles, with distribution as the prime example:

$$\begin{aligned}\langle N \rangle(\varphi \vee \psi, \xi) &\leftrightarrow \langle N \rangle(\varphi, \xi) \vee \langle N \rangle(\psi, \xi) \\ \langle N \rangle(\varphi, \psi \vee \xi) &\leftrightarrow \langle N \rangle(\varphi, \psi) \vee \langle N \rangle(\varphi, \xi)\end{aligned}$$

Universal constraints of earlier kinds will return as special axioms. Here is an example:

$$\langle N \rangle(\varphi, \psi) \wedge \neg \langle N \rangle(\varphi, \varphi) \wedge \neg \langle N \rangle(\psi, \psi) \wedge \langle N \rangle(\psi, \xi) \rightarrow \langle N \rangle(\varphi, \xi) \quad (\text{transitivity})$$

In the above definition the two clauses $\neg \langle N \rangle(\varphi, \varphi)$ and $\neg \langle N \rangle(\psi, \psi)$ are necessary to ensure that φ and ψ can be true only at a fixed distance from the current point. Omitting them results in an invalid principle, as it may very well be the case that $\langle N \rangle(\varphi, \psi) \wedge \langle N \rangle(\psi, \xi) \wedge \neg \langle N \rangle(\xi, \varphi)$ if φ is true at points at different distances from the current one. Another example of a universal constraint is almost-connectedness:

$$\langle N \rangle(\varphi, \psi) \wedge \neg \langle N \rangle(\varphi, \varphi) \wedge \neg \langle N \rangle(\psi, \psi) \wedge E\xi \rightarrow \langle N \rangle(\varphi, \xi) \vee \langle N \rangle(\xi, \psi) \quad (\text{almost-connectedness})$$

Irreflexivity seems harder to define (as usual in modal logics), but see below.

Special logics of nearness arise by looking at special structures, or at least, imposing more particular constraints. These can again be computed by correspondence techniques. In a similar way, one can modally express the *triangle inequalities*. But in fact, there is a more general observation to be made here. Note that our language can define that φ holds in a unique point:

$$E! \varphi \text{ iff } E(\varphi \wedge \neg \langle N \rangle(\varphi, \varphi))$$

Now observe the following.

5.2.1. PROPOSITION. *Every universal first-order property of N is modally definable.*

Proof Every such property is of the form

$$\forall x_1 \dots \forall x_k : BC(N(x_i, x_j, x_k))$$

where BC stands for any Boolean combination of nearness atoms. Now take proposition letters p_1, \dots, p_k and write

$$E!p_1 \wedge \dots \wedge E!p_k \rightarrow BC(N^\#(x_i, x_j, x_k))$$

where $N^\#(x_i, x_j, x_k)$ is defined as $E(p_i \wedge \langle N \rangle(p_j, p_k))$. It is evident that this is a modal frame correspondent. QED

This explains the definition of the triangle inequalities. Moreover, *irreflexivity* (whose first-order definition is $\forall x \forall y \neg N(x, y, y)$) is definable after all by

$$E!p_1 \wedge E!p_2 \rightarrow \neg E(p_1 \wedge \langle N \rangle(p_2, p_2))$$

5.2.2.3 Modal extensions

Useful modal extensions of the base language are partly as in the affine case. But there is also a novelty. In describing patterns, one may often want to say something like this:

for every φ -point around x , *there exists* some closer ψ -point.

Now this is not definable in our language, which uses uniform EE or AA quantifier combinations. Mixing universal and existential quantifiers is more like temporal ‘Until’ languages. Speaking generally, we want a new operator:

$$M, x \models \langle N^{\exists\forall} \rangle(\varphi, \psi) \text{ iff } \forall y (M, y \models \varphi \rightarrow \exists z (N(z, y, x) \wedge M, z \models \psi))$$

The general logic of this additional modality over a ternary relation is a bit more complex with respect to distribution and monotonicity behavior—but it can be axiomatized completely, at least minimally, over all abstract models.

Indeed, this universal-existential pattern is reminiscent of other modal logics naturally involving ternary frame relations. One example is *temporal logic* of Since and Until, which involves moving to some point around the current point in time, and then saying something about all points in between. One existential-universal variant of the preceding modality would indeed be a kind of spatial Until, stating that some point on a circle around the current point satisfies φ , while all points in the interior satisfy ψ . This is almost a metric analogue of the topological Until operator in Section 4.2.2, but the latter should have the whole circle boundary satisfy φ , which requires one more universal modality over equidistant points.

Another intriguing analogy is with a typical modal logic over comparative nearness, viz. *conditional logic*. The latter is mostly known in connection with counterfactuals and default reasoning [Lewis, 1973, Nute, 1983, Veltman, 1985]. In general conditional logic, one crucial binary modality reads

$$\varphi \Rightarrow \psi \text{ iff every closest } \varphi\text{-world is } \psi$$

This satisfies the usual Lewis axioms for conditional semantics in terms of ‘nested spheres’ (cf. [van Benthem, 1983a]):

$$\begin{aligned}
 & \varphi \Rightarrow \psi \rightarrow \varphi \Rightarrow \psi \vee \xi \\
 & \varphi \Rightarrow \psi \wedge \varphi \Rightarrow \xi \rightarrow \varphi \Rightarrow \psi \wedge \xi \\
 & \varphi \Rightarrow \psi \wedge \varphi \Rightarrow \xi \rightarrow \varphi \wedge \psi \Rightarrow \xi \\
 & \varphi \Rightarrow \psi \wedge \xi \Rightarrow \psi \rightarrow \varphi \vee \xi \Rightarrow \psi \\
 & ((\varphi \vee \psi) \Rightarrow \varphi) \vee (\neg((\varphi \vee \psi) \Rightarrow \xi)) \vee (\psi \Rightarrow \xi)
 \end{aligned}$$

The interesting open question here concerns modal-conditional reflections of the additional geometrical content of the $N(x, y, z)$ relation. Lewis’ complete system is just about ordering properties of comparisons from some fixed vantage point. This shows in the fact that there are no significant axioms for *iterated conditionals* which require shifts in vantage point. What is the conditional logic content of the triangle inequalities?

5.2.3 First-order theory of nearness

As for the complete first-order theory of relative nearness, we have no special results to offer, except for the promised proof of an earlier claim.

5.2.2. FACT. *The single primitive of comparative nearness defines the two primitives of Tarski’s Elementary Geometry in first order logic.*

Proof The following defines betweenness (see Figure 5.18):

$$\beta(yxz) \text{ iff } \neg \exists x' : N(y, x', x) \wedge N(z, x', x)$$

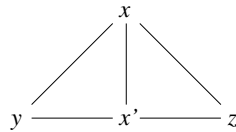


Figure 5.18: Defining betweenness via nearness.

This allows us to define parallel segments in the usual way, as having no intersection points on their generated lines.

$$\begin{aligned}
 xx' || yy' & \leftrightarrow \neg \exists c : \beta(xx'c) \wedge \beta(yy'c) \wedge \\
 & \neg \exists c' : \beta(c'xx') \wedge \beta(cyy') \wedge \\
 & \neg \exists c'' : \beta(xcx') \wedge \beta(ycy')
 \end{aligned}$$

Then, one defines equal segment length by

$$\delta(x, y, z, u) \text{ iff } \exists y' : xu \parallel yy' \wedge xy \parallel uy' \wedge \neg N(u, z, y') \wedge \neg N(u, y'z)$$

Intuitively, one moves one segment on the other matching end-points and preserving length via parallel lines. Then state that the other end-points are at the same distance from the joined point. See the depiction in Figure 5.19. QED

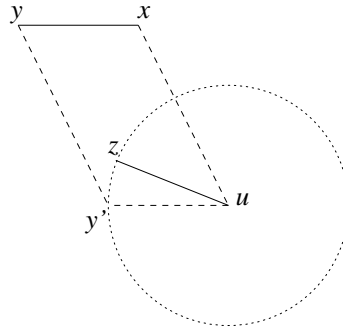


Figure 5.19: Equidistance in terms of betweenness.

Apart from this, much of our earlier discussion concerning affine first-order geometry applies. Incidentally, no claim is made here for the originality of this approach per se. There are many systems of logical geometry which have similar richness. A case in point is the axiomatization of constructive geometry in [von Plato, 1995].

5.3 Linear algebra

Our final example of modal structures inside a spatial theory is different in spirit from either topology or standard geometry. Connections between *linear algebra* and spatial representation are well-known from a major qualitative visual theory, viz. *mathematical morphology*. Our treatment follows the lines of [Aiello and van Benthem, 1999]—and especially [van Benthem, 2000], which also has further details. (A different connection between mathematical morphology and modal logic is found in [Bloch, 2000], which also includes a fuzzy version.) The flavor of this brand of spatial reasoning is different from what we had before—but similar modal themes emerge all the same.

Mathematical morphology, developed in the 60s by Matheron and Serra, [Matheron, 1967, Serra, 1982], underlies modern image processing, where it has a wide variety of applications. Compared with classical signal processing approaches it is more efficient in image preprocessing, enhancing object structure, and segmenting objects from the background. The modern mathematics behind this involves lattice theory: [Heijmans, 1994]. Logicians may want to think of ‘linear algebras’ [Girard, 1987], an abstract version of vector spaces:

5.3.1. DEFINITION (LINEAR ALGEBRA). $\langle X, \sqcap, \sqcup, \perp, \dashv, \star, \check{0}, \check{1} \rangle$ is a *linear algebra* if

- (i) $\langle X, \sqcap, \sqcup, \perp, \rangle$ is a lattice with bottom \perp ;
- (ii) $\langle X, \star, \check{1} \rangle$ is a monoid with unit $\check{1}$;
- (iii) if $x \leq x', y \leq y'$, then $x \star y \leq x' \star y'$ and $x' \dashv y \leq x \dashv y'$;
- (iv) $x \star y \leq z$ iff $x \leq y \dashv z$;
- (v) $x = (x \dashv \check{0}) \dashv \check{0}$ for all x .

In line with our spatial emphasis of this chapter, we will stick with concrete vector spaces \mathbb{R}^n in what follows. Images are regions consisting of sets of vectors. Mathematical morphology provides four basic ways of combining, or simplifying images, viz. *dilation*, *erosion*, *opening* and *closing*. These are illustrated pictorially in Figure 5.20. Intuitively, dilation adds regions together—while, e.g., erosion is a way of

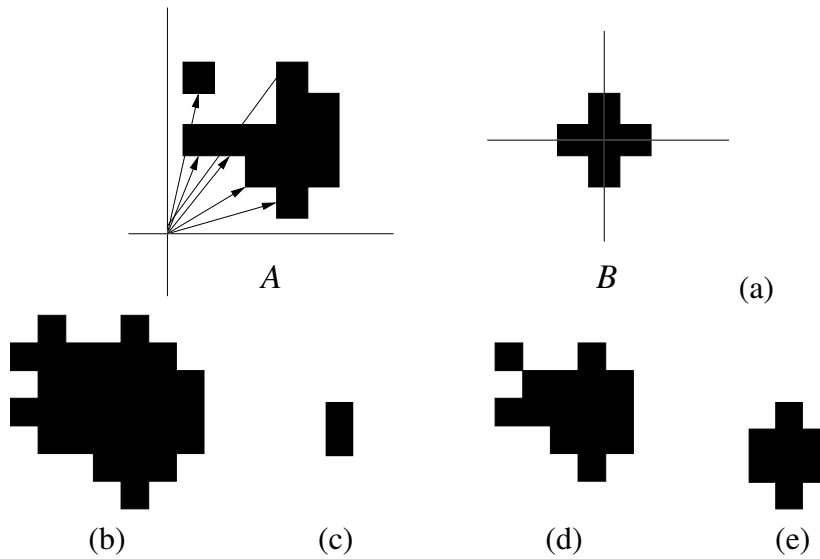


Figure 5.20: (a) Regions A and B , elements of the vector space \mathbb{N}^2 ; (b) dilating A by B ; (c) eroding A by B ; (d) closing A by B ; (e) opening A by B .

removing ‘measuring idiosyncrasies’ from a region A by using region B as a kind of boundary smoothener. (If B is a circle, one can think of it as rolling tightly along the inside of A ’s boundary, leaving only a smoother interior version of A .) More formally, dilation, or *Minkowski addition* \oplus is vector sum:

$$A \oplus B = \{a + b \mid a \in A, b \in B\} \quad \text{dilation}$$

This is naturally accompanied by

$$A \ominus B = \{a \mid a + b \in A, \forall b \in B\} \quad \text{erosion}$$

Openings and closing are combinations of dilation and erosions:

$$\begin{array}{ll} \text{the structural } \textit{opening} \text{ of } A \text{ by } B & (A \ominus B) \oplus B \\ \text{the structural } \textit{closing} \text{ of } A \text{ by } B & (A \oplus B) \ominus B \end{array}$$

In addition, mathematical morphology also employs the usual Boolean operations on regions: intersection, union, and complement. This is our third mathematization of real numbers \mathbb{R}^n in various dimensions, this time focusing on their vector structure. Evidently, the above operations are only a small sub-calculus, chosen for its computational utility and expressive perspicuity.

5.3.1 Mathematical morphology and linear logic

The first connection that we note lies even below the level of standard modal languages. The Minkowski operations behave a bit like the operations of *propositional logic*. Dilation is like a logical conjunction \oplus , and erosion like an implication \longrightarrow , as seems clear from their definitions ('combining an A and a B ', and 'if you give me a B , I will give an A '). The two were related by the following *residuation law*:

$$A \bullet B \subseteq C \text{ iff } A \subseteq B \longrightarrow C$$

which is also typical for conjunction and implication (cf. also clause (iv) in Definition 5.3.1). Thus, \longrightarrow is a sort of inverse to \oplus .

5.3.1.1 Resource logics

There already exists a logical calculus for these operations, invented under the *multiplicative linear logic* name in theoretical computer science [Troelstra, 1992], and independently as the *Lambek calculus with permutation* in logical linguistics, cf. [Kurtonina, 1995]. The calculus derives 'sequents' of the form $A_1, \dots, A_k \Rightarrow B$ where each expression A, B in the current setting stands for a region, and the intended interpretation, in our case, says that

the sum of the A 's is included in the region denoted by B .

Here are the derivation rules, starting from basic axioms $A \Rightarrow A$:

$$\frac{X \Rightarrow A \quad Y \Rightarrow B}{X, Y \Rightarrow A \bullet B} \qquad \frac{X, A, B \Rightarrow C}{X, A \bullet B \Rightarrow C} \quad \text{(product rules)}$$

$$\frac{A, X \Rightarrow B}{X \Rightarrow A \longrightarrow B} \qquad \frac{X \Rightarrow A \quad B, Y \Rightarrow C}{X, A \longrightarrow B, Y \Rightarrow C} \quad \text{(arrow rules)}$$

$$\frac{X \Rightarrow A}{\pi[X] \Rightarrow A} \text{permutation} \qquad \frac{X \Rightarrow A \quad A, Y \Rightarrow B}{X, Y \Rightarrow B} \text{cut} \quad \text{(structural rules)}$$

Derivable sequents typically include:

$$\begin{aligned} A, A \longrightarrow B &\Rightarrow B && \text{('function application')} \\ A \longrightarrow B, B \longrightarrow C &\Rightarrow A \longrightarrow C && \text{('function composition')} \end{aligned}$$

Here is an example of a derivation, just for the flavor of the system:

$$\frac{\frac{\frac{A \Rightarrow A \quad B \Rightarrow B}{A, A \longrightarrow B \Rightarrow B} \quad \overline{C \Rightarrow C}}{A, A \longrightarrow B, B \longrightarrow C \Rightarrow C}}{A \longrightarrow B, B \longrightarrow C \Rightarrow A \longrightarrow C}$$

Another key example are the two ‘Currying’ laws, whose proof uses the \bullet rules:

$$\begin{aligned} (A \bullet B) \longrightarrow C &\Rightarrow (A \longrightarrow (B \longrightarrow C)) \\ (A \longrightarrow (B \longrightarrow C)) &\Rightarrow (A \bullet B) \longrightarrow C \end{aligned}$$

This calculus is best understood in terms of *resources*. Think of each premise in an argument as a resource which you can use just once when ‘drawing’ the conclusion. In standard logical inference, the premises form a set: you can duplicate the same item, or contract different occurrence of it without any change in valid conclusions. This time, however, the premises form a *bag*, or *multi-set*, of occurrences: validating only ‘resource-conscious’ versions of the standard logical laws. E.g., ‘Modus Ponens’ $A, A \longrightarrow B \Rightarrow B$ is valid, but its variant $A, A, A \longrightarrow B \Rightarrow B$ is not: there is one unused resource left. A correct, and provable sequent using the latter resources is:

$$A, A, A \longrightarrow B \Rightarrow A \bullet B$$

Or consider the classically valid sequent $A, (A \longrightarrow (A \longrightarrow B)) \Rightarrow B$. Here the above calculus only proves $A, (A \longrightarrow (A \longrightarrow B)) \Rightarrow A \longrightarrow B$, and you must supply one more resource A to derive

$$A, A, (A \longrightarrow (A \longrightarrow B)) \Rightarrow B.$$

The related categorial grammar interpretation for this same calculus reads the product \bullet as syntactic juxtaposition of linguistic expressions, and an implication $A \longrightarrow B$ as a function category taking A -type to B -type expressions. The same occurrence-based character will hold: repeating the same word is not the same as having it once.

The major combinatorial properties of this calculus **LL** are known, including proof-theoretic cut elimination theorems, and *decidability* of derivability in NP time. Moreover, there are several formal semantics underpinning this calculus (algebraic, game-theoretic, category-theoretic, possible worlds-style [van Benthem, 1991a]). Still, no totally satisfying modeling has emerged so far.

5.3.1.2 Linear logic as mathematical morphology

Here is where the present setting becomes intriguing: mathematical morphology provides a new model for linear logic.

5.3.2. FACT. *Every space \mathbb{R}^n with the Minkowski operations is a model for all **LL**-provable sequents.*

This soundness theorem shows that every sequent one derives in **LL** must be a valid principle of mathematical morphology. One can see this for the above examples, or other ones, such as the idempotence of morphological opening $(A \oplus B) \oplus B$:

$$(((A \oplus B) \oplus B) \oplus B) \oplus B = ((A \oplus B) \oplus B)$$

In **LL**, the opening is $(A \longrightarrow B) \bullet A$, and the idempotence law is literally derivable using the above rules:

$$\begin{aligned} (A \longrightarrow B) \bullet A &\Rightarrow (A \longrightarrow ((A \longrightarrow B) \bullet A)) \bullet A \\ (A \longrightarrow ((A \longrightarrow B) \bullet A)) \bullet A &\Rightarrow (A \longrightarrow B) \bullet A \end{aligned}$$

The list might even include new principles not considered in that community. The converse seems an open completeness question of independent interest:

*Is multiplicative linear logic complete w.r.t the class of all \mathbb{R}^n 's?
Or even w.r.t. two-dimensional Euclidean space?*

Further, mathematical morphology laws ‘mix’ pure Minkowski operations \oplus, \longrightarrow with standard Boolean ones. E.g. they include the fact that $A \longrightarrow (B \cap C)$ is the same as $(A \cup B) \longrightarrow C = (A \longrightarrow C) \cap (B \longrightarrow C)$. This requires adding Boolean operations:

$$\begin{array}{ccc} \frac{X, A \Rightarrow B}{X, A \cap C \Rightarrow B} & \frac{X, A \Rightarrow B}{X, C \cap A \Rightarrow B} & \frac{X \Rightarrow A \quad X \Rightarrow B}{X \Rightarrow A \cap B} \\ \\ \frac{X \Rightarrow A}{X \Rightarrow A \cup B} & \frac{X \Rightarrow A}{X \Rightarrow B \cup A} & \frac{X, A \Rightarrow B \quad X, C \Rightarrow B}{X, A \cup C \Rightarrow B} \end{array}$$

Note the difference between the two conjunctions. Product \bullet and intersection have some similarities, but the rules are different. E.g., $A \longrightarrow (B \bullet C)$ does not derive $(A \longrightarrow B) \bullet (A \longrightarrow C)$, or vice versa. Conversely, dot product satisfied the ‘Curry laws’, but $(A \cap B) \longrightarrow C$ is certainly not derivably equivalent to $(A \longrightarrow (B \longrightarrow C))$. All these observations tally with known facts in mathematical morphology. Indeed, the extended calculus is still sound—while its completeness remains an open question.

The Boolean operations look a bit like the ‘additives’ of linear logic, but they also recall ordinary modal logic, which is where we are going now.

5.3.2 Richer languages

Evidently, the basic players in an algebra of regions in a vector space are the vectors themselves. For instance, Figure 5.20.a represents the region A as a set of 13 vectors departing from the origin. Vectors come with some natural operations, such as binary addition, or unary inverse—witness the usual definition of a vector space. A vector v in our particular spaces may be viewed as an ordered pair of points (o, e) , with o the origin and e the end point. Pictorially, this is an arrow from o to e . Now this provides our point of entry into modal logic.

5.3.2.1 Arrow logic

Arrow logic is a form of modal logic where the objects are transitions or arrows, structured by various relations. In particular, there is a binary modality for *composition* of arrows, and a unary one for *converse*. The motivation for this comes from dynamic logics, treating transitions as objects in their own right, and from relational algebra, making pairs of points separate objects. This allows for greater expressive power than the usual systems, while also lowering complexity of the core logics (see [Blackburn et al., 2001, van Benthem, 1996] for overviews). Consider in particular the pair-interpretation, with arrows being pairs of points (a_o, a_e) . Here are the fundamental semantic relations:

composition $C(a_o, a_e)(b_o, b_e)(c_o, c_e)$ iff $a_o = b_o$, $a_e = c_e$, and $b_e = c_o$,

inverse $R(a_o, a_e)(b_o, b_e)$ iff $a_o = b_e$, and $a_e = b_o$,

identity $I(a_o, a_e)$ iff $a_o = a_e$.

An abstract model is then defined as any set of arrows as primitive objects, with three relations as above, and a valuation function sending each proposition letter p to the set of the arrows where property p holds.

5.3.3. DEFINITION (ARROW MODEL). An *arrow model* is a tuple $M = \langle W, C, R, I, \nu \rangle$ such that $C \subseteq W \times W \times W$, $R \subseteq W \times W$, $I \subseteq W$, and $\nu : W \rightarrow P$.

Such models have a wide variety of interpretations, ranging from concrete models in linguistic syntax to abstract ones in category theory [Venema, 1996]—but of relevance to us is the obvious connection with vector spaces. Think of $Cxyz$ as $x = y + z$, Rxy as $x = -y$ and Ix as $x = 0$. To make this even clearer, we use a modal arrow language with proposition letters, the identity element 0, monadic operators \neg , $-$, and a dyadic operator \oplus . The truth definition reads:

$M, x \models p$	iff	$a \in \nu(p)$
$M, x \models 0$	iff	Ix
$M, x \models -\varphi$	iff	$\exists y : Rxy$ and $M, y \models \varphi$
$M, x \models \neg\varphi$	iff	not $M, x \models \varphi$
$M, x \models \varphi \vee \psi$	iff	$M, x \models \varphi$ or $M, x \models \psi$
$M, x \models A \oplus B$	iff	$\exists y \exists z : Cxyz \wedge M, y \models A \wedge M, z \models B$
$M, x \models A \ominus B$	iff	$\forall y \forall z : Cxyz \wedge M, z \models A \rightarrow M, y \models B$

This system can be studied like any modal logic. For the basic results in the area, we refer to the above-mentioned publications.

5.3.2.2 Arrow logic as linear algebra

Most modal topics make immediate sense in linear algebra or mathematical morphology. E.g., the above models support a natural notion of *bisimulation*:

5.3.4. DEFINITION (ARROW BISIMULATION). Let M, M' be two arrow models. A relation $\rightleftharpoons \subseteq W \times W'$ is an *arrow bisimulation* iff, for all x, x' such that $x \rightleftharpoons x'$:

base $x \in \nu(p)$ iff $x' \in \nu'(p)$,

C-forth $Cxyz$ only if there are $y'z' \in W'$ such that $C'x'y'z'$, $y \rightleftharpoons y'$ and $z \rightleftharpoons z'$,

C-back $C'x'y'z'$ only if there are $yz \in W$ such that $Cxyz$, $y \rightleftharpoons y'$ and $z \rightleftharpoons z'$,

R-forth Rxy only if there are $y' \in W'$ such that $R'x'y'$ and $y \rightleftharpoons y'$,

R-back $R'x'y'$ only if there are $y \in W$ such that Rxy and $y \rightleftharpoons y'$,

I-harmony Ix iff $I'x'$.

Arrow bisimulation is a coarser comparison of vector spaces than the usual linear transformations. It preserves all modal statements in the above modal arrow language, and hence provide a lower level of visual analysis in linear algebra similar to what we have found earlier for topology, or geometry.

Next, logics for valid reasoning also transfer immediately. Here is a display of the basic system of arrow logic:

$$(\varphi \vee \psi) \oplus \xi \leftrightarrow (\varphi \vee \phi) \oplus \xi \quad (5.11)$$

$$\varphi \oplus (\psi \vee \xi) \leftrightarrow (\varphi \oplus \phi) \vee (\varphi \oplus \xi) \quad (5.12)$$

$$-(\varphi \vee \psi) \leftrightarrow -\varphi \vee -\psi \quad (5.13)$$

$$\varphi \wedge (\psi \oplus \xi) \rightarrow \psi \oplus (\xi \wedge (-\psi \oplus \varphi)) \quad (5.14)$$

These principles either represent or imply obvious vector laws. Here are some consequences of (5.13), (5.14):

$$-(\neg A) \leftrightarrow \neg(-A)$$

$$-(A + B) \leftrightarrow -B + -A$$

$$A + \neg(-A + \neg B) \rightarrow B$$

The latter ‘triangle inequality’ is the earlier rule of Modus Ponens in disguise. On top of this, special arrow logics have been axiomatized with a number of additional frame conditions. In particular, the vector space interpretation makes composition *commutative* and *associative*, which leads to further axioms:

$$A \oplus B \leftrightarrow B \oplus A \quad \text{commutativity}$$

$$A \oplus (B \oplus C) \leftrightarrow (A \oplus B) \oplus C \quad \text{associativity}$$

These additional principles make the calculus simpler in some ways than basic arrow logic. The key fact about composition is now the vector law

$$a = b + c \quad \text{iff} \quad c = a - b$$

which derives the triangle inequality. And there are also expressive gains. E.g., the modal language becomes automatically ‘versatile’ in our earlier sense.

Again the soundness of the given arrow logic for vector algebra is clear, and we can freely derive old and new laws of vector algebra. But the central open question about arrow logic and mathematical morphology is again a converse:

What is the complete axiomatization of arrow logic over the standard vector spaces \mathbb{R}^n ?

In particular, are there differences of dimensionality that show up in different arrow principles across these spaces?

Continuing with earlier topics, *extending* the basic modal language of arrows also makes sense. E.g., in general arrow logic there may be many identity arrows, while in vector space there is only one identity element 0. To express this uniqueness, we need to move to some form of modal *difference logic* (cf. Chapter 4). Also, in mathematical

morphology, one finds a device for stating laws that are not valid in general, but only when we interpret some variables as standing for single vectors. An example is:

$$\begin{aligned} (X)_t - Y &= (X - Y)_t && \text{(MM-form)} \\ B \rightarrow (A + t) &\Leftrightarrow (B \rightarrow A) + t && \text{(LL-form)} \end{aligned}$$

From right to left, this is **LL** derivable as the general law $(S \longrightarrow X) \bullet Y \Rightarrow S \longrightarrow (X \bullet Y)$. The converse of this is not **LL** derivable, but it only works when Y is a singleton $\{t\}$. In the latter case, we have the special principle $S \Leftrightarrow (S + \{t\}) - \{t\}$, which we have to ‘inject’ into an otherwise fine **LL** derivation to get the desired result. This trick is exactly the same as using so-called *nominals* in extended modal logics, cf. [Areces, 2000], which are special proposition letters denoting just a single point. Other natural language extensions include an infinitary version of the addition modality \oplus , allowing us to close sets to *linear subspaces*.

Thus, the two fields are related, not just in their general structure, but also in their modus operandi, including tricks for boosting expressiveness. Of course, one would hope that the *algorithmic* content of arrow logics also makes sense under this connection, including its philosophy of ‘taming complexity’.

5.3.2.3 A worry about complexity

Issues of decidability and complexity have been largely ignored in this thesis. But one part of the ‘modal program’ is the balance between moderate expressive power and low complexity for various tasks: model checking, model comparison, and logical inference. In particular, arrow logics were originally designed to make the spectacular jump from undecidability in standard relational algebra to decidability. What happens to arrow logics in mathematical morphology? Even though the logic of the standard models appears to be effectively axiomatizable, i.e., recursively enumerable, *undecidability* is lurking! One bad omen is the validity of associativity, a danger sign in the arrow philosophy (cf. [van Benthem, 1996]).

Resorting to the tiling techniques introduced in [Harel, 1983], by encoding the problem of tiling the $\mathbb{N} \times \mathbb{N}$ grid in the arrow logic of vector spaces, one can show its undecidability. The idea of the proof is that of considering a denumerable set of colors C and a set of tiles $T = \{t_1, \dots, t_l\}$ (where each tile is a four-tuple of colors). Tiling is defined as a map $\rho : \mathbb{N} \times \mathbb{N} \rightarrow T$ such that the colors on touching edges coincide. The problem is known to be undecidable, [Robinson, 1971].

The tiling problem is encoded in the arrow logic of vector spaces as φ_T :

$$\begin{aligned}
 r \oplus u &= u \oplus r \\
 \top &= u \oplus \top \quad \text{where } u \leftarrow t_1 \vee \cdots \vee t_l \\
 \top &= r \oplus \top \quad \text{where } r \leftarrow t_1 \vee \cdots \vee t_l \\
 \bigwedge_{i=1}^l u \leftarrow t_i \ominus \bigvee \{t_j \mid \text{topColor}(t_i) = \text{bottomColor}(t_j)\} \\
 \bigwedge_{i=1}^l r \leftarrow t_i \ominus \bigvee \{t_j \mid \text{rightColor}(t_i) = \text{leftColor}(t_j)\}
 \end{aligned}$$

The argument to show undecidability becomes: T tiles $\mathbb{N} \times \mathbb{N}$ if and only if there exists a non-trivial vector space \mathbb{R}^k such that $\mathbb{R}^k \models \varphi_T$. The key of the proof is to show that indeed it is possible to encode the tiling problem in terms of the arrow logic of vector spaces with the formula φ_T .

We may have gone overboard in our desire to express the truth about vectors. Thus, the balance remains a continuing concern.

This chapter has shown modal structure in whichever direction one looks. There are natural fine-structured modal versions of affine and metric geometry, and linear algebra. These can be studied by general modal techniques—though much of the interest comes from paying attention to special spatial features. The benefits of this may be uniformity and greater sensitivity to expressive and computational fine-structure in theories of space. As a pleasant side-effect, a number of open problems of expressivity, complexity and complete axiomatization arise.

CHAPTER 6

A GAME-BASED SIMILARITY FOR IMAGE RETRIEVAL

6.1 Introduction

Image retrieval is concerned with the recovering of elements from a collection of images according to some set of desired properties. The properties of images are related to features which can be as diverse as textual annotations, color, texture, object shape, and spatial relationships among objects. The way the features from different images are compared, in order to have a measure of similarity among images, characterizes an image retrieval architecture.

Though quite a young field of computer vision, image retrieval already counts numerous frameworks, prototypes and commercial products. In [Smeulders et al., 2000], a method for systematizing approaches to image retrieval and a unifying framework for comparing systems is proposed. The work serves also as an excellent and up-to-date overview of the field. A similar purpose is served by the book [Del Bimbo, 1999].

The kind of topological relationship among objects we focus on are those at the qualitative level of mereotopology, that is, part-whole relations, topological relations and topological properties of individual regions. Other image retrieval systems are based on spatial relationships as the main retrieval feature. The work in [Tagare et al., 1995] is founded on transformation of Voronoi diagrams and that in [Petrakis et al., 2001] on graph matching. An older and known approach to image retrieval by spatial relationship is in [Chang and Liu, 1984]. This work considers the projections of regions onto two axes superimposed on the picture and simple interval relations over the projections over the axes. The approach suffers from not being orientation invariant and from the inability to deal with overlapping objects. On the positive side is the compactness of the topological representation of spatial relationships (called 2D strings). Other symbolic formalisms to handle qualitative topological relationships, which have been deployed for image retrieval, are those presented in [Egenhofer, 1991, Egenhofer and Franzosa, 1991, Del Bimbo et al., 1995]. Another trend in symbolic

approaches to image retrieval are the knowledge based ones. Originally concerned with the organization of knowledge *attached* to the image, the field has more recently seen efforts to organize and handle spatial information *obtained* from the image. Such recent systems [Russ et al., 1996, Aiello et al., 1999, Di Sciascio et al., 2000] make use of description logics, a modern reasoning tool closely related to modal logics.

Following the agenda in [Smeulders et al., 2000], we set the boundaries of our approach to image retrieval in the context of this chapter. Smeulders et al. identify five basic aspects of an image retrieval approach: image processing, features, interpretation and similarity, interaction, and system aspects. Here, we abstract from the image processing and from the system aspects, that is, we do not consider how to process images and extract features, nor how the images are stored and how systems are evaluated. We concentrate on a specific sort of features, we define for it a precise similarity measure and we consider only one set of modalities to interact with the system. Let us be more precise. The features we are interested in are the topological configurations of extended spatial entities and the topological spatial relations among different entities. The similarity is assessed with respect to a game theoretic comparison of the features. The interaction is based on query by example and query by sketch.

There are two requirements we desire to fulfill in our approach: on the one hand, the system should be based on a formal framework the properties of which must be well understood, on the other hand, the system should be actually implementable. The first part of the thesis is the place to dig for tools in order to satisfy the first requirement. We have seen a number of modal formalisms to handle space of which we have studied the formal properties such as their expressive power and their completeness. The second requirement must also be handled with care. We not only need correct representation and reasoning tools, we also need them to be compact and implementable. By implementable we intend both that the reasoning procedures should be decidable and should be decided in an amount of time acceptable for the average user. The inspiring model comes from the field of textual information retrieval (see for instance [Baeza-Yates and Ribeiro-Neto, 1999, Witten et al., 1999]): our aim is having a compact representations related to each picture such that all representations can be directly and rapidly compared in the retrieval phase.

The language we choose to express the main spatial information of an image is $\mathbf{S4}_u$, see Chapter 4. In the next section, we show why this language is adequate for expressing basic topological properties of patterns by highlighting its mereotopological strengths. Then we show how a similarity measure is built departing from the basic formal tools of the topo-approach. We can then identify a compact representation for images. Finally, we illustrate IRIS, a prototype based on the framework proposed.

6.2 A general framework for mereotopology

In Section 4.1, we have extensively studied the properties of the extended modal language $\mathbf{S4}_u$ in the context of the topo-approach. Before putting the language in action

on the task of image retrieval, we highlight its mereotopological expressive power. We now bring evidence to the claim that $S4_u$ is a general framework for mereotopological representation and reasoning. We also define our personal point of view on the connection with RCC identified in [Bennett, 1995].

6.2.1 Expressiveness

The language $S4_u$ is perfectly suited to express mereotopological concepts. The relation of parthood $P(A, B)$ of a region A being inside the region B holds whenever it is the case everywhere that A implies B :

$$P(A, B) := U(A \rightarrow B)$$

This captures exactly the set-inclusion relation of the models. As for connection C , two regions A and B are connected if there exists a point where both A and B are true:

$$C(A, B) := E(A \wedge B)$$

From here it is immediate to define all the basic eight RCC mereotopological predicates. Referring to Figure 6.1, let us recall the RCC8 relations (which we know to be definable in terms of $S4_u$, [Bennett, 1995]):

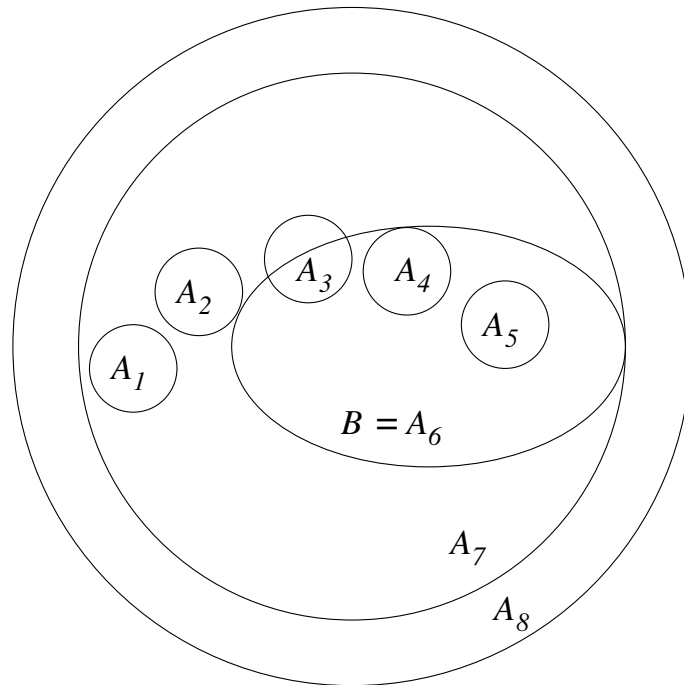


Figure 6.1: The RCC8 relations.

- A_1 is Disconnected from B ,
- A_2 is in External Connection with B ,
- A_3 Overlaps with B ,
- A_4 is Tangential P of B ,
- A_5 is Proper P of B ,
- A_6 is Equal to B ,
- A_7 contains tangentially (Tangential P^{-1}) B ,
- A_8 contains (Proper P^{-1}) B ,

Notice that the choice made in defining P and C is arbitrary. So, why not take a more restrictive definition of parthood? Say, A is part of B whenever the closure of A is contained in the interior of B ?

$$P(A, B) := U(\diamond A \rightarrow \Box B)$$

As this formula shows, $\mathbf{S4}_u$ is expressive enough to capture also this definition of parthood. In [Cohn and Varzi, 1998], the logical space of mereotopological theories is systematized. Based on the intended interpretation of the connection predicate C, and the consequent interpretation of P (and fusion operation), a type is assigned to mereotopological theories. More precisely, a *type* is a triple $\tau = \langle i, j, k \rangle$, where the first i refers to the adopted definition of C_i , j to that of P_j and k to the sort of fusion. The index i , referring to the connection predicate C, accounts for the different definition of connection at the topological level. Using $\mathbf{S4}_u$ one can repeat here the three types of connection:

$$\begin{aligned} C_1(A, B) &:= E(A \wedge B) \\ C_2(A, B) &:= E(A \wedge \diamond B) \vee E(\diamond A \wedge B) \\ C_3(A, B) &:= E(\diamond A \wedge \diamond B) \end{aligned}$$

Looking at previous mereotopological literature, one remarks that RCC uses a C_3 definition, while the system proposed in [Asher and Vieu, 1995] uses a C_1 . Similarly to connectedness, one can distinguish various types of parthood, again in terms of $\mathbf{S4}_u$:

$$\begin{aligned} P_1(A, B) &:= U(A \rightarrow B) \\ P_2(A, B) &:= U(A \rightarrow \diamond B) \\ P_3(A, B) &:= U(\diamond A \rightarrow \diamond B) \end{aligned}$$

In [Cohn and Varzi, 1998], the definitions of the C_i are given directly in terms of topology, and the definitions of P_j in terms of a first order language with the addition of a predicate C_i . Finally, a general fusion ϕ_k is defined in terms of a first order language

with a C_i predicate. Fusion operations are like algebraic operations on regions, such as adding two regions (product), or subtracting two regions. One cannot repeat the general definition given in [Cohn and Varzi, 1998] at the $S4_u$ level. Anyhow, one can show that various instances of fusion operations are expressible. E.g., the product $A \times_k B$:

$$\begin{aligned} A \times_1 B &:= A \wedge B \\ A \times_2 B &:= (\diamond A \wedge B) \vee (A \wedge \diamond B) \\ A \times_3 B &:= (\diamond A \wedge \diamond B) \end{aligned}$$

Usually uniform theories are found in the literature, that is, theories that combine definitions of C_i, P_i , and \times_i with the same index i . Though, there are some exceptions, e.g, [Cartwright, 1975] uses a C_2, P_1 combination. Non uniform theories separate the topological part from the purely mereological one requiring the definition of parthood and connection to be independent. Clearly, P_i cannot be defined in terms of C_j if $i \neq j$.

The above discussion has shown that $S4_u$ is a general language for mereotopology. All the different types $\tau = \langle i, j, k \rangle$ of mereotopological theories are expressible. Incidentally, notice that a mereotopological theory of space may combine definition of parthood and connection with different indices. For instance, it is possible to have a C_1, P_2 mereotopological theory.

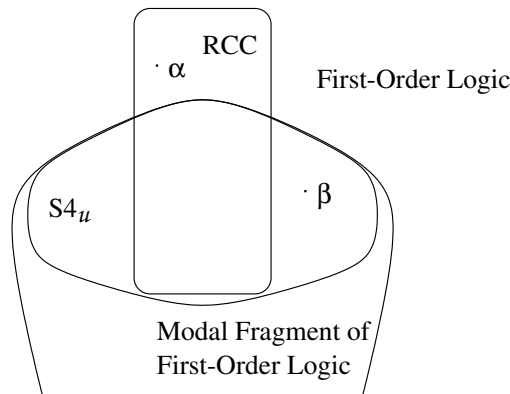


Figure 6.2: The positioning of $S4_u$ and RCC with respect to well-known logics.

The language $S4_u$ is a multi-modal language with nice computational properties. It is complete with respect to topological models, it is decidable, it has the finite model property. It captures a large and “well-behaved” fragment of mereotopology, though, it is not a first-order language. In other words, it is not possible to quantify over regions. A comparison with the best-known RCC is in order.

6.2.2 Comparison with RCC

RCC is a first order language with a distinguished connection predicate C_3 . The driving idea behind this qualitative theory of space is that regions of space are primitive objects

and connection is the basic predicate. This reflects in the main difference between RCC and the proposed system, which on the contrary builds on point-based topology.

RCC and $\mathbf{S4}_u$ capture different portions of mereotopology.

To show this, two formulas are given: an RCC formula which is not expressible in $\mathbf{S4}_u$ and, vice-versa, one expressible in $\mathbf{S4}_u$, but not in RCC. The situation is depicted in Figure 6.2. In RCC, one can write:

$$\forall A \exists B : P(A, B) \quad (\alpha)$$

meaning that every region is part of another one (think of the entire space). On the other hand, one can write a formula such as:

$$U(p \leftrightarrow \diamond \Box p) \quad (\beta)$$

which expresses the regularity of the region p . It is easy to see that α is not expressible in $\mathbf{S4}_u$ and that β is not in RCC.

This fact may be misleading. It is neither the motivations, nor the core philosophical intuitions that draw the line between RCC and $\mathbf{S4}_u$. Rather, it is the logical apparatus which makes the difference. To boost the similarities, consider again how the main predicates of RCC can be expressed within $\mathbf{S4}_u$. Indeed one can define the same predicates as RCC8. However, as remarked before the nature of the approach is quite different. Take for instance the non tangential part predicate. In RCC it is defined by means of the non existence of a third entity C :

$$\text{NTTP}(A, B) \text{ iff } P(A, B) \wedge \neg P(B, A) \wedge \neg \exists C [EC(C, A) \wedge EC(C, B)]$$

On the other hand, in $\mathbf{S4}_u$ it is simply a matter of topological operations. As in the previous table, for $\text{NTTP}(A, B)$ it is sufficient to take the interior of the containing region $\Box B$, the closure of the contained region $\diamond A$ and check if all points that satisfy the latter $\diamond A$ also satisfy the former $\Box B$.

6.3 Comparing spatial patterns

At the beginning of the chapter, we introduced the problem of image retrieval and its relying on similarity measures. Then, we advocated the adequacy of $\mathbf{S4}_u$ as a general language of mereotopology. We take the view that one should use $\mathbf{S4}_u$ to talk about spatial patterns in the context of image retrieval. Now there is a technical question. How does one answer questions such as *When are two spatial patterns the same?*, *When is a pattern a sub-pattern of another one?*, and, most importantly, *How different are two spatial patterns?*

6.3.1 Model comparison games distance

The answer to these questions comes from looking at the tools of the topo-approach with a different twist. Consider the definition of topo-bisimulation and topo-game, Definition 4.1.2 and Definition 4.1.5. Topo-bisimulations are an equivalence relation, so one may very well use them to define identity of patterns. Via simulations, one can also consider issues of a pattern being a sub-pattern of another one. Then, topo-games were introduced as a refining notion of topo-bisimulations. Therefore, one may use topo-games to define a measure of difference among spatial patterns. Think of it this way. The less it takes Spoiler to win a game, the more different must the spatial patterns be, the more *unsimilar*. On the opposite, the longer Duplicator can resist, the more *similar* are the spatial patterns. In the limit, if Duplicator can resist forever, i.e., in the infinite round game, the two patterns are topologically bisimilar. Now comes the technical problem. Topo-games are defined as a way of comparing two given topological models, exactly in the spirit of the original definition of first-order model comparison games à la Ehrenfeucht-Fraïssé, but we need a similarity measure on the whole class of models; we need a measure that behaves uniformly across all models for $\mathbf{S4}_u$.

The first intuition on turning model comparison games into a similarity measure may be misleading in a pessimistic direction. To get to a similarity measure, we need to define a distance in terms of topo-games. Distances require considering more than just two models at a time. Consider, for example, three models and the three model comparison games that can be played. The formulas, the points and open sets picked in the three games may be completely unrelated one game from each other, therefore, one may be discouraged and conjecture that model comparison games are not related across different models of the same class.

Even though the remark on the unrelatedness of the strategies for different games is true. It turns out that there is still an interrelation between model comparison games over two given models and the whole class. Most importantly, the relation can be defined to satisfy the three properties defining a distance measure. Here is how.

6.3.1. DEFINITION (ISOSCELES TOPO-DISTANCE). Consider the space of all topological models T . *Spoiler's shortest possible win* is the function $spw : T \times T \rightarrow \mathbb{N} \cup \{\infty\}$, defined as:

$$spw(X_1, X_2) = \begin{cases} n & \text{if Spoiler has a winning strategy in } TG(X_1, X_2, n), \\ & \text{but not in } TG(X_1, X_2, n-1) \\ \infty & \text{if Spoiler does not have a winning strategy in} \\ & TG(X_1, X_2, \infty) \end{cases}$$

The *isosceles topo-model distance* (*topo-distance*, for short) between X_1 and X_2 is the function $tmd: T \times T \rightarrow [0, 1]$ defined as:

$$tmd(X_1, X_2) = \frac{1}{spw(X_1, X_2)}$$

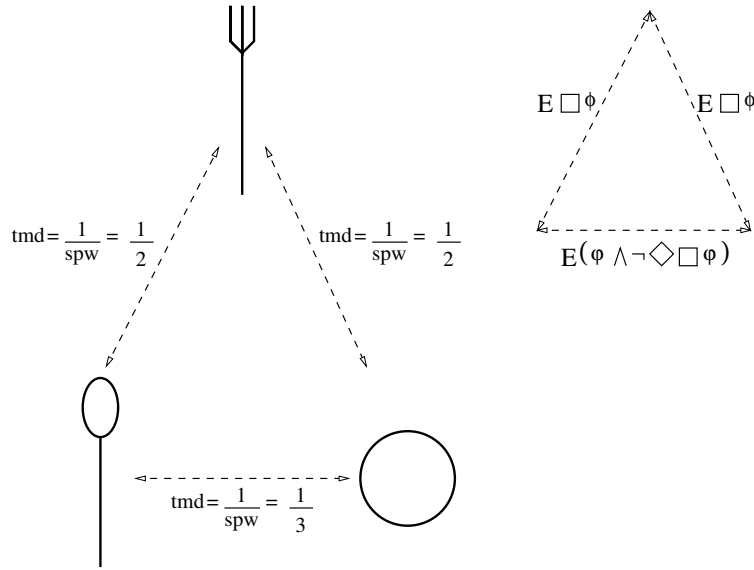


Figure 6.3: On the left, three models and their relative distance. On the right, the distinguishing formulas.

The distance was named ‘isosceles’ since it satisfies the triangular property in a peculiar manner. Given three models, two of the distances among them (two sides of the triangle) are always the same and the remaining distance (the other side of the triangle) is smaller or equal. On the left of Figure 6.3, three models are displayed: a spoon, a fork and a plate. Think these cutlery objects as subsets of a dense space, such as the real plane, which evaluate to ϕ , while the background of the items evaluates to $\neg\phi$. The isosceles topo-distance is displayed on the left next to the arrow connecting two models. For instance, the distance between the fork and the spoon is $\frac{1}{2}$ since the minimum number of rounds that Spoiler needs to win the game is 2. To see this, consider the formula $E\Box\phi$, which is true on the spoon (there exists an interior point of the region ϕ associated with the spoon) but not on the fork (which has no interior points). On the right of the figure, the formulas used by spoiler to win the three games between the fork, the spoon and the plate are shown. Next the proof that tmd is a distance function, in particular the triangular property, exemplified in Figure 6.3, is always satisfied by any three topological models.

6.3.2. THEOREM (ISOSCELES TOPO-MODEL DISTANCE). tmd is a distance measure on the space of all topological models.

Proof tmd satisfies the three properties of distances; i.e., for all $X_1, X_2 \in T$:

- (i) $tmd(X_1, X_2) \geq 0$ and $tmd(X_1, X_2) = 0$ iff $X_1 = X_2$
- (ii) $tmd(X_1, X_2) = tmd(X_2, X_1)$

$$(iii) \quad tmd(X_1, X_2) + tmd(X_2, X_3) \geq tmd(X_1, X_3)$$

As for (i), from the definition of topo-games it follows that the amount of rounds that can be played is a positive quantity. Furthermore, the interpretation of $X_1 = X_2$ is that the spaces X_1, X_2 satisfy the same modal formulas. If Spoiler does not have a w.s. in $\lim_{n \rightarrow \infty} TG(X_1, X_2, n)$ then X_1, X_2 satisfy the same modal formulas. Thus, one correctly gets

$$tmd(X_1, X_2) = \lim_{n \rightarrow \infty} \frac{1}{n} = 0.$$

Equation (ii), since for all X_1, X_2 , then $TG(X_1, X_2, n) = TG(X_2, X_1, n)$.

As for (iii), the triangular property, consider any three models X_1, X_2, X_3 and the three games playable on them,

$$TG(X_1, X_2, n), TG(X_2, X_3, n), TG(X_1, X_3, n) \quad (6.1)$$

Two cases are possible. Either Spoiler does not have a winning strategy in all 3 games (6.1) for any amount of rounds, or he has a winning strategy in at least one game.

If Spoiler does not have a winning strategy in all the games (6.1) for any number of rounds n , then Duplicator has a winning strategy in all games (6.1). Therefore, the three models satisfy the same modal formulas, $spw \rightarrow \infty$, and $tmd \rightarrow 0$. Trivially, the triangular property (iii) is satisfied.

Suppose Spoiler has a winning strategy in one of the games (6.1). Via Theorem 4.1.6 (adequacy), one can shift the reasoning from games to formulas: there exists a modal formula γ of multi-modal rank m such that $X_i \models \gamma$ and $X_j \models \neg\gamma$. Without loss of generality, one can think of γ as being in normal form:

$$\gamma = \bigvee \bigwedge [\neg] U(\varphi_{\mathbf{S4}}) \quad (6.2)$$

This last step is granted by the fact that every formula φ of $\mathbf{S4}_u$ has an equivalent one in normal form whose modal rank is equivalent or smaller to that of φ .¹ Let γ^* be the formula with minimal multi-modal depth m^* with the property: $X_i \models \gamma^*$ and $X_j \models \neg\gamma^*$. Now, the other model X_k either satisfies γ^* or its negation. Without loss of generality, $X_k \models \gamma^*$ and therefore X_j and X_k are distinguished by a formula of depth m^* . Suppose X_j and X_k to be distinguished by a formula β of multi-modal rank $h < m^*$: $X_j \models \beta$ and $X_k \models \neg\beta$. By the minimality of m^* , one has that $X_i \models \beta$, and hence, X_i and X_k can be distinguished at depth h . As this argument is symmetric, it shows that either

- one model is at distance $\frac{1}{m^*}$ from the other two models, which are at distance $\frac{1}{i}$ ($\leq \frac{1}{m^*}$), or

¹In the proof, the availability of the normal form is not strictly necessary, but it gives a better impression of the behavior of the language, see Section 4.1.

- one model is at distance $\frac{1}{h}$ from the other two models, which are at distance $\frac{1}{m^*}$ ($\leq \frac{1}{h}$) one from the other.

It is a simple matter of algebraic manipulation to check that m^*, l and h, m^* (as in the two cases above), always satisfy the triangular inequality. QED

The nature of the isosceles topo-distance triggers a question. Given three spatial models, why is the distance between two pairs of them always the same?

First an example, consider a spoon, a chop-stick and a sculpture by Henry Moore. It is immediate to distinguish, via *tmd*, the Moore's sculpture from the spoon and from the chop-stick. The distance between them is high and the same. On the other hand, the spoon and the chop-stick look much more similar, thus, their distance is much smaller. Mereotopologically, it may even be impossible to distinguish them (null distance).

In fact one is dealing with models of a qualitative spatial reasoning language of mereotopology. Given three models, via the isosceles topo-distance, one can easily distinguish the very different patterns. In some sense they are far apart as if they were belonging to different equivalence classes. Then, to distinguish the remaining two can only be harder, or equivalently, the distance can only be smaller.

The division in classes of equivalence and the isosceles nature of the topo-distance should not be interpreted as the topo-distance having only a finite number of values. In general, the distance between any two patterns can be any value between 1 and $\frac{1}{n}$ with $n \in \mathbb{N}$. One way of seeing this is considering two non-equivalent $\mathbf{S4}_u$ formulas. Such formulas can be chosen of any modal depth. Therefore, the distance could have any value in the interval $[0, 1]$. What is true is that the distribution of the values is not linear in the interval $[0, 1]$, but rather it becomes increasingly more dense towards 0.

6.4 Computing similarities

The definition of a distance based on model comparison games is an important step, but how can we complete our journey towards practice? We need to compute the topo-distance. First, we give a general methodology, then we provide an algorithm for the concrete case that is of most interest to us.

6.4.1 Methodology

A general methodology for the computation of the topo-distance among two topo-models M, N might work as follows. First, one translates the topo-models M, N into equivalent Kripke models (as we did in Section 2.1.3), then one checks the models for traditional bisimilarity [Dovier et al., 2001]. If the models are not bisimilar, one checks all the points for which a bisimulation can not be established. The inverse of the minimal modal depth of the formulas distinguishing these points is the topo-distance.

An alternative and more direct approach is that of relating the points in the topological spaces. The worry with topological semantics might be an exponential explosion

due to the number of open sets (rather than the accessibility relation for ordinary Kripke models). Still, one might measure input complexity in terms of both points and opens, and still get a polynomial time algorithm to compute the topo-distance.

Here we shall not enter into the details of such general methods for computing topo-distance, but rather concentrate on a specific case of practical interest.

6.4.2 Polygons of the plane

We make an ontological commitment to finite polygons of the real plane. This is common practice in various application domains such as geographical information systems (GIS), in many branches of image retrieval and of computer vision, or in robot planning, just to mention the most common.

Consider the real plane \mathbb{R}^2 , any line in \mathbb{R}^2 cuts it into two half-planes. We call a half-plane *closed* if it includes the cutting line, *open* otherwise.

6.4.1. DEFINITION (REGION). A *polygon* is the intersection of finitely many open and closed half-planes. An *atomic region* of \mathbb{R}^2 is the union of finitely many polygons.

An atomic region is denoted by one propositional letter. More in general, any set of atomic regions, simply called *region*, is denoted by a $\mathbf{S4}_u$ formula. The polygons of the plane equipped with a valuation function, denoted by $M_{\mathbb{R}^2}$, are in full rights a topological model as in Definition 2.1.1, a basic topological fact. A similar definition of region can be found in [Pratt and Lemon, 1997]. In that article Pratt and Lemon also provide a collection of fundamental results regarding the plane, polygonal ontology just defined (actually one in which the regions are open regular).

From a model theoretic point of view, the advantage of working with $M_{\mathbb{R}^2}$ is that we can prove a logical finiteness result and thus give a terminating algorithm to compute the topo-distance between any two regions.

6.4.2.1 Finiteness

In general, there are infinitely many non equivalent $\mathbf{S4}$ formulas and one can identify appropriate Kripke models to show this (cf. [Blackburn et al., 2001]). In Section 3.4.1, however we have seen how finite unions of convex intervals of the real line yield a finite number (64) of modally different formulas (Theorem 3.4.10). Similar results (though with a larger upper bound) hold for the plane where in place of intervals one considers rectangles, cf. Section 3.4.4. The further extension needed here is to move from such rectangles to generic polygons with a finite number of sides.

First, let us consider an example. Figure 6.4.a shows a model composed of two closed polygons: one denoted by r and one by q . Relevant points of the union of these two polygons are those on the frontiers, on the intersections of the frontiers and in the interiors of each polygon. A distinguishing formula of minimal modal rank true at each of these relevant points is shown in Figure 6.4.b.

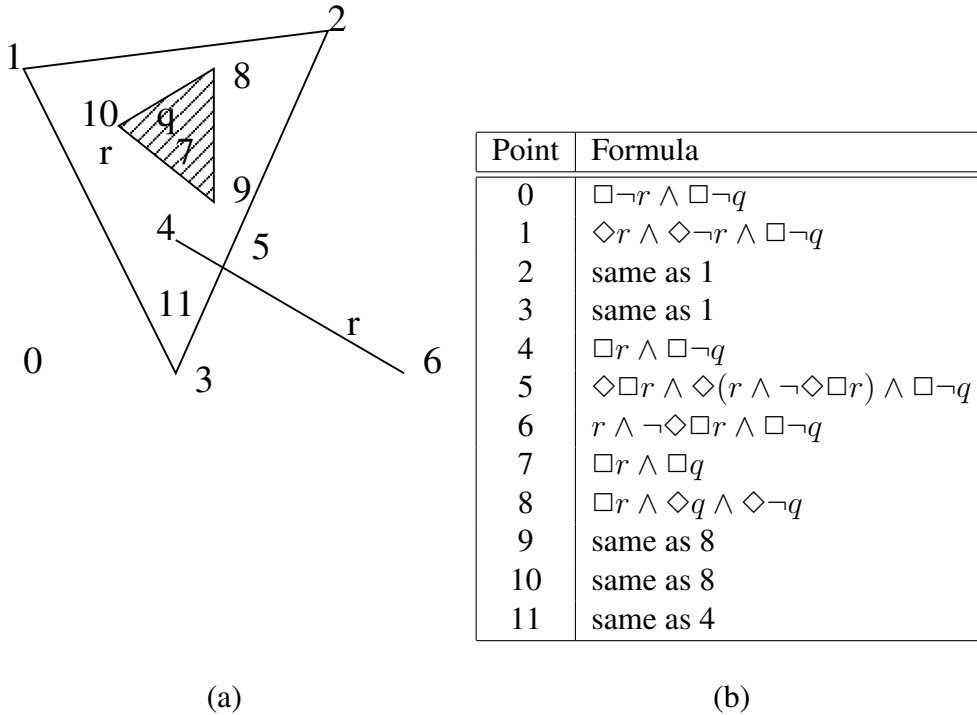


Figure 6.4: (a) A simple polygonal model of the plane; (b) relevant formulas.

Given the limitation on finiteness of the number of polygons of the space in our definition of a polygonal model of the plane, one can get a grasp of why there are only finitely many definable formulas. We shall not give a precise proof here: it would go more or less like the one in Section 3.4.4, but now taking oblique orientations into account, instead we sketch some concrete steps toward the result.

6.4.2. LEMMA (FINITENESS). There are only finitely many modally definable subsets starting from any finite set of regions viewed as atoms.

We work by enumerating cases, i.e., considering Boolean combinations of planes, adding to an ‘empty’ space one half-plane at the time, first to build one region r , and then to build a finite set of regions. The goal is to show that only finitely many possibilities exist. We begin by placing a closed half plane denoted by r on an empty bidimensional space, Figure 6.5.a. Let us follow what happens to points in the space from left to right. On the left, points satisfy the formula $\Box r$. This is true until we reach the closed frontier point of the half-plane, where $\Diamond r \wedge \Diamond\neg r \wedge r$ holds. Left of the frontier, the points satisfy the formula $\Box\neg r$. Similarly the formulas are defined for the negate region in Figure 6.5.–a, notice that this time the polygon is open. In fact, by considering negation the roles of r and $\neg r$ switch. Consider now a second plane in the picture:

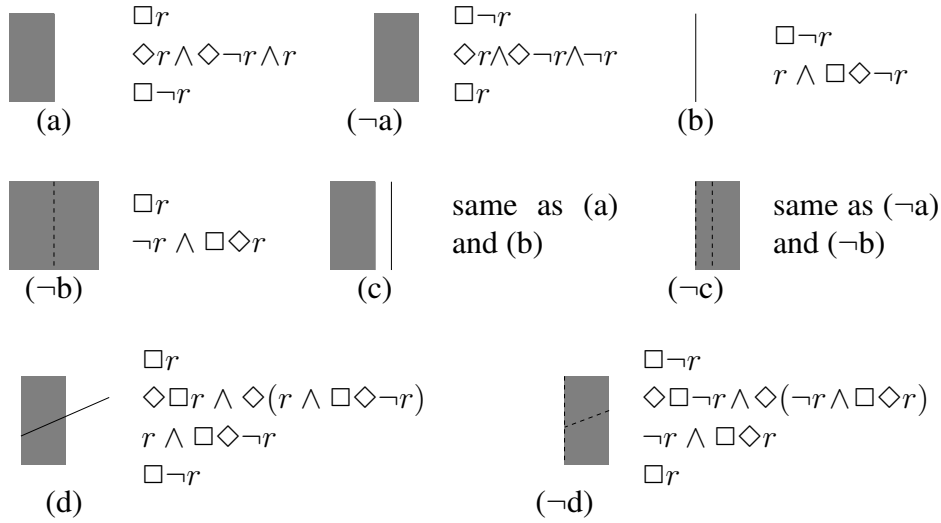
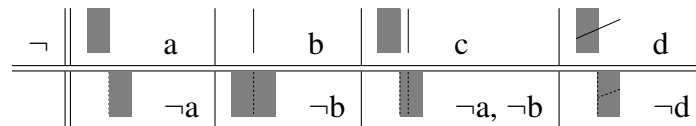


Figure 6.5: Basic formulas defined by one region r .









- **Intersection:** the intersection may be empty (no new formula), may be a polygon with two sides and vertices (no new formula, the same situation as with one polygon), or it may be a line, the case of two closed polygons that share the side (in this last case depicted in Figure 6.5.b—*spike*—we have a new formula, namely, $r \wedge \square \diamond \neg r$).
- **Union:** the union may be a polygon with either one or two sides (no new formula), two separated polygons (no new formula), or two open polygons sharing the open side (this last case depicted in Figure 6.5.¬b—*crack*—is like the spike, one inverts the roles r and $\neg r$ in the formula: $\neg r \wedge \square \diamond r$).

Finally, consider combining cases (a) and (b). By union, we get Figure 6.5.a, 6.5.c, 6.5.d. The only situation bringing new formulas is the latter. In particular, the point where the line intersects the plane satisfies the formula: $\diamond \square r \wedge \diamond (r \wedge \square \diamond \neg r)$. By intersection, we get a segment, or the empty space, thus, no new formula.









The four basic configurations just identified yield no new configuration from the $S4_u$ point of view. To see this, consider the Boolean combinations of the above configurations. We begin by negation (complement):



Union straightforwardly follows (where a stands for both a and ¬a, as both configurations always appear together):

U	 a	 b	 c	 d
 a	a, ¬b, ¬d	a, c, d	a, ¬b, c, d, ¬d	a, ¬b, d, ¬d
 b	a, c, d	b	c, d	d
 c	a, ¬b, c, d, ¬d	c, d	a, ¬b, c, d, ¬d	a, ¬b, c, d, ¬d
 d	a, ¬b, d, ¬d	d	a, ¬b, c, d, ¬d	a, ¬b, d, ¬d

The table for intersection follows, with the proviso that the combination of the two regions can always be empty (not reported in the table) and again a and ¬a are represented simply by a:

\cap	 a	 b	 c	 d
 a	a, b, c, d	b	a, b, c	a, b, d
 b	b	b	b	b
 c	a, b, c	b	a, b, c, d	a, b, c, d
 d	a, b, d	b	a, b, c, d	a, b, c, d

We have thus gave a taste for the finiteness of the polygons of the real plane for **S4**. But we have designed the whole topo-distance for the richer language **S4_u**. This extension is not a problem. Recalling the availability of a normal form for **S4_u** (Section 4.1), one sees that the finiteness result simply extends. The formulas above are simply preceded by an existential operator stating the existence of such a point in the model. For instance in case a we have: $E\Box r$, $E(\Diamond r \wedge \Diamond \neg r \wedge r)$, and $E(\Box \neg r)$.

Since the information related to a region is finite, we can compactly represent it. We call *topo-vector* associated with the region r , notation \vec{r} , an ordered sequence of Boolean values. The values represent whether the region r satisfies or not a fixed sequence of **S4_u** formulas:

$$\boxed{E\Box r \mid E\Box \neg r \mid \dots \mid E(\neg r \wedge \Box \Diamond r) \mid \dots \mid E(\Diamond \Box \neg r \wedge \Diamond(\neg r \wedge \Box \Diamond r))}$$

The formulas are those identified in Figure 6.5 preceded by an existential operator. For example, the topo-vector associated with a plate—a closed square r in the plane—is:

$$\boxed{\text{true} \mid \text{true} \mid \dots \mid \text{false} \mid \dots \mid \text{false}}$$

Adding half-planes with different denotations r_2, r_3, \dots increases the number of defined formulas. The definition of topo-vector is extended to an entire $M_{\mathbb{R}^2}$ model.

The topo-vector is built such that the modal rank of the formulas is not decreasing going from the positions with lower index to those with higher. The size of the topo-vector grows by $c \cdot 2^i$, where c is a constant value and i is the number of the regions in the model. This might seem a serious drawback. But then, the topo-vector has to store all the information relevant for $\mathbf{S4}_u$ about the model. Furthermore, its size is often considerably smaller than that of the whole topological model. As a final consideration, one should add that in practical situations the size of topo-vectors still remains manageable.

6.4.3 The topo-distance algorithm

The topo-vector is a compact representation of a spatial pattern. One way of seeing this is the following. Take any spatial pattern. Reduce it to the smallest topo-bisimilar model, using the technique of Section 2.1.3. Consider all definable formulas from the proposition letters present in the pattern and consider whether each formula is true somewhere in the reduced model. This information is collected in the topo-vector.

Before giving an algorithm to compute the topo-vector for $M_{\mathbb{R}^2}$, and in turn the topo-distance, let us reconsider the example of Figure 6.4. The region q contributes to the topo-vector only with three possible behaviors. Either the points are outside of it $\Box\neg q$, or they are q points. In the latter case all these points are also inside $\Box r$ and one only distinguishes the case for which the points are in the interior or on the boundary. For the region r there is a bit more variety, because there is a spike. The spike contributes because it yields a non regular portion, and because it intersects a regular region. Summarizing, the topo-vector for the region looks like this:

$E(\Box r \wedge \Box q)$	$E(\Box r \wedge \Box\neg q)$	$E(\Box\neg r \wedge \Box q)$	$E(\Box\neg r \wedge \Box\neg q)$...
true (7)	true (4)	false	true (0)	...

...	$r \wedge \neg\Diamond\Box r \wedge \Box q$	$r \wedge \neg\Diamond\Box r \wedge \Box\neg q$...
...	false	true (6)	...

where we have marked the point satisfying the part of the formula after the existential operator. The name of the points refers to Figure 6.4.

Next we present the algorithm to compute the topo-vector for a generic pattern of $M_{\mathbb{R}^2}$ with respect to $\mathbf{S4}_u$. Given an $M_{\mathbb{R}^2}$ model M , $\text{topo-vector}(M)$ returns the topo-vector associated with M .

```

topo-vector( $M$ )
   $\vec{v} \leftarrow$  initialized to all false values
  loop on regions  $r$  of  $M$  with index  $i$ 
    loop on atomic regions  $a$  of  $r(i)$  with index  $j$ 
      loop on vertices  $v$  of  $a(j)$  with index  $k$ 
        update  $\vec{v}$  with the  $\bigwedge \square[\neg]g$  information
        update  $\vec{v}$  with the point  $v(k)$ 
        if  $v(k)$  is not free
          update  $\vec{v}$  with the  $\bigwedge \square[\neg]g$  information
          loop on intersections  $x$  of  $a(j)$  with all
            regions of  $M$  with index  $l$ 
            update  $\vec{v}$  with the point  $x(l)$ 
        return  $\vec{v}$ 
  return  $\vec{v}$ 

```

If a point $v(k)$ of an atomic region $a(j)$ is contained in any polygon different from $a(j)$ and it is not contained in any other region, then the condition $v(k)$ is not free is satisfied. Standard computational geometry algorithms exist for this task, [de Berg et al., 2000]. When the “update \vec{v} with the point p ” function is called, one checks in which case p is (as shown after Lemma 6.4.2), then one considers the position in the topo-vector corresponding to the formula satisfied by the point. Then one sets the values for that entry to true.² When the “update \vec{v} with the $\bigwedge \square[\neg]g$ ” function is called, one checks in the interior of which regions the current point is and updates accordingly the $\square g_1 \wedge \square \neg g_2 \wedge \dots$ formulas (e.g., those for the points 0, 4, 7, and 11 in Figure 6.4).

Consider again the simple model of Figure 6.4, repeated in Figure 6.6 for convenience. After initialization, the region r is considered and one starts looping on the vertices of its polygons, first the point 1. The point is `free`, it is the vertex of a full polygon (not a segment) and therefore the topo-vector is updated with a true value in the positions corresponding to $E(\square r \wedge \square \neg q)$, $E(r \wedge \diamond r \wedge \diamond \neg r \wedge \square \neg q)$, and $E(\square \neg r \wedge \square \neg q)$. The points 2 and 3 update the values for the same formulas and thus have no effect. The point 4 falls inside the first polygon of r , the topo-vector does not need update. Intersections are then computed and the point 5 is found. The point needs to update the vector for the formula $E(\diamond \square r \wedge \diamond (r \wedge \neg \diamond \square r) \wedge \square \neg q)$. The point 6 is considered and the point needs to update the formula $E(r \wedge \neg \diamond \square r \wedge \square \neg q)$. The algorithm proceeds by considering the second region, q and its vertices 8, 9, and 10. The three vertices all fall inside the region r and provide for the satisfaction of the formulas $E(\square r \wedge \square q)$, $E(\square r \wedge \diamond q \wedge \diamond \neg q)$, and $E(\square r \wedge \square \neg q)$.

²An obvious optimization to the algorithm is to avoid checking points for which all the associated entries are already set to true.

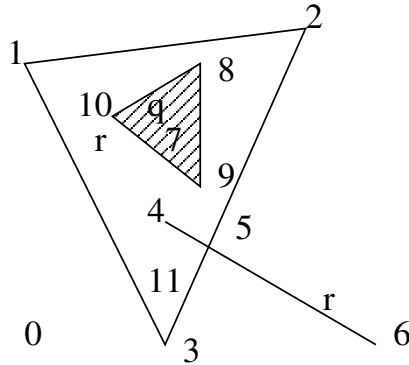


Figure 6.6: Computing the topo-vector on a simple model.

The final goal is to have an algorithm to compute the topo-distance not simply the topo-vector. One can compute the topo-distance among two models by comparing the respective topo-vectors. Here is the algorithm taking as input two $M_{\mathbb{R}^2}$ models M_1, M_2 and outputting the value of the topo-distance between them.

```

topo-distance ( $M_1, M_2$ )
   $\vec{v}_1 \leftarrow$  topo-vector ( $M_1$ )
   $\vec{v}_2 \leftarrow$  topo-vector ( $M_2$ )
  align  $\vec{v}_1$  and  $\vec{v}_2$ 
  loop on  $\vec{v}_1 \vec{v}_2$  with index  $i$ 
    if  $\vec{v}_1(i) \neq \vec{v}_2(i)$ 
      return  $\frac{1}{\text{modal rank}(\vec{v}_1(i))}$ 
  return 0

```

The idea is to retrieve the topo-vectors associated with the two input models and then loop over their elements. The inequality check can also be thought of as a **xor**, since the elements of the array are Booleans. If the condition is never satisfied, the two topo-vectors are identical, the two-models are topo-bisimilar and thus the topo-distance is zero. The `align` command makes the topo-vectors of the same length and aligns the formulas of the two in a way such that to the same index in the vector corresponds the same formula. If a topo-vector contains a formula that the other one does not, the entry is added to the vector missing it with a false value.

The basic properties of the topo-distance algorithm are the following.

6.4.3. LEMMA (TERMINATION). The topo-distance algorithm terminates.

The property is easily shown by noticing that a segment (a side of a polygon) can have at most one intersection with any other segment, that the number of polygons forming a region of $M_{\mathbb{R}^2}$ is finite, and that the number of regions of $M_{\mathbb{R}^2}$ is finite.

6.4.4. LEMMA (CORRECTNESS). For any $M, N \in M_{\mathbb{R}^2}$, $\text{topo-distance}(M, N) = k$ iff the actual topo-distance between M and N is k .

Proof First, consider the case of bisimilar models. If M and N are topo-bisimilar, by Definition 6.3.1 the topo-distance is 0. By Theorem 4.1.3, if the models are topo-bisimilar they satisfy the same modal formulas, thus, $\text{topo-vector}(M)$ and $\text{topo-vector}(N)$ are identical, which in turn means that $\text{topo-distance}(M, N) = 0$.

As for the other direction, if the $\text{topo-distance}(M, N) = 0$ then the topo-vectors are identical. But since the topo-vectors comprise all non-equivalent modal formulas for M and N , and since they are finite, Theorem 4.1.3 may be applied to the finite bisimulation contractions. That is, the latter models are topo-bisimilar. By Definition 6.3.1, if the models are topo-bisimilar, then topo-distance is 0.

Second, consider the case in which the models are not topo-bisimilar. The idea is similar. One uses the adequacy theorem for model comparison topo-games in place of the theorems for topo-bisimulations.

If the distance between M and N is $k > 0$, by Definition 6.3.1, Spoiler has a winning strategy in any game of length at least $\frac{1}{k}$. By Theorem 4.1.6, all the entries for the formulas of modal depth smaller than $\frac{1}{k}$ in the $\text{topo-vector}(M)$ and $\text{topo-vector}(N)$ must be the same. Since the topo-vectors comprise all non-equivalent modal formulas for M and N , there must be an entry for at least one modal formula of depth $\frac{1}{k}$ which differentiate the two topo-vectors allowing Spoiler to win. By the topo-distance algorithm, this means that $\text{topo-distance}(M, N) = k$.

As for the other direction, suppose that the $\text{topo-distance}(M, N) = k > 0$, then the topo-vectors must be identical for all entries associated to formulas of modal depth smaller than $\frac{1}{k}$, and there must be a difference for at least one entry associated with a formula of modal depth $\frac{1}{k}$. Since the topo-vectors comprise all non-equivalent modal formulas of minimal modal rank for M and N , the differentiating formula of minimal modal rank for M and N has modal rank $\frac{1}{k}$. By Theorem 4.1.6, this means that Spoiler's shortest winning strategy needs exactly $\frac{1}{k}$ rounds. By Definition 6.3.1, the latter implies that the topo-distance between M and N is k . QED

By Lemma 6.4.3 and Lemma 6.4.4, we obtain the following result.

6.4.5. THEOREM (DECIDABILITY OF THE TOPO-DISTANCE). In the case of polygonal topological models $M_{\mathbb{R}^2}$ over the real plane, the problem of computing the topo-distance among any two models is decidable.

Given our further definitions, and the connection between Duplicator's winning strategies in infinite topo-games and topo-bisimulations, (cf. [Barwise and Moss, 1996]), we also have the following result.

6.4.6. COROLLARY (DECIDABILITY OF TOPO-BISIMULATIONS). In the case of polygonal topological models over the real plane, the problem of identifying whether two models are topo-bisimilar is decidable.

6.5 The IRIS prototype

The ultimate step toward practice of the spatial framework presented in the chapter is the actual implementation of the similarity measure in a prototype. The topo-distance is a building block of an image retrieval system, named IRIS Image RetrIeval based on Spatial relationships, coded in Java and enjoying a Swing interface (Figure 6.8). An overview of the programming behind IRIS with the presentation of the most relevant source code passages can be found in Appendix C.

The actual similarity measure is built in IRIS to both index and retrieve images on the basis of:

- (i) The spatial intricacy of each region,
- (ii) The binary spatial relationships between regions, and
- (iii) The textual description accompanying the image.

Referring to Figure 6.7, one can get a glimpse of the conceptual organization of IRIS. A spatial model, as in Definition 2.1.1, and a textual description (central portion of the figure) are associated with each image of the collection (on the left). Each topological model is represented by its topo-distance vector, as built by the algorithm in Section 6.4 and by a matrix of binary relationships holding between regions. Similarly, each textual description is indexed holding a representative textual vector of the text (right portion of the figure). In Figure 6.8, a screen-shot from IRIS after querying a database of about 50 images of men and cars is shown. On the top-right is the window for sketching queries. The top-center window serves to write textual queries and to attach information to the sketched regions. The bottom window shows the results of the query with the thumbnails of the retrieved images (left to right are the most similar). Finally, the window on the top-left controls the session.

We remark again the importance of moving from games to a distance measure and of identifying the topo-vectors for actually being able to implement the spatial framework. In particular, in IRIS once an image is placed in the database the topo-vector for its related topological model is computed, thus off-line, and it is the only data structure actually used in the retrieval process. The representation is quite compact both if compared with the topological model and with the image itself. In addition, the availability of topo-vectors as indexing structures enables us to use a number of information retrieval optimizations, [Frakes and Baeza-Yates, 1992].

6.5.1 Implementing the similarity measure

In IRIS, the similarity measure is built on three components:

$$\text{similarity}(I_q, I_j) = \frac{1}{k_n} (k_u^{\text{topo}} \cdot d_{\text{topo}}(I_q, I_j) + k_u^{\text{b}} \cdot d_{\text{b}}(I_q, I_j) + k_u^{\text{text}} \cdot d_{\text{text}}(I_q, I_j))$$

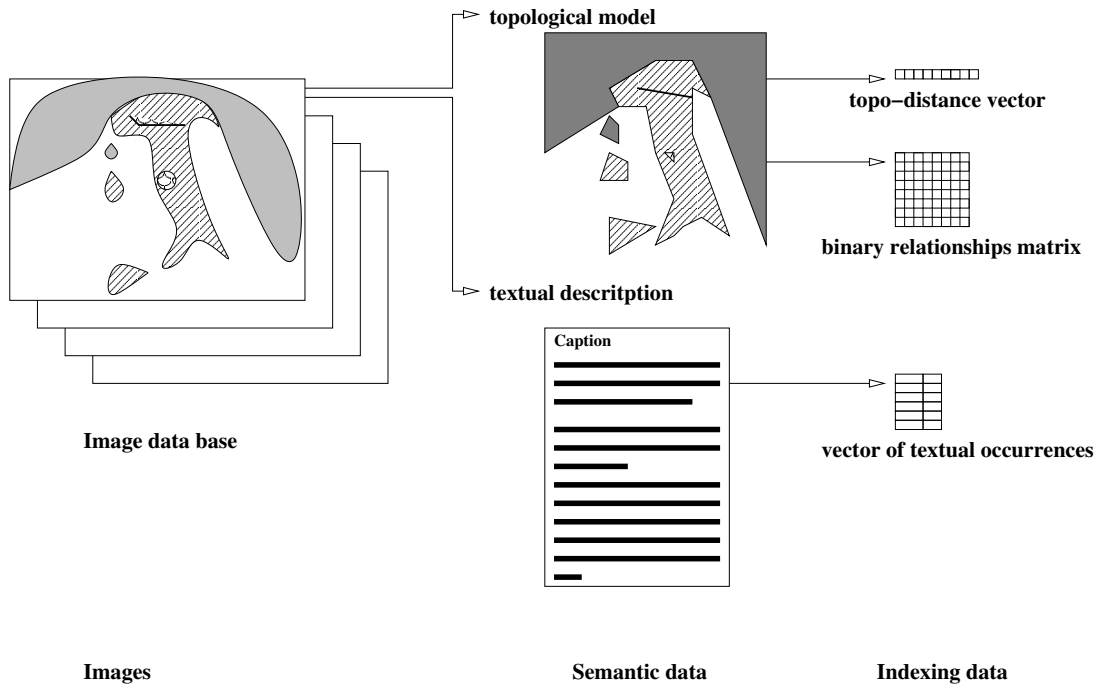


Figure 6.7: The organization of IRIS together with the indexing data structures.

where I_q is the query image (equipped with its topological model and textual description), I_j is the j -th image in the visual database, k_u^{topo} , k_u^{b} , and k_u^{text} are user defined factors to specify the relative importance of topological intricacy, binary relationships and text in the querying process, k_n is a normalizing factor, $d_{\text{topo}}(I_q, I_j)$ is the topo-distance between I_q and I_j , $d_{\text{b}}(I_q, I_j)$ and $d_{\text{text}}(I_q, I_j)$ are the distances for the binary spatial relationships and for the textual descriptions, respectively. In the context of IRIS, the textual component is considered independent of the two spatial ones, while the binary relationship and the topo-distance are also independent. In fact, the comparison of the topological configuration of a given region does not affect the comparison of its relations with other regions (it does not matter if a region is open regular or a spike when considering if it is contained in another region or not). The user defined factors k serve for experimentation purpose. So one can experiment with the relevance of a factor in the retrieval process. Ideally, one should find the perfect balance between the three components of the similarity measure and then fix these three parameters once and for all (or fix them for a specific domain).

The entire Section 6.4 is concerned with the computation of $d_{\text{topo}}(I_q, I_j)$. The topo-distance component is simply:

$$d_{\text{topo}}(I_q, I_j) = \text{topo-distance}(\text{t-vec}(I_q), \text{t-vec}(I_j))$$

The second component $d_{\text{b}}(I_q, I_j)$ of the similarity measure accounts for the binary spatial relationships between objects. When an image is indexed, a matrix is built.



Figure 6.8: The result of querying a database of men and cars.

This is a square matrix whose indices range over the regions present in the model. The generic entry $e_{i,j}$ of the matrix represents the spatial relationship between region i and region j and can be one of the following: disconnected, externally connected, overlap, equal, tangential part, non-tangential part, and the inverses of the last two (RCC8). Following [Egenhofer, 1997], we define a topological distance using RCC8 in the following way. Any two relations are at distance n if there is a path of length n in the graph in Figure 6.9 connecting the two nodes representing the relations. Our distance is slightly different from that in [Egenhofer, 1997] since we use a modification of its original graph, though the underlying idea is the same. In the similarity measure, one compares matrices $b(M_1, M_2)$:

$$d_b(I_q, I_j) = b(b_matrix(I_q), b_matrix(I_j))$$

where $b_matrix(I_j)$ is the matrix of binary s8 relations associated with the regions identified in the j -th image.

The third and last component $d_{\text{text}}(I_q, I_j)$ of the similarity measure deals with textual annotation. The motivation comes from captions accompanying images in paper documents or present ‘near’ images in hyper-media documents. We employ quite standard textual information retrieval techniques, see for instance [Frakes and Baeza-Yates, 1992], and therefore omit further explanation of this part of the similarity measure be-

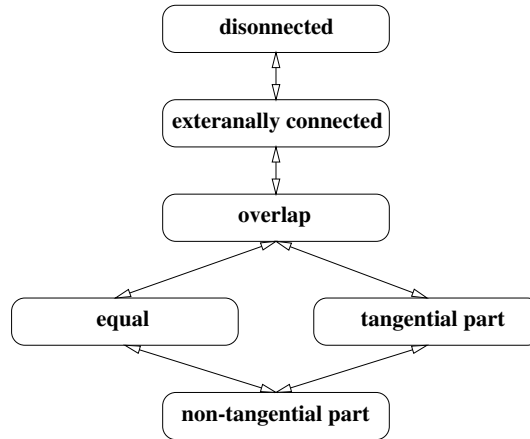


Figure 6.9: The binary relationships graph.

half for the standard definition of ‘textual distance’ between two image descriptions:

$$d_{\text{text}}(I_q, I_j) = \left(1 - \frac{\text{weighted_occurrences}(\text{text_vector}(I_q), \text{text_vector}(I_j))}{\text{length}(\text{text_vector}(I_q))}\right)$$

where $\text{text_vector}(I_j)$ is the list of meaningful words found in the description of the j -th image, $\text{weighted_occurrences}$ counts the number of instances of a word appearing in two textual vectors weighted by a factor indicating the indexing power of the word. A word is more powerful if it discriminates more, which in turns means that it occurs in less descriptions in the whole collection of image captions. The $d_{\text{text}}(I_q, I_j)$ follows a common way of defining a cosine distance among word vectors, see for instance [Witten et al., 1999].

6.6 Discussion

There are two abstractions on the idea of topo-distance that are worth noticing:

1. the transformation of model comparison games into distance measures for languages different from $\mathbf{S4}_u$,
2. the extension of the framework topo-bisimulation, topo-game, topo-distance to modal spatial languages more expressive than the simple $\mathbf{S4}_u$.

1. The theoretical framework proposed is much more general than what we have shown here. We were interested in a mereotopological framework and have therefore used the language $\mathbf{S4}_u$ interpreted on topological models, but an isosceles distance can be used for any modal language equipped with negation, for which one has adequate notions of model comparison games and bisimulation. Even the restriction to modal logic

is not necessary, one can think of first-order logic, of the usual Ehrenfeucht-Fraïssé games, of elementary equivalence in place of bisimulation, and an isosceles distance is then definable. The decidability result for the distance is the only thing that does not necessarily extend, rather one has to consider the class of models and the logic case by case. Of particular importance is then how the adequate topological games are defined. The technique employed in Theorem 6.3.2 for the language $\mathbf{S4}_u$ is, as we have just mentioned, much more general. A question interesting *per se*, but out of the scope of the present dissertation, is: which is the class of games (over which languages) for which a notion of isosceles distance holds? We believe the class of such languages and model comparison games to be quite vast.

2. The second abstraction step is, in some sense, an instance of the previous one. The idea is to take the framework topo-bisimulation, topo-game, topo-distance, algorithm to compute the distance to more expressive languages than $\mathbf{S4}_u$. The starting point is then to identify an appropriate language of, say, qualitative shape with adequate model comparison games. Then a newer version of IRIS can be built. We have seen a number of modal languages more expressive than $\mathbf{S4}_u$ together with adequate notions of model comparison games in Chapter 4 and 5. All these languages are excellent candidates for extending IRIS. The difficulty of the extension will mostly lie in the identification of efficient algorithms to compute the various distance measures.

A separate remark regarding experimentation is in order. Having implemented a system based on the topo approach is also an important step in the presented research. Experimentation is essential to assess applicability, but some preliminary considerations are possible. We have noticed that the prototype is very sensible to the labeling of segmented areas of images, i.e., to the assignment of proposition letters to regions. We have also noticed that the mereotopological expressive power appears to enhance the quality of retrieval and indexing over pure textual searches, but the expressive power of $\mathbf{S4}_u$ is still too limited. Notions of qualitative orientation, shape or geometry appear to be important, especially when the user expresses his desires in the form of an image query or of a sketch.

All in all, the idea of designing a spatial similarity measure based on formal model comparison games is both intriguing and rewarding from the intellectual point of view. Though, the gap between our implemented system and actually practical systems is still to be filled. There is no indication that the topo-distance gives human-intuitive meanings to the similarity of images, because a numerical distance can hide very different types of visual distinction.

Another major concern is the following. The system proposed may result to be very brittle when experimenting on real world images segmented automatically. The misclassification of a region, or the misinterpretation of a boundary, not to mention noise in the original image, can have a devastating impact on the values of the similarity measure. Solutions to fill these gaps between our system and more effective ones are bound to involve some genuine extensions of our pure topological framework.

CHAPTER 7

THICK 2D RELATIONS FOR DOCUMENT UNDERSTANDING

7.1 Introduction

When Dave placed his own drawing in front of the ‘eye’ of HAL—in 2001: A Space Odyssey—HAL showed to have correctly comprehended and interpreted the sketch. “That’s Dr. Hunter, isn’t it?” [Rosenfeld, 1997]. But what would have happened if Dave used the first page of a newspaper in front of the eye and started discussing its contents? Considering HAL a system capable of AI, we expect HAL to recognize the document as a newspaper, to understand how to extract information and to understand its contents. Finally, we expect Dave and HAL to begin a conversation on the contents of the document. In short, HAL has to be able to perform *document image analysis*.

Document image analysis is the set of techniques to recover syntactic and semantic information from images of documents, prominently scanned versions of paper documents. An excellent survey of document image analysis is provided in [Nagy, 2000] where, by going through 99 articles having appeared in the IEEE’s Transactions on Pattern Analysis and Machine Intelligence (PAMI), Nagy reconstructs the history and state of the art of document image analysis. Research in document images analysis is useful and studied in connection with document reproduction, digital libraries, information retrieval, office automation, and text-to-speech.

One may have different goals when performing document image analysis. For instance, one may be interested in the reconstruction of the reading order of a document from its image. One way to achieve this is by performing the following intermediate steps. First, one identifies the basic components of the document, the so-called *document objects*. Second, one identifies the logical function of the document objects within the document (e.g., title, page number, caption). This is called *logical labeling*. Last, one infers the order in which the user is to read the document objects. This phase is called the *reading order detection*. In the process, one moves from basic geometric information of the document composition, the *layout structure*, to semantic

information, the *logical structure*. document objects and their spatial arrangement are prototypical examples of elements of the layout structure, while the reading order is an instance of the logical structure.

In Figure 7.1, we illustrate possible flows of information in document image analysis. The first row represents the flow from the document image to its reading order. The following row represents the flow from the image to the identification of the document class. Discovering to which scientific publication belongs a given document image is an example of document classification. One should interpret the arrows in the figure as possible choices. It is perfectly normal to move from one row to another, or to stop the analysis at the layout structure level. For example, systems for mail delivery do not need to perform any document classification, or reading order detection.

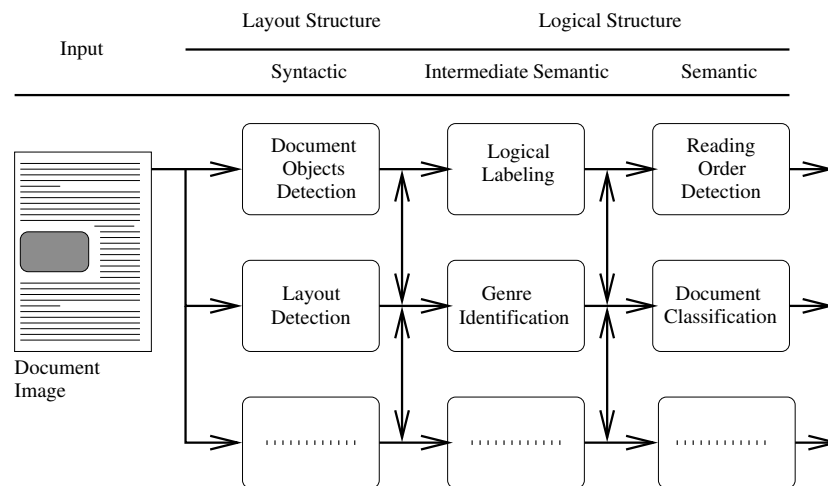


Figure 7.1: Various tasks in document image analysis and understanding. Left to right, from input data towards semantic content.

The first document image analysis systems were built to process documents of a specific class, e.g., forms for telegraph input. One of the recent trends is to build systems as flexible as possible, capable of treating the widest variety of documents. This has led to categorize the knowledge used in a document image analysis system into: class specific and general knowledge (e.g., [Cesarini et al., 1999]). In addition, such knowledge can be explicitly available or implicitly hard-coded in the system.

Lee and Choy [2000] present a system to analyze technical journals of one kind (PAMI) based on explicit knowledge of the specific journal. The goal is that of region segmentation and identification (logical labeling). The knowledge is formalized in “IF-THEN” rules applied directly to part of the document image and “IF-THEN” meta rules. Though the idea of encoding the class specific knowledge of a document is promising, it is not clear whether the proposed approach is scalable and flexible. Given the specific form of the IF-THEN rules, the impression is that the system is not suited for the analysis of documents different from PAMI. Experimental results show good

performance in the task of logical labeling, especially in the detection of formulas and drawings embedded in the main text.

There are a number of problems related to the rule based approaches found in the literature. The most prominent is the high specificity of the rules. The specificity makes it hard or impossible to extend such systems to documents of a class different from the one for which the system was originally designed. Another problem is the lack of proof of correctness or termination. Recent rule-based approaches for layout and logical structure detection are presented in [Klink and Kieninger, 2001, Lee et al., 2000, Niyogi and Srihari, 1996] while an older one is [Tsujimoto and Asada, 1992].

Given the difficulty in designing appropriate rules for the analysis of documents, approaches based on learning are interesting. The document classification components of the WISDOM++ system [Altamura et al., 2001] are based on first-order learning algorithms [Esposito et al., 2000]. Another advantage of such systems is their flexibility compared to the non-learning based systems. By training the system on a different class of documents with similar layout, it should be possible to reuse the same architecture. On the negative side, the rules learned are not intuitive. More often than not, these rules are impossible to modularize for further use on different document classes.

An important aspect of a document image analysis system working at the logical structure level is the representation of the information extracted from the document. The key here is a modularity and standardization of the representation. Markup languages are a good example of representation means with such qualities. The system presented in [Worring and Smeulders, 1999] uses HTML as its final output form, while [Altamura et al., 2001] uses XML. More abstract representations are labeled and weighted graphs. These have been used in various systems such as, for instance, the ones presented in [Li and Ng, 1999, Cesarini et al., 1998, Walischewski, 1997].

As we are investigating practical applications of spatial reasoning formalisms, it is relevant to review approaches using these kind of formalisms. In particular, we consider bidimensional extensions of Allen's interval relations, that is, rectangular relations. To the best of our knowledge, bidimensional Allen relations have been used in document image analysis in three cases [Klink et al., 2000, Singh et al., 1999, Walischewski, 1997]. In all these approaches, bidimensional Allen relations are used as geometric features descriptors, at times as labels for graphs and at other times as layout relations among document objects. Thus, the use of Allen relations is relegated to feature comparison and it is not used for performing any other kind of reasoning.

We present a methodology based on inference with bidimensional qualitative spatial relations for logical structure detection of document images. In particular, the methodology addresses the issue of detecting the reading order in documents from an heterogeneous collection without using any document specific knowledge.

The methodology is implemented in a prototype system named *SpaRe* (Spatial Reasoning component) part of a larger architecture for logical structure detection in a broad class of documents. In the next section, we give an overview of the architecture. In Section 7.3, we describe the methodology based on the concept of document encoding rule and of thick boundary interpretation of bidimensional Allen relations.

Section 7.4 is dedicated to the experimental results and their discussion. Directions for future work and a discussion of the methodology are presented in Section 7.5.

7.2 A logical structure detection architecture

In [Todoran et al., 2001a], we presented a logical structure detection architecture. Departing from a pre-processed document image the goal of such an architecture is that of logically labeling the document objects and subsequently identify the reading order. The system uses general document knowledge only, hence, it is applicable to documents of different classes.

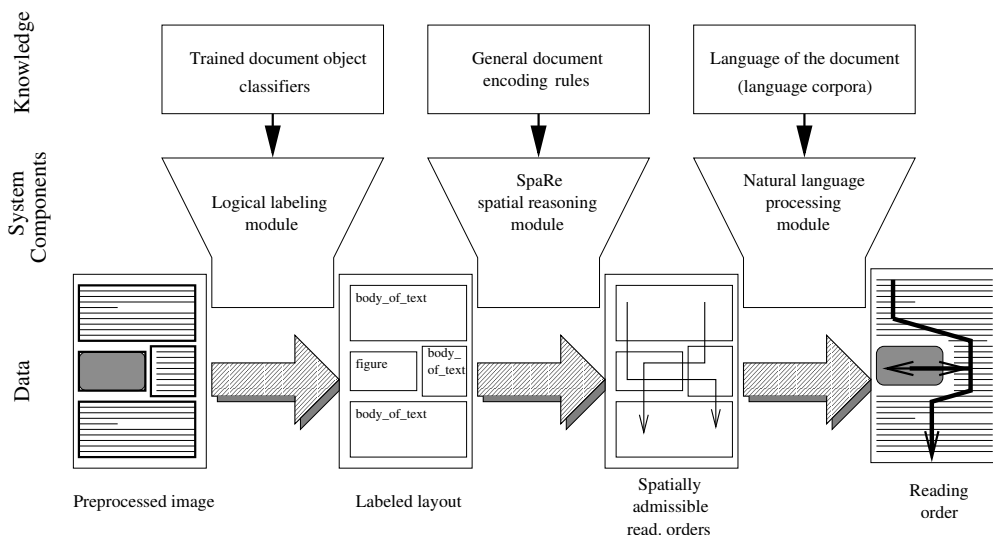


Figure 7.2: The flow of knowledge and data in the logical structure detection architecture presented in [Todoran et al., 2001a].

Referring to Figure 7.2, one has a glimpse of the architecture presented in [Todoran et al., 2001a]. The input is a pre-processed document image in which the document objects have been segmented, local textual content recognized and font information identified. The original document can be of any class as long as it is acceptable that document objects are represented by rectangles. Overlapping document objects are accepted by the system.

There are three modules: a logical labeler, a spatial reasoning reading order detector, and a natural language processing ‘disambiguator’. logical labeling on the pre-processed image is achieved via pre-trained classifiers.

The spatial reasoning module starts from the logically labeled layout of the document and, using general document encoding rules, it outputs a number of reading orders. The module is the subject of the remainder of the chapter.

The natural language processing module starts from the spatially admissible reading orders and the textual content of each one of the textual document objects. It uses this information to prune the set of spatially admissible reading orders of those which are linguistically not acceptable. This is performed by applying a combination of statistical methods and shallow parsing techniques. The statistical tools are trained on a large corpora of text. The training corpora is based on [Hersh et al., 1994] and [Baayen et al., 1995] which are independent from the document classes analyzed. Details of this module are presented in [Todoran et al., 2001b].

The output of the system is a reading order for the input document image. To be more precise, the output is a list of reading orders for the document ranked in order of linguistic plausibility (a probability is assigned to each reading order). Experimental results on each module and on the whole system have been presented in [Aiello et al., 2000, Todoran et al., 2001b, Todoran et al., 2001a].

7.3 Methodology

We focus on the spatial reasoning module of the architecture presented in the previous section. Figure 7.3 is a zoom-in of the spatial reasoning component in Figure 7.2 highlighting details. First, the generic document knowledge in the form of document encoding rules may have different origins. Second, the spatial reasoning module *SpaRe*, is actually composed of two sub-modules. The first one, which performs inference on the spatial relations of the layout and on the document encoding rules, is based on constraint satisfaction techniques. The second one is a module to sort graphs, that is, directed transitive cyclic ones. In the following sections, we analyze each of these items.

7.3.1 Document encoding rules

A *document encoding rule* is a principle followed by the author of a document to convey an intent of the author by layout details. document encoding rules can be one of two types: general or class specific. Document encoding rules can be expressed in a informal or in a formal manner. Informal rules are proposed in natural language or by sketch. Examples are found in books such as [Reynold, 1979]. Examples of generic and specific, and formal and informal rules are presented in Figure 7.4.

Let us consider a number of formal ways to express document encoding rules.

L^AT_EX is a compiled markup language. Typically, there is a number of source files with the main marked-up text (the `.tex` files), a number of style definition files (`.sty`, `.cls`) and a compiler. The document encoding rules can reside as macros in the `.tex` file, but the most common solution is that document encoding rules reside inside the style files. Consider the figure environment in the class file for generating transactions for the ACM.¹

¹M. Aiello. (2001). http://www.acm.org/pubs/submissions/latex_style/

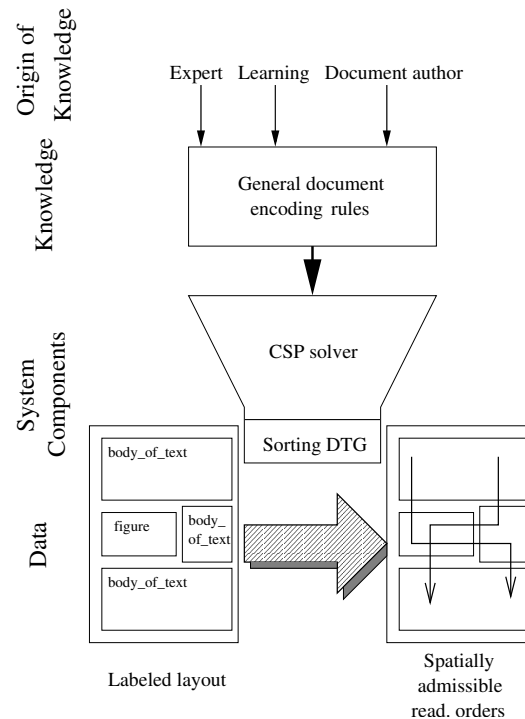


Figure 7.3: The flow of knowledge and data in the spatial reasoning module *SpaRe*. The document encoding rules originate from an expert, or from previous learning or are given directly by the document author. The module itself is composed of a constraint satisfaction problem solver and a handler for directed transitive cyclic graphs.

```

\newcounter{figure}
\def\thefigure{\@arabic\c@figure}
\def\fps@figure{tbp}
\def\ftype@figure{1}
\def\ext@figure{lof}
\def\fnum@figure{Fig.\ \thefigure}
\def\figure{\let\normalize\footnotesize \normalsize
  \@float{figure}}\let\endfigure\end@float
\@namedef{figure*}{\@dblfloat{figure}}
\@namedef{endfigure*}{\end@dblfloat}

```

The above definition, among other things,² defines the figure as belonging to a float environment [Goossens et al., 1994] whose default major features are: a float occupies the top of a page; a float does not have to appear where it is

acmtrans2m.cls. The class file currently in use at ACM, an extension of the acmtrans2e.cls version.

²See [Knuth, 1984] for details over the syntax and semantics of \TeX .

	general	class specific
informal	“the caption neighbors its figure”	“the caption starts with the word “Fig.” with font size 12pt, green text, and it is centered”
formal	$\forall f \in \text{figure} \exists c \in \text{caption}: \text{neighbors}(f, c)$	$\forall c \in \text{caption} \rightarrow \text{text_starts}(c, \text{“Fig.”}) \wedge \text{point}(c, 12) \wedge \text{font_color}(c, \text{‘Green’}) \wedge \text{centered}(c)$

Figure 7.4: Examples of generic and specific, and formal and informal rules. The formal rules are expressed in a first-order like language for documents whose semantics should be self-evident.

declared in the source; a float should not occupy more than 70% of the page otherwise it is moved after the first `\clearpage` instruction; if a caption is present it cannot be split across pages. The ACM transactions style file provides further class specific definitions for displaying the caption which overwrite the corresponding \LaTeX definitions.

```

\long\def\@makecaption#1#2{\vskip 1pc
  \setbox\@tempboxa\hbox{#1.\hskip 1em\relax #2}
  \ifdim \wd\@tempboxa >\hsize #1. #2\par \else
  \hbox to\hsize{\hfil\box\@tempboxa\hfil}
  \fi}
\def\nocaption{\refstepcounter\@capttype \par
  \vskip 1pc \hbox to\hsize{\hfil \footnotesize
  Figure \thefigure\hfil}}

```

The second example of a document rule in \LaTeX places the word “Figure” the figure counter immediately below the picture, placing a vertical space of 1pc units (i.e., 12pt). The size of such text is set to the value of `\footnotesize`.

More abstractly, the document encoding rule for a figure says that a figure is left to float in the main text, its preferred position is on top of a page and the caption is placed immediately below. The figure and caption always appear in the above/below spatial relation on one page.

WYSIWYG are computer systems in which the input of the user corresponds almost exactly to the final layout of the document (WYSIWYG stands for ‘what you

see is what you get'). A prototypical example is Microsoft's Word. In these sort of systems, it is hard to distinguish between the syntactic and the semantic portion of document encoding rules, as they are hidden in the implementation. The only control that the user has over the formal document encoding rules is through the functionalities provided by an interface allowing the user to change rule parameters.

In a common way of employing the WYSIWYG-style there are few strict document encoding rules, while the user still enforces elements of style. In a typical way of doing, captions may be put underneath a figure and also typically be indented (apart from the one place where the user forgot to implement that).

However, when observing the text one could learn document encoding rules, for example, that captions are always below the figure and immediately following it provided there is one. In that case, one would require rules which can express topological relationships with some form of tolerance as the user will implement notions like alignment and marking with a limited precision. In addition, one would require rules which express topographic relationships as they can be implemented in the freedom to move around on the 2D-screen where the WYSIWYG-program runs, implying that the caption is always close to the figure. Finally, to address the inconsistencies of ad hoc rule implementation and the lack of discipline to enforce them would require rules with a less than strict character.

SGML languages are a family of interpreted markup languages, whose best known members are HTML and XML. The eXtensible Markup Language, XML for short, achieves a clear separation between content (the `.\indexn{XML}` file), syntactic document encoding rules (`.css`, `.xsl`, `.dtd`) and semantics of the document encoding rules (the browser's interpretation of the document encoding rules). For instance, the document encoding rule for a caption like `<CAPTION>` A figure `</CAPTION>` could be the following:

- (*syntax*): inside a `.css` file
`CAPTION`
`{display: block; font-size: 12pt; color: #000000; text-align: center}`
- (*semantics*): the browser will display the text "A figure" in one block of text, in black color, using the default font, using the font size 12pt, and center it.

To the same degree SGML as WYSIWYG offers the possibility to move around the images of the document objects and hence implement document encoding rules by habit rather than by a priori rules. As the user has no visual feedback, the factual encoding rules are more informal than in the WYSIWYG paradigm. Hence, here are needed topological and topographical rule sets to describe the power of SGML but even more forgiving than in the WYSIWYG style.

Abstract formal languages can also serve as document encoding languages, for instance, first-order logic. The syntax and semantics are the usual ones for first-order logic, taking special care in giving adequate semantics to spatial relations and predicates.

A final example of a general document encoding rule stated informally in natural language is the following:

“in the Western culture, documents are usually read top-bottom and left-right.” (7.1)

A problem of stating rules in natural language is ambiguity. In fact, we do not know if one should interpret the “and” as commutative or not. Should one first go top-bottom and *then* left-right? Or, should one apply any of the two interchangeably? It is not possible to say from the rule merely stated in natural language.

In the next section, we define an abstract propositional formal language to express qualitative spatial relations among document objects to formally express document encoding rules.

7.3.2 Relations adequate for documents

Considering relations adequate for documents and their components, requires a preliminary formalization step. This consists of regarding a document as a formal model. At this level of abstraction a document is a tuple $\langle D, R, l \rangle$ of document objects D , a binary relation R , and a labeling function l . Each document object $d \in D$ consists of the coordinates of its bounding box (defined as the smallest rectangle containing all elements of that object)

$$D = \{d \mid d = \langle id, x_1, y_1, x_2, y_2 \rangle\}$$

where id is an identifier of the document object and (x_1, y_1) (x_2, y_2) represent the upper-left corner and the lower-right corner of the bounding box of the document object. In addition, we consider the logical labeling information. Given a set of labels L , logical labeling is a function l , typically injective, from document objects to labels:

$$l : D \rightarrow L$$

In the following, we consider an instance of such a model where the set of relations R is the set of bidimensional Allen relations and where the set of labels L is $\{\text{title, body_of_text, figure, caption, footer, header, page_number, graphics}\}$. We shall refer to this model as a *spatial [bidimensional Allen] model*. Bidimensional Allen relations consist of 13×13 relations: the product of Allen’s 13 interval relations [Allen, 1983, van Benthem, 1983b] on two orthogonal axes. (Consider an inverted coordinate system for each document with origin (0,0) in the left-upper corner. The x axis spans horizontally increasing to the right, while the y axis spans vertically towards the bottom.) Each relation $r \in A$ is a tuple of Allen interval relations of the

form: *precedes*, *meets*, *overlaps*, *starts*, *during*, *finishes*, *equals*, and *precedes_i*, *meets_i*, *overlaps_i*, *starts_i*, *during_i*, *finishes_i*. We shall refer to the set of Allen bidimensional relations simply as A and to the propositional language over bidimensional Allen relations as \mathcal{L} the remainder of the chapter. Since Allen relations are jointly exhaustive and pairwise disjoint, so is A . This implies that given any two document objects there is one and only one A relation holding among them.

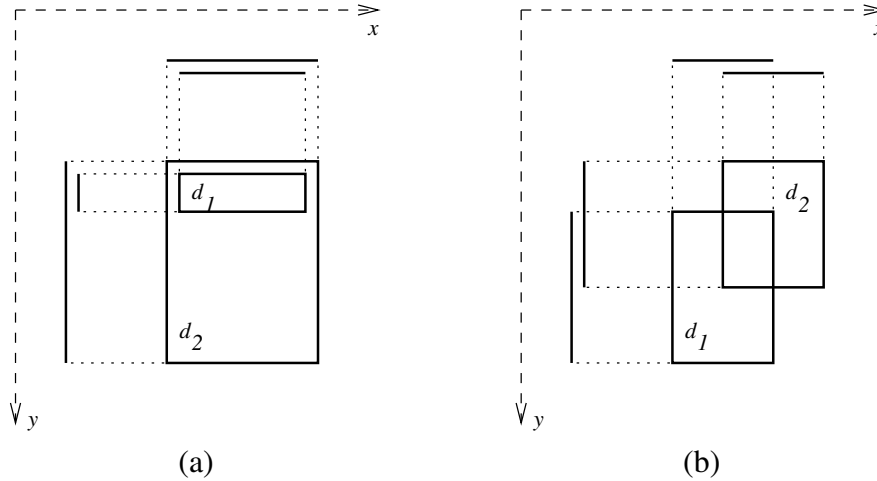


Figure 7.5: (a) The document object d_1 is Part of d_2 , as the projection of d_1 on both axes is *during* the projection of d_2 ; (b) The document object d_2 Overlaps with d_1 , as the projection on x of d_1 *overlaps* that of d_2 and on y it *overlaps_i* that of d_2 .

Document objects are represented by their bounding boxes and the relative position of these objects plays a key role in the interpretation of the meaning of the document. For example, if a document object is above another one it is likely that it should be read before. If a document object is contained in another one, it is likely that the contained one is a ‘part’ of the containing one, for instance the title of a remark inside a frame. document objects can be also overlapping. This last feature is more common when the document has different colors and colored text runs over pictures, logos and drawings.

All relations of the examples above are expressible in terms of \mathcal{L} . For instance, ‘being part of’ is

$$\text{Part}(d_1, d_2) \text{ iff } (during_x(d_1, d_2) \vee starts_x(d_1, d_2) \vee finishes_x(d_1, d_2)) \wedge (during_y(d_1, d_2) \vee starts_y(d_1, d_2) \vee finishes_y(d_1, d_2)) \quad (7.2)$$

To analyze the expressive power of \mathcal{L} , we encode the basic RCC5 [Randell et al., 1992] relations:

- $\text{Part}^{-1}(d_1, d_2) = \text{Part}(d_2, d_1)$,
- $\text{Equal}(d_1, d_2) = equal_x(d_1, d_2) \wedge equal_y(d_1, d_2)$,

- $\text{Disconnected}(d_1, d_2) = \text{precedes}_x(d_1, d_2) \vee \text{precedes}_{i_x}(d_1, d_2) \vee \text{precedes}_y(d_1, d_2) \vee \text{precedes}_{i_y}(d_1, d_2)$,
- $\text{Overlap}(d_1, d_2) = \neg \text{Part}(d_1, d_2) \wedge \neg \text{Part}^{-1}(d_1, d_2) \wedge \neg \text{Equal}(d_1, d_2) \wedge \neg \text{Disconnected}(d_1, d_2) \wedge \neg \text{ExternalConnection}(d_1, d_2)$.

Similarly, one can encode RCC8 in \mathcal{L} . Examples of document objects satisfying the part and the overlap relations are presented in Figure 7.5.

Restricting attention to RCC relations one loses a feature of \mathcal{L} of great importance, namely, its ordering expressivity with respect to the axes. Take for instance the `Disconnected` relation. There are various ways in which two document objects can satisfy this relation. If either $\text{precedes}_x(d_1, d_2) \wedge \text{equal}_{y_x}(d_1, d_2)$ or $\text{precedes}_{i_x}(d_1, d_2) \wedge \text{equal}_{y_x}(d_1, d_2)$ holds, then it is true that the RCC8 predicate `Disconnected`(d_1, d_2) holds, but the two situations are most different. In the first case, d_1 is to the left of d_2 , in the second case it is to the right. In other words, in the first case it is likely that d_1 is to be read before than d_2 in the document, while in the second case d_2 is to be read before d_1 . This is one of the key features that we exploit in using \mathcal{L} to define document encoding rules.

Consider again the example of the relation between a figure and its caption in the \LaTeX ACM transactions class file. This spatial relation is \mathcal{L} definable:

$$(\text{during}_x(\text{figure}, \text{caption}) \vee \text{equals}_x(\text{figure}, \text{caption})) \wedge \text{precedes}_y(\text{figure}, \text{caption})$$

The spatial relation between the word “Figure” and the figure counter is also \mathcal{L} definable:

$$\text{meets}_x(\text{“Figure ”}, \text{figure_counter}) \wedge \text{equals}_y(\text{“Figure ”}, \text{figure_counter}) \vee \text{during}_{i_y}(\text{“Figure ”}, \text{figure_counter})$$

Other features of the \LaTeX definitions are not \mathcal{L} definable: trivially, all font and textual features. But also size and distance features are not \mathcal{L} definable, e.g., the fact that the white space between a figure and a caption is of a fixed amount (`1pc`).

7.3.2.1 Document encoding rules with \mathcal{L}

The language \mathcal{L} is adequate to express mereotopological and ordering relations among rectangles. Here, we show how to use this power to express formal unambiguous document encoding rules.

Take the informal document encoding rule (7.1) expressed in natural language. Consider the layout of a document as presented in Figure 7.6.a, where the numbering of the document objects is provided counterclockwise. After having read the document object 2, to which one should the reader move? Only having the layout and not the content of the text there is not a unique choice. One would either move to the block of

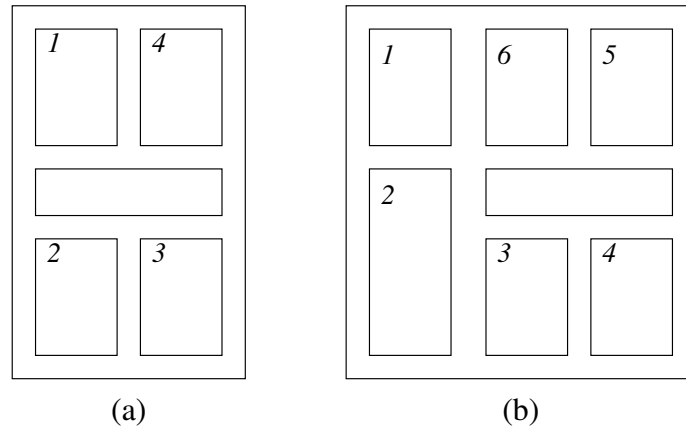


Figure 7.6: Layouts of documents considering text objects only.

text 4 or to block 3. In the first case, one has followed the left-right rule, in the latter the top-bottom rule. No one would have proposed to move to block 1, this because it is in violation of the top-bottom rule.

The top-bottom, left-right document rules are expressible in the language \mathcal{L} by:

$$\begin{aligned} \text{before_in_reading}(d_1, d_2) \text{ iff } & \text{precedes_x}(d_1, d_2) \vee \text{meets_x}(d_1, d_2) \vee \\ & \text{overlaps_x}(d_1, d_2) \vee \text{precedes_y}(d_1, d_2) \vee \\ & \text{meets_y}(d_1, d_2) \vee \text{overlaps_y}(d_1, d_2) \end{aligned} \quad (7.3)$$

The equation reads “the document object d_1 is ‘before in the reading order’ of the document object d_2 if the a Boolean combination of basic \mathcal{L} relations are satisfied.” The rule (7.3) is the formal counterpart to (7.1). Though the generality of (7.3) is also its weakness. Too many document objects satisfy it, calling for the design of rules balancing between being more restrictive and being general. Consider the layout proposed in Figure 7.6.b. It is hard to judge if one would follow the reading 1, 2, 6, 3, 5, 4 or the reading 1, 6, 5, 2, 3, 4, but the reading 1, 6, 2, 3, 5, 4 surely seems odd. Without knowing the content of the document, we are inclined to consistently apply a column-wise or row-wise rule. Therefore, a candidate for a general and yet more restrictive rule in comparison with (7.3) is a *column-wise* document rule. In this case, one *first* goes top-bottom, *then* left-right. A rule to encode this behavior is again expressible with \mathcal{L} . It has the following form:

$\text{before_in_reading}^{\text{col}}(d_1, d_2)$ iff

$$\begin{aligned}
 & \text{precedes}_x(d_1, d_2) \vee \text{meets}_x(d_1, d_2) \vee \\
 & (\text{overlaps}_x(d_1, d_2) \wedge \\
 & (\text{precedes}_y(d_1, d_2) \vee \text{meets}_y(d_1, d_2) \vee \text{overlaps}_y(d_1, d_2))) \vee \\
 & ((\text{precedes}_y(d_1, d_2) \vee \text{meets}_y(d_1, d_2) \vee \text{overlaps}_y(d_1, d_2)) \wedge \\
 & (\text{precedes}_x(d_1, d_2) \vee \text{meets}_x(d_1, d_2) \vee \text{overlaps}_x(d_1, d_2) \vee \\
 & \text{starts}_x(d_1, d_2) \vee \text{finishes}_i_x(d_1, d_2) \vee \text{equals}_x(d_1, d_2) \vee \\
 & \text{during}_x(d_1, d_2) \vee \text{during}_i_x(d_1, d_2) \vee \text{finishes}_x(d_1, d_2) \vee \\
 & \text{starts}_i_x(d_1, d_2) \vee \text{overlaps}_i_x(d_1, d_2)))
 \end{aligned} \tag{7.4}$$

The declarative code implementing this rules is presented on page 175. An analogous *row-wise* rule is obtained by inverting the axes in (7.4).³

7.3.2.2 Thick boundary interpretation

The direct application of systems based on Allen or similar relations results in brittle systems. This is because Allen relations rely on the precise identification of a boundary of the interval. This means that some relations never occur in practical situations. One goes directly from *precedes* to *overlaps* and from *overlaps* to *during* without ever identifying an instance of *meets*, *starts*, or *finishes*. To solve this drawback of Allen-like relations, we provide a less brittle interpretation of the relations.

Instead of considering two interval extremes to be equal when they have the same coordinates, we consider them equal if they are closer than a fixed threshold distance \mathbf{T} . This can be seen as if the bounding boxes of the document objects have a *thick boundary*. We name the set of thirteen Allen's relations thus interpreted *thick boundary rectangle relations*.

The thickness of the boundary is assumed identical for all objects in the document. It is fixed with respect to the page size. The optimal value is found through experimentation. There is a constraint on the \mathbf{T} with respect to the size of the smallest document object: it should not exceed half the size of the shortest side of all bounding boxes. Referring to Figure 7.7, one sees how the A relations with their thick interpretation are more tolerant in the establishment of a relation between two intervals. For example, interval a meets interval b not only if $x_2^a = x_1^b$, but also if $x_1^b - T \leq x_2^a \leq x_1^b + T$. With the thick boundary interpretation, Allen's relation maintain the jointly exhaustive and pairwise disjoint property, see [Todoran et al., 2001a] for a proof. The declarative code with the clauses defining A with the thick boundary interpretation are reported in Appendix C on pages 177–179.

³Its implementation is presented on page 175.

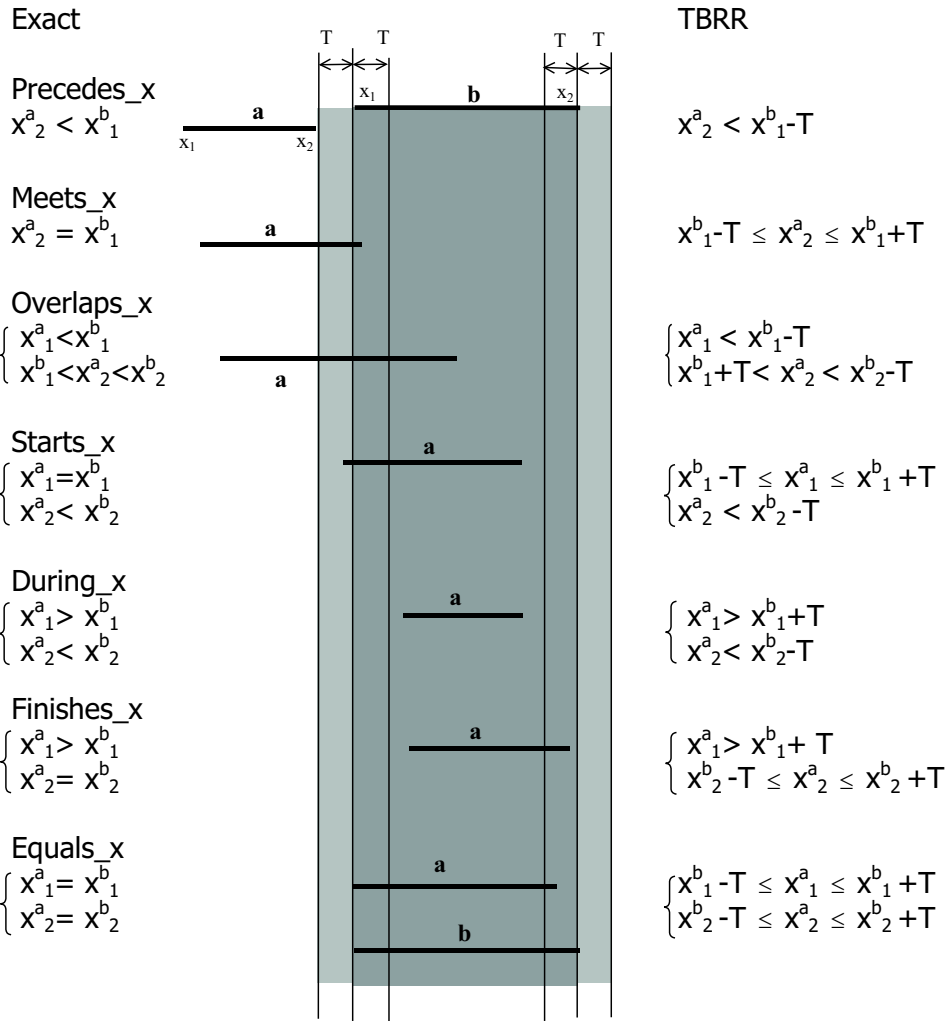


Figure 7.7: The thick boundary interpretation of Allen’s relations. The interval b is considered fixed and the threshold T is highlighted on its extreme points. The interval a varies in all 13 possible positions. On the left, the equation of the standard interpretation of Allen’s relations. On the right, the thick boundary interpretation.

7.3.2.3 Theoretical excursus

One might wonder about the connection between \mathcal{L} and the family of languages presented in the first half of the thesis. The connection is strong, as we have already remarked by showing the encoding of RCC8 in terms of \mathcal{L} . But there is more.

The A relations are mereotopological relations, but they are also weak geometrical relations. It is possible to define a notion of betweenness, see Section 5.1.2, in terms

of \mathcal{L} . Consider the following definition

$$\begin{aligned} \beta(d_1, d, d_2) \text{ iff } & \neg \text{precedes_i_x}(d_1, d) \wedge \neg \text{meets_i_x}(d_1, d) \wedge \\ & \neg \text{overlaps_i_x}(d_1, d) \wedge \neg \text{precedes_i_x}(d, d_2) \wedge \\ & \neg \text{meets_i_x}(d, d_2) \wedge \neg \text{overlaps_i_x}(d, d_2) \wedge \\ & \neg \text{precedes_i_y}(d_1, d) \wedge \neg \text{meets_i_y}(d_1, d) \wedge \\ & \neg \text{overlaps_i_y}(d_1, d) \wedge \neg \text{precedes_i_y}(d, d_2) \wedge \\ & \neg \text{meets_i_y}(d, d_2) \wedge \neg \text{overlaps_i_y}(d, d_2) \end{aligned}$$

We call this notion *Manhattan betweenness*, in the spirit of the Manhattan distance. An

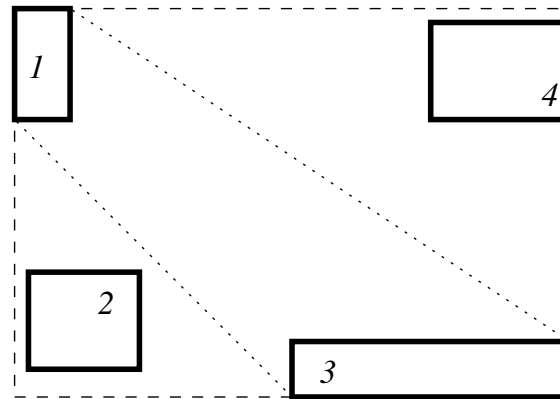


Figure 7.8: The document objects 2 and 4 lie ‘in between’ the document objects 1 and 3. 2 is strictly in between 1 and 3, while 4 is a limit case.

example of \mathcal{L} -betweenness holding among three rectangles is presented in Figure 7.8. One can check that it satisfies the universal betweenness axioms (Section 5.1) with one minor adjustment. The identity axiom becomes $\beta(d_1, d, d_1) \rightarrow \text{Part}(d, d_1)$, that is, the equality relation in **A2** is replaced by the ‘part’ relation.

To move from \mathcal{L} to a modal logic of rectangles is possible. The techniques used to perform the same move for the one-dimensional case are the most promising. The idea of chopping intervals [Venema, 1991] could be extended to chopping rectangles. Also the technique of Halpern and Shoham [Halpern and Shoham, 1991] should work for rectangles.

7.3.3 Inference

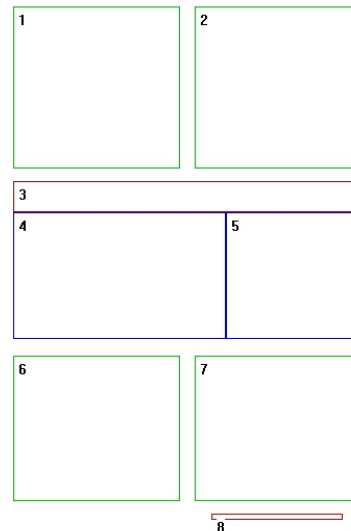
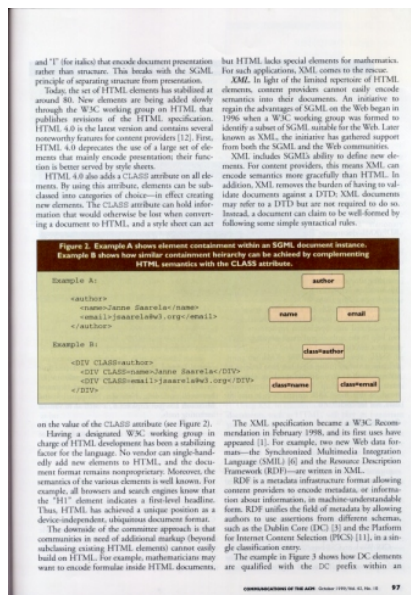
Equipped with a qualitative spatial language for document objects \mathcal{L} , with document encoding rules and the layout and logical labeling information, we are now in the position to perform inference in order to achieve ‘understanding’ of a document. Following is the definition of document understanding in this context.

First, we define the notion of an admissible transition between document objects. Given a pair of document objects d_1 and d_2 , a document model $\langle D, R, l \rangle$ and a set of document encoding rules S , we say that (d_1, d_2) is an *admissible transition* with respect to R iff the bidimensional Allen relation $(d_1, d_2) \in R$ is consistent with S .

A *spatially admissible reading order* with respect to a document model $\langle D, R, l \rangle$ and a set of document encoding rules S is a total ordering of document objects in D with respect to the admissible transitions.

The *understanding* of the document with respect to a document model $\langle D, R, l \rangle$ and a set of document encoding rules S is the set of spatially admissible reading orders.

Following the above definitions, we see that inference is performed by two following steps. The first one is a constraint satisfaction step in which instances of bidimensional Allen relations are matched against document encoding rules expressed in \mathcal{L} . The second one is a graph sorting procedure similar to topological sorting.



(a)

(b)

Figure 7.9: A page from the Communications of the Association for Computing Machinery and a possible layout segmentation of it.

Consider the image from the magazine Communications of the Association for Computing Machinery presented in Figure 7.9.a. A possible segmentation of its layout (Figure 7.9.b) is formally represented by

- [1, body_of_text, [13, 23, 93, 101], Times, 11, 0, 16]
- [2, body_of_text, [100, 23, 180, 101], Times, 11, 0, 16]
- [3, caption, [13, 107, 180, 122], Arial, 11, 0, 16]
- [4, graphics, [13, 122, 115, 183], Courier, 11, 16, 0]
- [5, figure, [115, 122, 180, 183], None, 11, 0, 16]

```
[6, body\_of\_text, [13, 191, 93, 261], Times, 11, 0, 16]
[7, body\_of\_text, [100, 191, 180, 261], Times, 11, 0, 16]
[8, footer, [108, 267, 171, 270], Arial, 7, 0, 16]
[9, page\_number, [175, 267, 180, 270], Arial, 12, 0, 16]
```

where each element of the list represents one document object together with its layout and logical labeling information. The first element is a unique identifier, the second is the logical label, the third is the upper-left corner and the bottom-right corner of the bounding box, the fourth is the font of the text (if applicable), then the size of the font, the color of the font, and the last element is the color of the background.

Consider using the general document encoding rule (7.3). For all pairs of document objects labeled by "body_of_text", we consider their bidimensional Allen relation. Then we input these together with (7.3) into a constraint satisfaction solver. Obtaining the following set of admissible transitions

```
[1, 2], [1, 6], [1, 7], [2, 6], [2, 7], [6, 2], [6, 7]
```

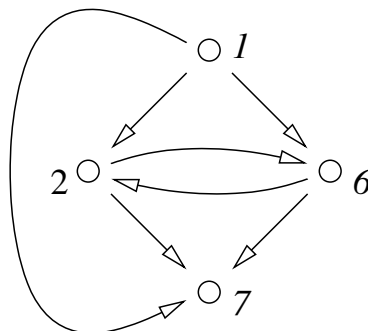


Figure 7.10: The graph of spatially admissible transitions for the body_of_text document objects of the document in Figure 7.9.

One can view this as a directed graph of spatially admissible transitions, Figure 7.10. There are two possible complete total orderings of this graph. They are

```
[1, 6, 2, 7]      [1, 2, 6, 7]
```

Following the above definition, the two spatially admissible reading orders constitute the ‘understanding’ of the document in Figure 7.9.b with respect to the set of document encoding rules $\{(7.3)\}$. Once the set of spatially admissible transitions is identified, the task is that of totally sorting the graph. The algorithm to perform the sorting of directed transitive cyclic graphs is presented in Appendix B.

7.4 Evaluation

The methodology proposed has been implemented in a prototype system: *SpaRe*. The core of *SpaRe* is implemented in the declarative programming language Eclipse,⁴ making use of the finite domain constraint satisfaction libraries. Relevant passages of Eclipse code are presented in Appendix C.3.

To test *SpaRe*, we used the Media Team Data-Base (MTDB) from the University of Oulu, [Sauvola and Kauniskangas, 2000]. The data set consists of scanned documents of various types: technical journals, newspapers, magazines, and one-page commercials. Elements from the data set are presented in Figure 7.11. We only used the documents in English, resulting in a data set of 34 documents having 171 pages. The MTDB data set has a ground truth at the document object level. Every document object has a layout label and a logical label. The reading orders are part of the ground truth. Of the 171 pages, 133 have a unique reading order, 32 have two independent reading orders, 5 have three, and 1 has four. We considered the layout information from the ground truth as the input to our system. As there is no ground truth for textual content and font information, we used the TextBridge OCR package⁵ to extract these.

For evaluation purposes, the documents in the data set were split into three main groups, based on their complexity:

- *trivial* documents containing up to 3 textual document objects;
- *regular* documents containing between 4 and 8 textual document objects;
- *complex* documents containing more than 8 textual document objects;

Out of 171 document pages, 98 are of type *trivial*, 66 of type *regular* and 7 are of type *complex*.

The goal of the experimentation was to evaluate whether *SpaRe* is effective in the detection of the reading order given the layout information. As subtasks, we were interested in evaluating the performance with different document encoding rules and with different values of the threshold for the thick boundary interpretation of bidimensional Allen relations.

The experiments consisted of three cases. In the first case, we have used the layout and labeling information from the ground truth and the general document encoding rule (7.3), denoted as *General Rule on Ground Truth data*. In the second case, we have used the layout and labeling information from the ground truth and the column and row-wise document encoding rules (7.4), denoted as *Column/Row Rules on Ground Truth data*. In the last case, we have used the layout and labeling information from an existing logical labeler (see Section 7.2) and the column and row-wise document encoding rules (7.4), denoted as *Column/Row Rules on the logical labeler data*. For each one of these we have varied the threshold of the thick boundary interpretation from 0 to 400 dots.

⁴<http://www-icparc.doc.ic.ac.uk/eclipse>.

⁵TextBridge SDK 4.5, ScanSoft, <http://www.scansoft.com>.

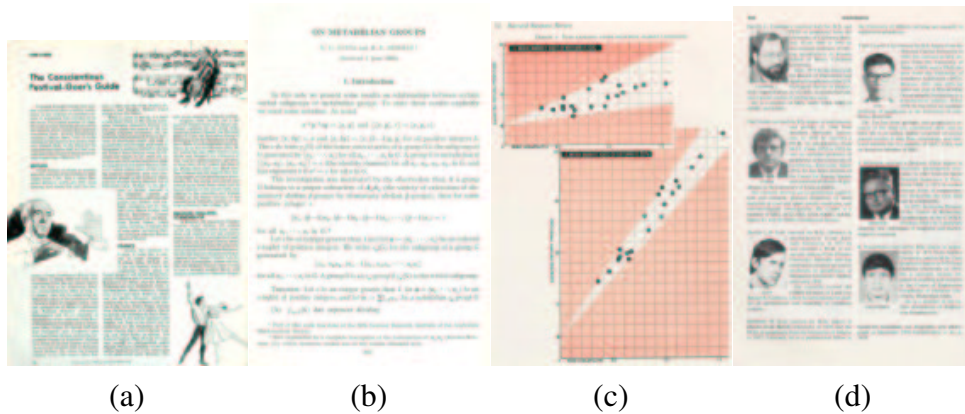


Figure 7.11: Sample images from the MTDB data set.

7.4.1 Criteria

To evaluate *SpaRe*, we use precision and recall [Baeza-Yates and Ribeiro-Neto, 1999]. The set of reading orders detected (D) is compared to the ground truth. For 38 documents, the ground truth defines independent reading orders on non-intersecting subsets of the textual objects within the same document. In these cases, the reading orders are composed by one main sequence of document objects and one or two blocks to be read independently; e.g., a page containing a frame with independent text. To account for this portion of documents with multiple reading orders (20% of the whole data set), we consider a reading order correct if it is identical to at least one permutation of the independent reading orders as defined in the ground truth.

We refer to the set of permutations of the ground truth as the set of correct reading orders (C). Then, the precision and recall are defined as follows:

$$p = \frac{|D \cap C|}{|D|} \quad r = \frac{|D \cap C|}{|C|} \quad (7.5)$$

The values lie between 0 and 1 inclusive, where 0 indicates the worst possible performance and 1 the best possible one. Because there is only one reading order, the recall can only be 1 if the correct reading is among the ones detected, or 0 if it is not. This makes the recall less informative of the overall behavior of the system.

7.4.2 Results

We have evaluated the results in terms of the average precision and recall defined in Equation 7.5.

General Rule on Ground Truth data. We have used the general document encoding rule (7.3) on the ground truth layout and logical labels of the MTDB documents.

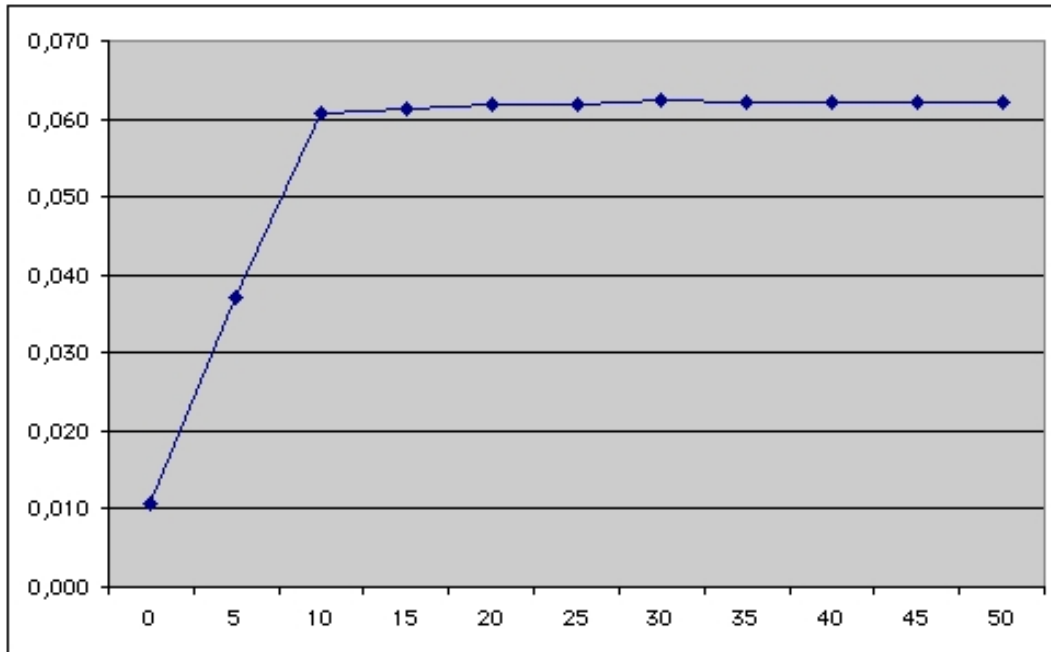


Figure 7.12: Average precision for increasing threshold values (between 0 and 50) using the general rule on the ground truth of the MTDB data set. The maximum value is for the threshold value of 30.

The values of average precision with respect to increasing values of the threshold are shown in Figure 7.12.

The average precision and recall of the system for the entire MTDB data set for the threshold value of 15 are:

<i>Document group</i>	<i>Number of Documents</i>	SpaRe	
		p	r
trivial	98	0.96	0.99
regular	66	0.31	0.97
complex	7	0.003	1.00
<i>average</i>	171	0.06	0.98

SpaRe detected 2714 reading orders for the 171 document pages in the data set. In the case of a very rich and complex document, 2157 reading orders were detected. For other four documents, 140, 50, 37 and 15 reading orders were detected. For the remaining collection the average of reading orders detected was of 1.74. In two cases, none of the reading orders as detected were correct.

Column/Row rule on Ground Truth data. We have used together the column and row-wise document encoding rules on the ground truth layout and logical labels

of the MTDB documents. The values of average precision with respect to increasing values of the threshold are shown in Figure 7.13. The maximum value of precision is for the threshold value of 15.

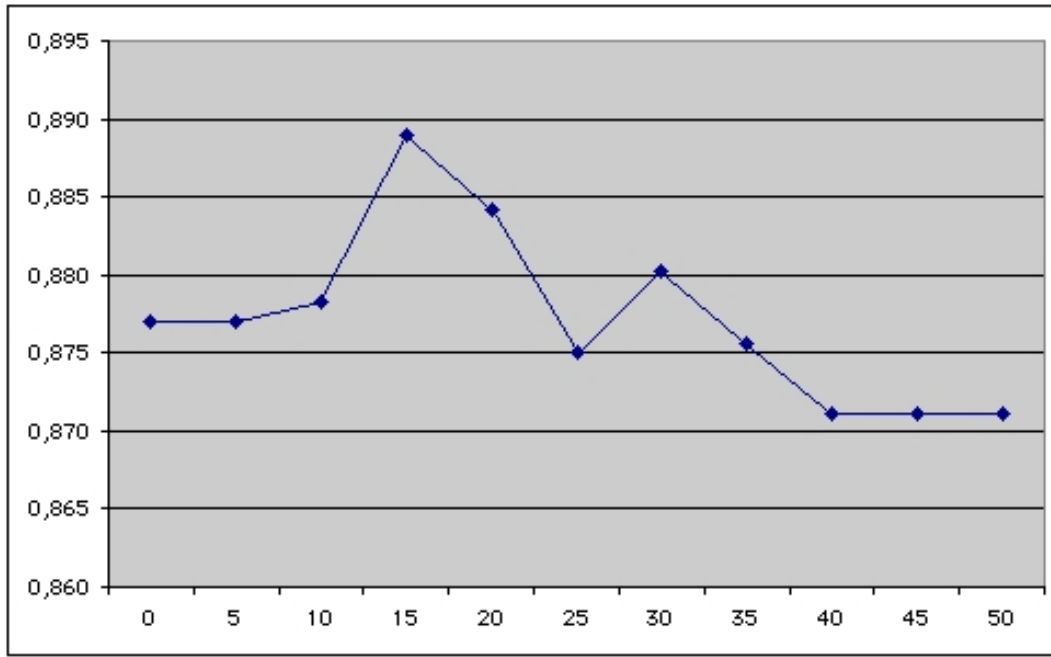


Figure 7.13: Average precision for increasing threshold values (between 0 and 50) using the column/row rule on the ground truth of the MTDB data set.

The average precision and recall of the system for the entire MTDB data set for the threshold value of 15 are:

<i>Document group</i>	<i>Number of Documents</i>	SpaRe	
		p	r
trivial	98	0.97	0.99
regular	66	0.79	0.97
complex	7	0.88	1.00
<i>average</i>	171	0.89	0.98

SpaRe detected 190 reading orders for the 171 document pages in the data set. For 16 documents 2 reading orders were detected, including the correct one. In one case, none of the two reading orders as detected were correct. For one document, 4 possible reading orders were detected and none of them was correct. For the rest of 154 documents, SpaRe detected one reading order only and in one case this was not correct.

In the case of a two column scientific article composed of 6 textual document objects, SpaRe detected 4 reading orders. These were all wrong because a short subtitle (“Acknowledgments”) was too close to a white space in the neighboring column and was considered the title of the neighboring row in a row-wise reading. This row-wise connection was possible in four different ways, all incorrect. In case of a first page of an article in a magazine composed of 3 textual document objects, the title was on the left of the main text and centered vertically. In a reading order, the title was considered by SpaRe to be a subtitle of one of the two main bodies of text. It was placed incorrectly in the center of the reading order instead of on top of it. For one document composed of 4 textual document objects organized in one column with two subtitles and poorly typeset, SpaRe wrongly detected the reading order. The reason is that the subtitles were almost embedded in the main text and in *overlap* relation in the x axes instead of *meet*. The problem disappears when increasing the threshold value above 25 points.

The column-wise document rule has as one of its conditions that two blocks meet on the x axis. But with the boundary’s thickness set to 0, this never occurs in the data set. On the other hand, allowing thickness, the *meet* relation holds among some neighboring document objects.

Column/Row on the logical labeler data. We have used the column and row-wise document encoding rules on the output of a logical labeling system on the MTDB documents. The values of average precision with respect to increasing values of the threshold are shown in Figure 7.14. The maximum value of precision is for the threshold value of 15.

The average precision and recall of the system for the entire MTDB data set for the threshold value of 15 are:

<i>Document group</i>	<i>Number of Documents</i>	SpaRe	
		p	r
trivial	98	0.92	0.94
regular	66	0.74	0.92
complex	7	0.86	1.00
<i>average</i>	171	0.84	0.94

SpaRe detected 192 reading orders for the 171 document pages in the data set. For 18 documents 2 reading orders were detected where the ground truth indicates only one. For one document, 4 possible reading orders were detected and none of them was correct. For the rest of 152 documents, SpaRe detected one reading order only. For 11 documents the correct reading order was not detected by SpaRe. In particular, for the simple documents 2 extra reading orders were detected and the number of wrongly understood documents was of 6. For the regular documents, the number of wrong detections was 5. For the 7 complex documents, there were no errors.

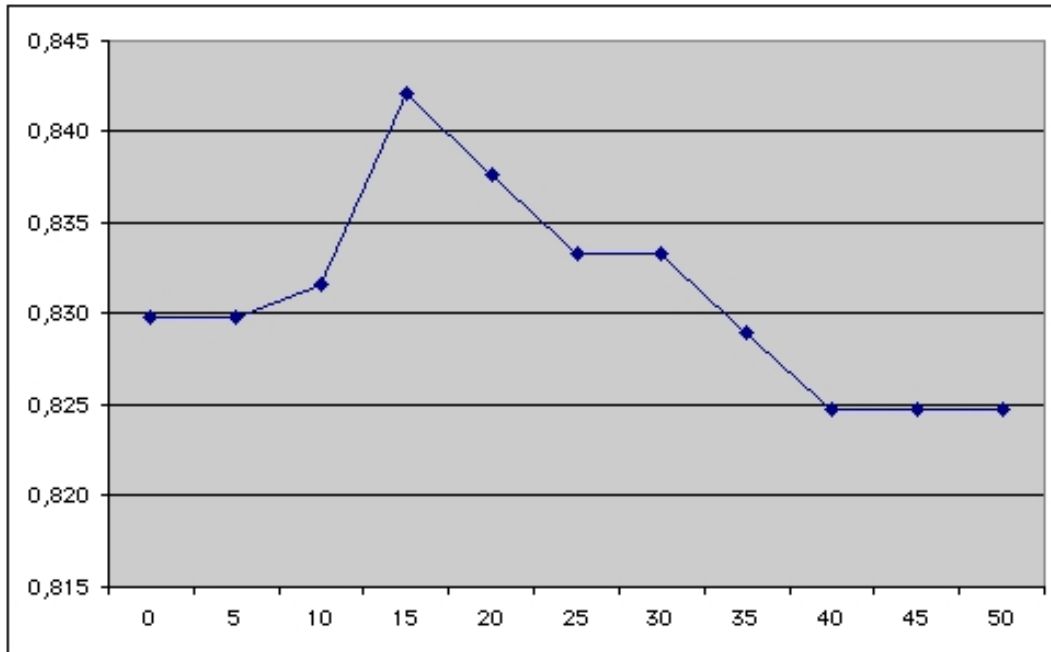


Figure 7.14: Average precision for increasing threshold values (between 0 and 50) using the column/row rule on the data from the logical labeler.

All additional misdetections of the reading order using the logical labeler data in place of the ground truth data are due to the misclassification of title objects. They are confused with footers, captions or rulers. The misclassification in the logical labeler data propagates to *SpaRe*. Eight additional documents are interpreted erroneously.

7.4.3 Discussion of the results

Varying the threshold in the thick boundary interpretation of Allen bidimensional relations does influence the overall performance considerably. In Figure 7.15, we compare the values of precision and recall for the three experimental cases increasing the threshold from 0 (no thickness) to 400 points. We notice that the precision increases considerably when the threshold goes from 0 to 5-10 points. Then it stabilizes showing minor variation over a wide range of thicknesses.

Moving the thickness from 0 to the maximum values corrects the situations in which boundary detection is not ideal. The reason for the stabilization of the precision between 15 and 100 points can be interpreted as follows. In a document, document objects need not be found perfectly aligned. As far as the variation is small, the document layout is still intelligible. The acceptable variation depends on the specific document. For example, in a multicolumn document without overlapping frames, it is

necessary to allow a small variation because the elements of a column will never be perfectly aligned; on the other hand, the variation should not go beyond half of the size of the white space between two adjacent columns otherwise columns will be confused.

Letting the thickness grow much beyond 100, makes the precision fall down as the thickness becomes too big with respect to the average document block size. The document objects become ‘blurred’ entities and *overlap* becomes the most frequent relation. Performance degrades rapidly.

Considering the maximum values in Figure 7.12, Figure 7.13, and Figure 7.14, we notice that the maximum value is different for different rules.

The recall is stable and has always a high score between 0.9 and 1.0. This makes this measure of little interest in the presented experimentation. The reason for this high values resides in the fact that only one reading order is considered for the documents.

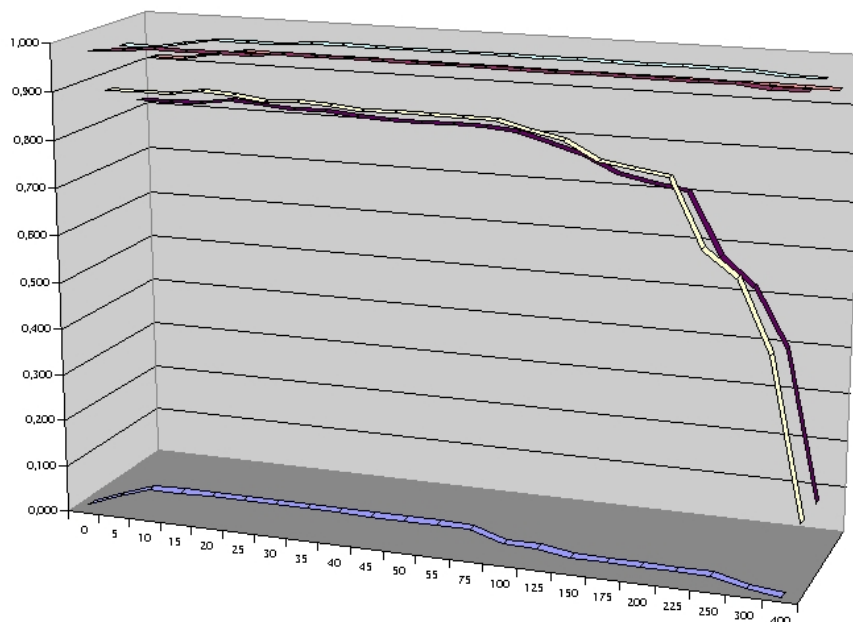


Figure 7.15: Comparing precision and recall for the three experimental cases with respect to increasing threshold (from 0 to 400). From foreground to background, the recall for the general rule on ground truth data, the recall for column/row rules on ground truth data, the recall for column/row rules on the logical labeler data, the precision for the general rule on ground truth data, the precision for column/row rules on ground truth data, and the precision for column/row rules on the logical labeler data.

From the comparison of the use of the column and row-wise rules on the ground truth and on the logical labeler data (with threshold set to 15), one notices a small degradation of the overall performance. On the whole collection this means an appreciable decrease in performance, but not a total brake-down of the approach, as the

precision goes from 0.89 to 0.84 and the recall from 0.98 to 0.94.

Considering the use of the general and the column and row-wise document encoding rules, one notices a big difference with respect to precision. The problem with the general document encoding rule is its generality. It loses almost none of the correct readings of a document, but it finds too many. For instance, for a three column document with an image in the central column composed of 14 textual document objects, the general rule gives 2714 admissible reading orders while using the column-wise rule one gets only the correct one. When performing the experiment with the column and row-wise rules, we appreciate the sharp increase in precision, while the recall remains unmodified. This means that the rules are less general to detect less reading orders, but are not too specific to degrade the performance. Even on a heterogeneous collection of documents such as the MTDB, the column and row-wise rules have high values of recall and, most notably, precision. It is safe to conclude that the general rule is of no interest when compared with the column and row-wise rules.

The average execution time of *SpaRe* is appreciably fast. On a standard Sparc 300 Mhz machine, it takes about 28 seconds of wall clock time to process the whole data set. The median execution value for a document is of 10 milliseconds. The execution time increases more than linearly with the number of document objects. Therefore, there is a practical upper bound to the complexity and richness of document components that can be analyzed.

7.5 Concluding remarks

We have shown the feasibility, and efficacy, of applying a symbolic approach to logical structure detection in the context of document image analysis and understanding. The approach is based on a spatial language of rectangles and basic mereotopological rectangle relations (bidimensional Allen relations). Inference is achieved via constraint satisfaction techniques.

We have shown a bidimensional Allen based language to have appropriate expressive power for the task of document understanding. Though, what the language misses is a notion of neighbourhood or some other kind of weak metric expressivity. Considering the 11% of the documents understood erroneously using the column and row-wise rules on the ground truth, one may argue that the correct order would have been captured by using a rule preferring neighboring text objects. Something not expressible in bidimensional Allen. In [Todoran et al., 2001a], we move the first steps in this direction by using Voronoi diagrams.

The logical labeler adds 4% of misclassified reading orders. Little can be modified in *SpaRe* to overcome these failures. When logical labels do not correspond with the actual logical function of the objects, any symbolic approach shows brittleness.

Two notable features of the presented symbolic approach are its flexibility and modularity. *SpaRe* is flexible enough to treat a wide variety of documents, including scientific articles, newspapers, magazines and commercial hand-outs, in a single run.

To increase the number of document classes handled, future work includes an extension to explicitly deal with independent reading orders. Independent reading orders are the case of complex documents, such as newspapers where pieces of text independent of one another coexist on the same sheet. The foreseeable key point of such an extension lies in the identification of appropriate document rules.

Regarding the issue of execution time for rich documents, there are more efficient alternatives. In [Aiello, 2002a], we propose the use of model checking techniques.

In conclusion, we do not know if HAL was equipped with a symbolic document image analysis system or with one based on different technologies. The only thing we know is that whenever a HAL-like machine will be available we expect it to read and understand the contents of any printed document brought to its attention.

8.1 Where we stand

Spatial structures and visual reasoning, in its broader sense, are the subject of this thesis. Our personal take on the matter is the attempt to bring together two research areas: the standard mathematical approach (topology, geometry, and linear algebra) with a computational analysis of visual processing tasks. To build such a bridge, we proposed a modal logic approach, which connects up with both:

- (i) more tractable levels of spatial structure inside mathematical theories, and
- (ii) more expressive power in computational tasks.

The results in the thesis show the connection meaningful by providing a number of tools which are both useful for ‘deconstructing mathematics’ and for the analysis and redesign of computational tasks. Next, we briefly summarize the main points.

Topo-approach. We proposed a framework for topological reasoning with a modal language of visual patterns, emphasizing bisimulation and comparison games as a means of calibrating similarity of visual scenes. Moreover, a pleasing side-effect was a new take on elementary topology. Laying the basis for a more ambitious program of ‘modal geometry’, exploring new fine-structure of tractable fragments of geometry; just as modal logic itself does for first-order logic.

Logical extensions. We proposed and reconsidered a number of languages to increase the expressive power of $\mathbf{S4}$ within the bounds of the topo-approach. This has included a new analysis of the universal language $\mathbf{S4}_u$ of Bennett [1995]. We showed $\mathbf{S4}_u$ to be a language of connected spaces whose simulations preserve the truth of existentially quantified formulas (the connection with connected spaces has also been presented in [Shehtman, 1999], the results were independently obtained). We introduced an even more expressive formalism: a spatial Since and Until logic.

Geometrical extensions. Our walk through geometrical spaces showed modal structures wherever one looks. There are natural fine-structured modal versions of affine and metric geometry. These can be studied by general modal techniques—though much of the interest comes from paying attention to special spatial features. The benefits of this may be uniformity and greater sensitivity to expressive and computational fine-structure in theories of space.

Logical fragments of mathematical morphology. We established preliminary connections between logical axiomatizations of mathematical morphology. The links were built on linear and on arrow logics, as both capture relevant fragments of what is a fundamental qualitative theory of shape.

Games as similarity measures. We introduced a similarity measure for spatial patterns based on model comparison games and implemented it in a image retrieval system.

Symbolic approach to document understanding. We showed the applicability of a symbolic approach to document image understanding. The use of a thick boundary interpretation of rectangular Allen relations has proven to be at the right expressive level to perform reading order detection. We implemented a system based on the framework which shows high accuracy when tested on heterogeneous collections of document images for which no specific document knowledge is available.

8.2 Final remarks on theory and practice

The words theory and practice may be dangerous. The risk we take is that the terms are considered in contraposition rather than as distinct aspects of the same research process; which has been our own experience. Still, we found a few concerns that differentiate more theoretical branches of spatial reasoning from more practical ones.

Ontology: regions vs. points. A long debated matter in temporal reasoning is the opposition of instant based ontologies with interval based ones, cf. [van Benthem, 1983b]. A similar dichotomy holds for spatial reasoning, opposing point-based theories to region based ones; the latter are more frequent in philosophy, artificial intelligence and cognitive science. Unfortunately, mathematical region-based theories are much scarcer than those based on points (cf. [Johnstone, 1977, Johnstone, 1982, Sambin, 1987, Vickers, 1994]), leaving researchers with few tools to approach the subject.

Our own experience shows that a theory of space must work with regions. What matters is that one can refer to regions and their properties. Our modal approach was designed to do just that. For instance, consider the languages $\mathbf{S4}$ and $\mathbf{S4}_u$. The first can express only properties of a point and its neighbourhoods. This has no immediate practical application. On the other hand, $\mathbf{S4}_u$ expresses properties of regions, their spatial structure, and their relations with other regions. These basic spatial descriptors make $\mathbf{S4}_u$ a promising candidate for applications, as we saw in Chapter 6. The design of the

formalism for document image analysis (Chapter 7) also needs to express properties of regions. In analyzing a document image, the prominent properties are those of the spatial arrangement of extended objects detected in the document, not those of specific points inside the document.

Boundaries: inferred vs. detected. The theories of space based on topology (Chapters 2–4) tend to put special emphasis on boundaries. After all, topology can be seen as the theory of connected entities and a connected entity is a collection of points up to a boundary. Now, boundaries are puzzling spatial entities which are located in space, but which do not take any space [Casati and Varzi, 1999, Aiello, 2001b]. Thus, while our spatial topological theories heavily rely on boundaries, the chances of detecting a boundary (in the sense of formal topology) in real life images and spatial patterns are none. A system relying on that detection is likely to be brittle and unsuccessful.

Of the two prototypes presented in the thesis, *SpaRe* is most affected by boundaries. As it works with real images, and uses topological regions, it is very sensitive to the precise location of boundaries. After a first round of experimentation, we realized that a number of erroneous analyses were due to boundary problems. We solved this by giving a different interpretation of boundaries, cf. Chapter 7.

Model classes: across vs. within. Theoretical research in spatial logics is interested in results *for* a specific class of models or *across* such classes. Take completeness: McKinsey and Tarski [1944] efforts went in showing completeness of **S4** with respect to the real line; [Shehtman, 1999] showed completeness of **S4_u**+ (the connectedness axiom) for connected topological spaces. Another example are Ehrenfeucht-Fraïssé games, which are typically used to compare across different structures.

In our applications, one is more interested in restricting attention *within* some particular class of models, and then use tools which behave uniformly on it. A typical example is our use of Ehrenfeucht-Fraïssé games to compare different images, viewed as constellations of regions in the same kind of mathematical space. In particular, the key step from theory to practice in Chapter 6 is a move from a general model comparison game to a distance measure within a fixed class of spatial structures.

Our analysis of space and of applications of spatial theories is only a small step which generates more questions than answers. We identified many new open problems along the way in the thesis. Thus, our work also serves as a pilot study for a broader modal geometry developed with a view to potential applications.

Most likely, the next spatial reasoning task that awaits us consists of closing the dissertation in hand and laying it down on a flat solid surface. Alternatively, by appropriate ‘point and click’-ing we shall get rid of the window containing the current text. Whatever we do next, there is just no way of avoiding spatial reasoning.

A topological space, in its general definition, is just a set with a tiny bit of extra structure. It is a collection of elements, a membership function and, in addition, a family of subcollections with three simple properties.

A.0.1. DEFINITION (TOPOLOGICAL SPACE). A *topological space* is a pair $\langle X, O \rangle$, where X is a set and $O \subseteq \mathcal{P}(X)$ a family of subsets of X such that:

1. $\emptyset \in O$ and $X \in O$,
2. O is closed under arbitrary unions,
3. O is closed under finite intersections.

Related definitions to that of a topological space follow.

- (i) An element of O is called an *open*. A subset A of X is called *closed* if $X - A$ is open.
- (ii) A point $s \in X$ is a *limit point* of a subset A of X if for each $o \in O$ such that $s \in o$, $(o - \{s\}) \cap A$ is not empty.
- (iii) The *interior* of a set $A \subseteq X$ is the union of all open sets contained in A .
- (iv) The *closure* of a set $A \subseteq X$ is the intersection of all closed sets containing A or, equivalently, the union of the set A with all its limit points.
- (v) Given a set A , the set of points y such that for any open set o containing y both $o \cap A \neq \emptyset$ and $o \cap (X - A) \neq \emptyset$ hold, is called the *frontier*, or *boundary*, of A .
- (vi) A family of open sets B is a *base* of the space X if all open sets are unions of members of B . Such a family is a *subbase* of X , if the collection of all finite intersections of elements of B is a base for X .

A.0.1. EXAMPLE (TOPOLOGICAL SPACES). Typical examples of topological spaces are the indiscrete topology, the discrete topology, metric spaces, and Cantor space.

- (i) **indiscrete topology** $\langle X, \{\emptyset, X\} \rangle$
- (ii) **discrete topology** $\langle X, \mathcal{P}(X) \rangle$
- (iii) **metric spaces** every metric space is a topological space. A base that builds up the topology is the family of sets $\{x \mid \text{distance}(x, p) < r\}$ for arbitrary points p of the space and nonnegative r . This is called the *standard* topology.
- (iv) **Cantor space** all infinite sequences of 0, 1. A base that builds up the topology is the family of sets A_σ , consisting of all the sequences extending the finite initial segment σ .

As a point of notation, when considering intervals in one dimensional metric spaces, we write (a, b) for $\{x \mid a < x < b\}$. Square brackets denote that the frontier point belongs to the interval, e.g. $(a, b]$ stands for $\{x \mid a < x \leq b\}$.

A.0.2. DEFINITION (CONNECTED SPACE). A topological space X is *connected* if the only sets which are both open and closed are \emptyset and X .

A.0.2. EXAMPLE (CONNECTED SPACE). Examples of connected spaces are the metric spaces \mathbb{R}^n with the standard topology, for any positive integer n . Non-connected spaces are the rationals \mathbb{Q} . E.g., consider the two non-empty open and closed sets $(-\infty, \sqrt{2})$ and $(\sqrt{2}, \infty)$.

A.0.3. DEFINITION (COMPACT SPACE). Let X be a topological space. A collection $V_i \in \mathcal{P}(X)$ is a *covering* of X if $\bigcup_i V_i = X$. It is an *open covering* if all the V_i are open. A topological space X is said to be *compact* if every open covering has a finite subcovering.

A.0.3. EXAMPLE (COMPACT SPACE). No space \mathbb{R}^n is compact. But all (and only) their *bounded* subsets are compact.

A.0.4. DEFINITION (DENSE). A set A in a topological space X , is said to be *dense* in X , if all points of X are a point or a limit point of A . A topological space is said to be *dense* if all its points are limit points for itself.

Another interesting way to discern topological spaces uses their richness in terms of points and open sets. If there are enough of them one can ‘separate’ points. This formally shows in so-called ‘separation axioms’:

A.0.5. DEFINITION (SEPARATION AXIOMS). A topological space X is called

- (i) T_0 if for any two distinct points x_1 and $x_2 (\in X)$, there exists an open set $o \in X$ containing one but not the other,
- (ii) T_1 if for any two distinct points x_1 and $x_2 (\in X)$, there exist an open set $o_1 \in X$ containing x_1 but not x_2 and there exists an open set $o_2 \in X$ containing x_2 but not x_1 ,
- (iii) T_2 (*Hausdorff*) as T_1 with the additional requirement that $o_1 \cap o_2 = \emptyset$,
- (iv) T_3 (*regular*) as T_1 and for every closed set and point not contained in it there exist two disjoint open sets containing the point and the closed set respectively,
- (v) T_4 (*normal*) as T_1 and for every two closed disjoint sets there exists two disjoint open sets each containing one of the closed sets.

The fundamental way to move from a topological space to another space is through continuous mappings. Those preserving, among all, the property of openness.

A.0.6. DEFINITION (CONTINUITY). A map $f : X \rightarrow X'$ between two topological spaces $\langle X, O \rangle, \langle X', O' \rangle$ is *continuous* if for all opens $o' \in O'$, $f^{-1}[o']$ is in O : i.e., inverse images of open sets are open.

Continuous mappings are the building block for defining the equivalence of topological spaces. If two continuous mappings exist that composed, either way, yield the identity on each space, then the two spaces are topologically equivalent. The equivalence is named in topology *homeomorphism*.

A.0.7. DEFINITION (HOMEOMORPHISM). Two topological spaces $\langle X, O \rangle, \langle X', O' \rangle$ are *homeomorphic* if there are continuous maps $f : X \rightarrow X'$ and $g : X' \rightarrow X$ such that $f \circ g, g \circ f$ are both identity maps.

A basic topological fact about homeomorphic spaces is that of having the same cardinality. But the converse is not generally true: topology demands more structure than pure counting.

A.0.4. EXAMPLE (HOMEOMORPHISM). The two subsets $(0, 1)$ and $(1, \infty)$ of the metric space \mathbb{R} with the standard topology are homeomorphic. The two inverse functions $f(x) = g(x) = \frac{1}{x}$ are continuous and compose to identity maps both ways. By a similar construction of homeomorphisms, the real plane \mathbb{R}^2 and a unit circle $x \in \mathbb{R}^2 : d(x, 0) < 1$ are homeomorphic. Also Cantor space is homeomorphic to $[0, 1]$.

Two non-homeomorphic spaces are the real plane \mathbb{R}^2 and a three dimensional unit ball $x \in \mathbb{R}^3 : d(x, 0) < 1$.

Topology also provides a more general notion than homeomorphism.

A.0.8. DEFINITION (HOMOTOPY). Let X and X' be topological spaces, and let f_0 and f_1 be continuous maps from X to X' . f_0 is *homotopic* to f_1 (notation $f_0 \simeq f_1$) if there exists a continuous map $F : X \times I \rightarrow X'$ such that for all x $F(x, 0) = f_0(x)$ and $F(x, 1) = f_1(x)$, where I is $[0, 1]$. F is called an *homotopy* from f_0 to f_1 .

A.0.9. DEFINITION (HOMOTOPY TYPE). Two topological spaces X and X' are of the same *homotopy type* if there exists two continuous maps $f : X \rightarrow X'$ and $g : X' \rightarrow X$ such that $g \circ f$ is homotopic to the identity mapping on X and $f \circ g$ is homotopic to the identity mapping on X' .

A.0.5. EXAMPLE (HOMOTOPY). Homeomorphic spaces are also homotopic. Therefore, an example of homotopic spaces is the real line and the real unit interval (see Example A.0.4). A more interesting example is the homotopy between a single point and any real metric space \mathbb{R}^n .

The real plane without its origin $\mathbb{R}^2 - (0, 0)$ and the unit circle are an example of non-homotopic spaces.

APPENDIX B

SORTING TRANSITIVE DIRECTED GRAPHS

We extend the notion of *topological sorting* a directed acyclic graph [Knuth, 1968, Knuth and Szwarcfiter, 1974]. Instead of a directed ‘acyclic’ graph, we sort a directed ‘cyclic’ graph whose edge relation is transitively closed. We call the latter *directed transitive cyclic graph*. More formally, a directed transitive cyclic graph is a graph $G = \langle V, E \rangle$ such that if $(i, j) \in E$ and $(j, k) \in E$, then $(i, k) \in E$. In what follows, we assume that there are n vertices $|V| = n$ and m edges $|E| = m$. The problem of sorting a directed transitive graph G consists of creating sequences of nodes of the graph such that for any pair of nodes u and v in G appearing in any sequence, then (u, v) must be an edge of G .

Algorithms to perform topological sorting of directed acyclic graphs work iterating the following procedure until all nodes have been visited. First, a node v with no predecessors

$$\forall u \neq v \neg \exists (u, v) \in E$$

is identified. The node v is placed in the output. Then, all the edges (v, u) such that $\forall u \neq v (v, u) \in E$ are removed from the graph. In other words, the set of edges E of the graph is replaced by its subset $E / \{(v, u) \in E\}$ without the edges departing from the node v . If the original graph is acyclic, then the algorithm outputs a topological sorting of the input graph, otherwise the output is incorrect. The complexity of this sort of algorithms is $O(m + n)$. Notice that the algorithm does not return any clue on the incorrectness of the output in the case the input graph is cyclic. This is rather natural when considering the complexity of topological sorting and that of identifying cycles in directed graphs. It is well known that the latter is in NL-hard (see, for instance, [Toda, 1990]).

The algorithm for sorting transitive cyclic directed graphs takes as input a connected graph $G = \langle V, E \rangle$ and outputs a sequence of nodes $v_1 \cdot v_2 \cdot v_3 \cdot \dots \cdot v_n$ such that:

1. for all $i: v_i \in V$,
2. $|v_1 \cdot v_2 \cdot v_3 \cdot \dots \cdot v_n| = |V|$,

3. for all $i \neq j$: $v_i \neq v_j$,
4. if $i < j$: $(v_i, v_j) \in E$.

One starts by removing all self-loops $(v, v) \in E$ to setup the graph. Then the main cycle of the algorithm begins by considering all the nodes and counting the number of edges departing from each one, also known as *degree* of the node: $\text{deg}(v) = |\{(v, w) \in E \mid w \in V\}|$. Then one chooses a node with the highest degree, which has to be the same as the number of nodes of the graph minus one. In other words, the node is related — is ‘before’ — all other nodes of the graph. As we allow for cycles, there can be more than one node satisfying this condition. Once a node with maximal degree has been chosen, we remove it from the graph together with all the edges connected to it, both outgoing and incoming, and repeat the procedure on the remaining subgraph.

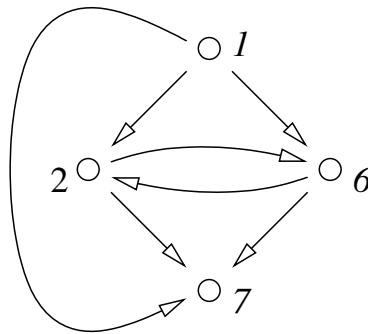


Figure B.1: A simple directed transitive cyclic graph.

Consider the simple example in Figure B.1. The input graph is $G = \langle \{1, 2, 6, 7\} \{(1, 2), (1, 6), (1, 7), (2, 6), (2, 7), (6, 2), (6, 7)\} \rangle$, it is easy to check that it meets the input conditions. The first step of the algorithm is to create a list of nodes and their occurrences: $L = \{(1, 3), (2, 2), (6, 2), (7, 0)\}$. The node 1 is selected as first node of the output, as its degree is $3 = |V| - 1$. The list L is then updated to $L = \{(2, 2), (6, 2), (7, 0)\}$. Two choices are possible at the following iteration: either 2 or 6. Suppose the first item is chosen, then L becomes $\{(6, 1), (7, 0)\}$. Finally, the output is updated with 6 and 7, respectively, yielding the final output of $\{1, 2, 6, 7\}$ (also $\{1, 6, 2, 7\}$ is a correct solution, and it can be computed by backtracking to the point in which v_2 was chosen in place of v_3).

Let us now proceed with a more precise definition of the algorithm. The preliminary step of the algorithm consists of the construction of a list L of pairs (v, o) , where o is the degree of v , i.e., $o = \text{deg}(v)$. In pseudo-code, we have:

```

fail ← false;
for all  $v$  such that  $(v, v) \in E$ ;
     $E \leftarrow E/(v, v)$ ;
while ( $|V| > 0$  and (not fail))
    sort  $L$  in descending order of occurrences
    % let  $(v^*, l)$  be the first element of  $L$ 
    if ( $l \neq |V| - 1$ )
        fail ← true;
    else
        output ← output +  $v^*$ ;
         $V \leftarrow V/v^*$ ;
        for all  $w$  such that  $(v, w) \in E$  or  $(w, v) \in E$ ;
             $E \leftarrow (E/(v^*, w)) \cap (E/(w, v^*))$ ;
        update ( $L$ );

```

Given that sorting a set of up to n values, each of which is an integer in the interval $[0, n-1]$ can be performed with a bucket sort in $O(n)$, one can conclude that the complexity of the proposed algorithm is in the $O(n^2)$ class.¹ If the algorithm terminates with `fail` set to false, then a correct sorting of the original directed transitive graph G will be found in the variable `output`. If no check is performed on the input graph, nothing can be said in case the algorithm returns true for the variable `fail`. On the other hand, if the input graph is tested to be transitively closed, then `fail` set to true indicates that no sorting for the input graph G exists. Algorithms to transitively close a graph can be found in the literature [Warshall, 1962, Munro, 1971, Arlazarov et al., 1970], and are also relatively inexpensive: $O(n^3)$, $O(n^{2.376})$, and $O(\frac{n^3}{\ln(n)})$, respectively.

¹It is possible to devise an algorithm for directed transitive graph sorting with lower asymptotic complexity, though this is beyond the scope of the presented material. The steps of such an algorithm consist of: 1) finding strongly connected components of the graph, which is in $O(n + m)$; 2) consider the graph of the strongly connected components; 3) topologically sort the new graph. This algorithm has $O(n + m)$ complexity where n is the number of nodes and m the number of edges.

This appendix consists of a number of short system descriptions and the presentation of selected declarative code. The systems described are:

1. the applet `topax` for the visualization of the selective unraveling technique as presented in Chapter 3,
2. the image retrieval prototype `IRIS` described in Chapter 6, and
3. the document analysis prototype `SpaRe` described in Chapter 7.

The full source code for these systems is available electronically at <http://www.aiellom.it/phd/source>. Other implementations related to the thesis can be found at <http://www.aiellom.it/java>, including one of Ehrenfeucht-Fraïssé games. The latter is described and motivated in [Agostini and Aiello, 1999].

C.1 Topax

What follows is the content of the web page <http://www.aiellom.it/java/topax>. It is a Java applet for the visualization of the selective unraveling presented in Section 3.3.1, together with instructions on how to use the applet and some motivations. The centered text in `typewriter` font does not appear in the web page and was added for this presentation. The colors refer to the web-page and the electronic pdf version of the thesis (the hard copy of the thesis has only gray-levels). The contents of the web-page start on the next page.

This page presents a Java applet for the visualization of the modal logic construction presented in [Aiello et al., 2001]. Instructions on how to use the applet and its motivation can be found below on this page.

NOTE, if a window has not popped-up, it means that the applet is not running properly. Please refer to the troubleshooting section.

[In the web-page, the applet appears here.]

Motivations and use

The aim of the applet is the visualization of a construction relating Kripke semantics and topological semantics for modal logics, in particular, for the modal logic S4.

A bit of history

The first completeness result for the modal logic S4 (we refer to a standard book in modal logic for its syntax and its standard Kripke semantics such as [Blackburn et al., 2001]) was given by Tarski in the late 30s [Tarski, 1938]. Later, together with McKinsey [McKinsey and Tarski, 1944], Tarski showed S4 to be complete with respect to any metric space without isolated points. The topological interpretation was somewhat abandoned when the possible worlds semantics was introduced for modal logics thanks to the independent efforts of a number of researchers, including Kripke. The graph like possible worlds semantics made modal logics more accessible and easy to use, completely replacing in common practice the topological semantics for modal logics. Recently the topological interpretation has received new attention in relation to spatial representation and reasoning (e.g., [Bennett, 1995]).

Standard Kripke models for S4

A known fact for the logic S4 is that its models can be viewed as trees of mutually accessible clusters. This means that a model can be partitioned into a number of clusters (cliques) of worlds which are all mutually accessible. An example of a cluster of 4 worlds is presented on the right (Figure C.1). The various clusters are ordered from a higher cluster that can access all other clusters to those which can access none. An example is given by the picture below (Figure C.2).

The model is not a tree with respect to the possible worlds (the blue circles), but rather a directed graph. Though, if considering the clusters (the red rectangles) as the basic elements and considering the green arrows, then the S4 model is a tree.

Evaluating colors

A model comprises a valuation function usually assigning a propositional letter to each world. Now viewing every propositional letter as a (different) color, we can visualize a valuation as a coloring of each world (may be in many different colors). In the two

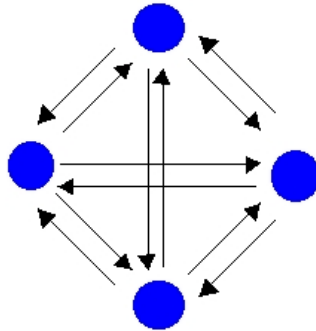


Figure C.1: The blue circles are possible worlds of the models and the arrows represent an element of the accessibility relation. This is a cluster of 4 mutually accessible worlds.

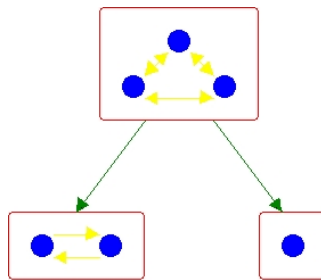


Figure C.2: The red rectangles denote the clusters of mutually accessible nodes (for which the accessibility relation is given by the yellow arrows). A green arrow means that all the worlds in the origin cluster are related to all the worlds in the endpoint cluster.

pictures above, the valuation function related all worlds with the color blue. In the applet, we can evaluate a world of the starting tree of clusters to any color.

Towards topological spaces

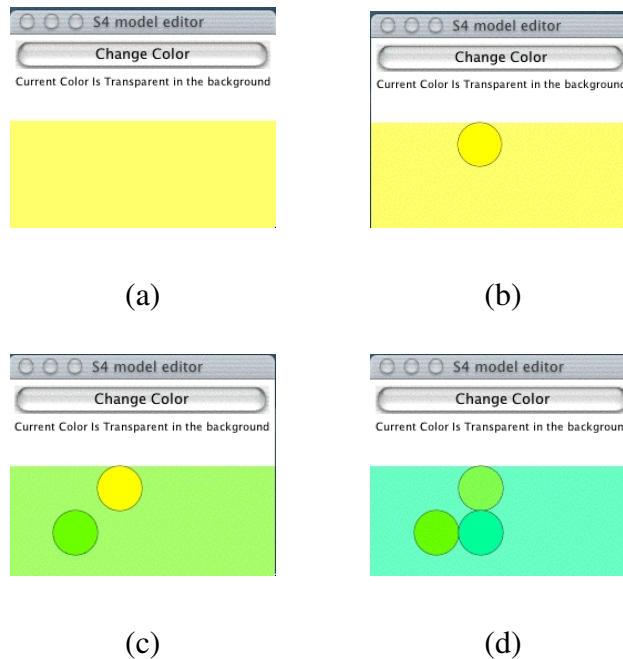
The tree of clusters of mutually accessible points can also be regarded as a topological space. In fact, it is an Alexandroff's space. A possible world becomes a point of the Alexandroff space, while all accessible worlds from a given one define its least open neighborhood [Vickers, 1989]. One can achieve more, and move from an Alexandroff space to the Cantor space, and then to the real line. To achieve completeness on the Cantor space, we selectively unravel an Alexandroff space generated by a tree of clusters into the Cantor space (an infinite complete binary tree). The full formal

description of selective unraveling is presented in [Aiello et al., 2001], while the tool to show the correctness of the transition is that of topo-bisimulations, introduced in [Aiello and van Benthem, 1999]. Note that in the topological space model of S4, the worlds are not the nodes of the Cantor tree, but rather the full infinite branches. So the colors visualized by the applet are only the preliminary colors assigned to the branch (not the colors assigned to obtain the exact topo-bisimulation, again refer to [Aiello et al., 2001] for the details of the construction).

Using the applet

First, one creates an S4 model as a tree of clusters of colors. The window S4 model editor serves this purpose. The background color of the window is the current color. By clicking on the window a cluster is created with the current color as the color of one of the worlds of the cluster. By pressing the Change Color button, a color chooser window pops up. In this manner one can choose the new current color. The current color becomes the new background color of the window. By clicking on an existing cluster a node of the current color is added to the cluster. If clicking outside any cluster, a new one is created. The clusters are represented by the average color of all their worlds. The tree is built considering as root the upper cluster (if more clusters are on the highest row of the window, a dummy cluster with one world of white color is the root), then the clusters below are considered as children in the tree and associated to the closest cluster going first up and the right.

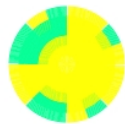
We present an example of using the applet. First, one builds the S4 model as a tree of clusters.



In the S4 model editor window with yellow background, we click in the window

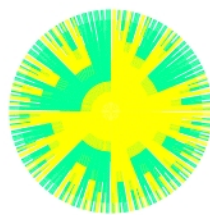
and we create a first cluster with one world of yellow color, (a) and (b). Then we press the Change Color button and select a new color: green. We click below the yellow cluster and create a new cluster with one world of green color (c). We then select blue as the current color. We click on the yellow cluster (which now gets as color the average of yellow and blue, i.e., green). Finally, we click below the yellow/blue cluster to obtain a new blue cluster (d).

When we close the S4 model editor window, the model we have just built becomes the current model. By pressing the Paint button the model is selectively unraveled into the Cantor space. In particular the default values for Depth of rendering and for Visualization mode are used.

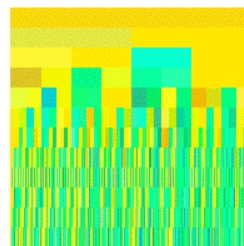


The default value for the depth of rendering is 5, while the default value for the visualization mode is circle. The latter means that the nodes of the binary tree are represented as thick circumferences. The root is the central circle. Then the second level of the binary tree is the surrounding circumference, which is divided in two half circumferences. The left son of the root is the left half circumference, the right son is the right half circumference.

By setting the depth of rendering to 12 and the Color Randomness to 0.4 we obtain the renderings



(e)

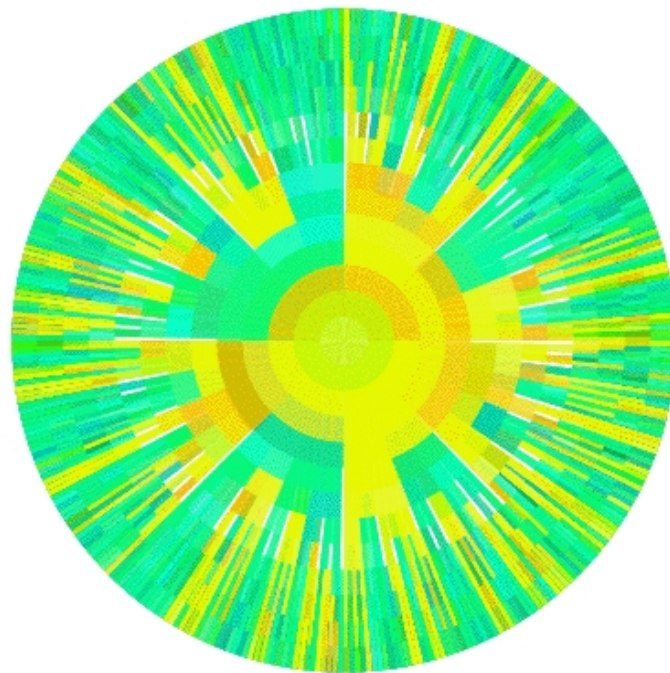


(f)

depending on the Visualization mode. (e) uses the circle mode, while (f) uses square mode. In the latter, the root of the Cantor tree is on top and the sons of each node are in the line below. The length of a son is the half of the one of the parent.

The color randomness lets a node be rendered with a little variance from the original color in the tree of clusters. In this way, it is possible to better identify the single nodes in the Cantor tree and to obtain more appealing renderings. (Please be careful in setting the depth of rendering. Depending on the computing power of your machine, you may be waiting for a long time or even get an out of memory error. On my Mac 12 is the limit, while on the Sun at work it is 18.)

Finally, by bringing the color randomness to 0.5 and moving the visualization mode to big circle we obtain the following rendering of the selective unraveling



Two final remarks:

- by rearranging the colors one can move further on to the real interval $(0,1)$. For details we refer to [Aiello et al., 2001].
- the Cantor space obtained has a fractal structure. It is easy to identify patterns that repeat themselves. Subtrees of the unraveled model are identical to the whole model and different subtrees at different depths are also identical with respect to each other. For instance, one can see the green pattern to repeat itself in the circle visualization mode above. It starts on the top-left part of the circle and it repeats moving counter-clockwise on the figure.

Troubleshooting

The applet will only run in Java2 enabled browsers. If your browser does not support Java2, you may try to download the whole applet on your machine and run it with

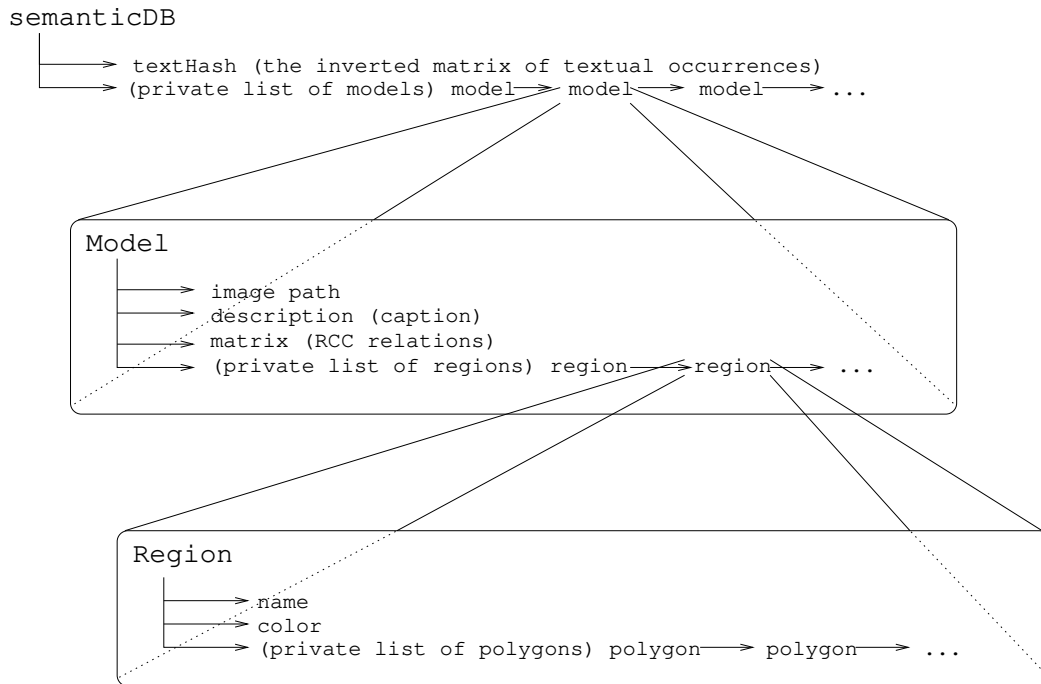


Figure C.3: The design of the spatial data structures.

appletviewer. In alternative, you can try the page with the Java-plug-in.

Contact and bibliographic information follows.

C.2 IRIS

The image retrieval prototype IRIS, presented in Chapter 6 is implemented in Java. Here we present the main data structures behind the implementation.

The spatial data structures

The spatial data structures are implemented according to the schema presented in Figure C.3. Classes are presented together with their most relevant variables. See Figure 6.7 for a more functional view of the data structures within IRIS.

```

public class semanticDB
extends java.lang.Object
implements java.io.Serializable
  
```

Field Summary

textHash	<i>textMatrix</i>
----------	-------------------

Constructor Summary

semanticDB()

Method Summary

void	addModel(model m)
java.awt.Graphics	draw(java.awt.Graphics g)
int	getFree()
java.lang.String	paramString()
void	remove())
int	scanModels())
model	scanModels(int i)

```
public class model
extends java.lang.Object
implements java.io.Serializable
```

Field Summary

java.lang.String	<i>description</i>
java.lang.String	<i>imagePath</i>
int[][]	<i>matrix</i>

Constructor Summary

model()

Method Summary

void	addRegion(region r)
void	computeMatrix()
int	contains(java.lang.String name)
java.awt.Graphics	draw(java.awt.Graphics g)
int	freeRegions()
java.lang.String	paramString()
void	printMatrix()
void	remove()
region	scanRegions(int i)

```
public class region
extends java.lang.Object
implements java.io.Serializable
```

Field Summary

<code>java.awt.Color</code>	<i>color</i>
<code>java.lang.String</code>	<i>description</i>
<code>int</code>	<i>freePolygons</i>
<code>java.lang.String</code>	<i>name</i>
<code>boolean[]</code>	<i>open</i>
<code>java.awt.Polygon[]</code>	<i>polygons</i>
<code>boolean</code>	<i>transparent</i>
<code>boolean[]</code>	<i>vector</i>

Constructor Summary

`region()`
Method Summary

<code>void</code>	<code>addPolygon(java.awt.Polygon p, boolean o)</code>
<code>void</code>	<code>computeVector()</code>
<code>java.awt.Graphics</code>	<code>draw(java.awt.Graphics g, int xshift, int yshift)</code>
<code>void</code>	<code>endRegion()</code>
<code>boolean[]</code>	<code>getVector()</code>
<code>boolean</code>	<code>printVector()</code>
<code>void</code>	<code>remove()</code>
<code>void</code>	<code>reset()</code>

C.3 SpaRe

SpaRe consists of an Eclipse program using the finite domain library¹ and a number of Perl scripts. The Perl scripts, which are not documented here, serve for the analysis of the output and to coordinate SpaRe with the other modules of the document image analysis system. We present source code in the thesis as an useful companion to Chapter 7. Being declarative code it should be fairly readable.

Selected Eclipse passages

The following listing of Eclipse clauses is not the full SpaRe implementation, but just the most relevant portions. It starts with the invocations of Eclipse libraries, then sets the type of analysis to perform and which document rules to adopt. Then there is the main clause as called by the overall document analysis system (`gO`) which takes as a input a list of documents (i.e., a list of document object positions and labels) and returns a list of lists of admissible reading orders. Following there is the body of clauses necessary to check the document rules on the given input. Then, there is

¹<http://www-icparc.doc.ic.ac.uk/eclipse>.

the encoding of the various document rules to be used in the analysis. Finally, Allen relations are defined for both axes. Note the use of a threshold in the definition, which slightly deviates from the usual interpretation of Allen's relations.

```

:-lib(fd).
:-lib(listut).

%*****
%
% type of analysis, set from calling function
%
%*****

threshold(15).

% this can be one of the following:
%   general, verticalColumns, horizontalColumns

rule_set(verticalColumns).

% this can be one of the following:
%   general, small caption, big caption
rule_figure(general).
rule_title(general).

%*****
%
% main call to SpaRe
%
%*****

go_each(H, Stream):-
    scan(25,H,Texts,Stream),
    scanNoWrite(20,H,Titles,Stream),
    merge_titles(Titles,Texts,TBcouples),
    text_analyze(Texts,[],OutputVer,verticalColumns),
    text_analyze(Texts,[],OutputHor,horizontalColumns),
    text_analyze(Titles,[],Output3,general), !,
    quicksort_couples(OutputVer,SortedVer), !,
    quicksort_couples(OutputHor,SortedHor), !,
    quicksort_couples(Output3,Sorted3), !,
    elements(Texts,Texts_elements),
    elements(Titles,Titles_elements),
    findall(PathVer,path(Texts_elements,SortedVer,PathVer),
        BVreading_orders),
    findall(PathHor,path(Texts_elements,SortedHor,PathHor),
        BHreading_orders),

```

```

findall(Path3, path(Titles_elements, Sorted3, Path3),
        Treading_orders),
append(BVreading_orders, BHreading_orders,
        BHVreading_orders),
merge_title_text_all(BHVreading_orders, Treading_orders,
        Tbcouples, OutputDuplicates),
remove_dups(OutputDuplicates, OutputEmptyList),
remove([], OutputEmptyList, Output),
length(Output, M),
nl, write('Number of pahts '), write(M), nl,
length(Titles, Ltitles),
length(Texts, Ltexts),
Blocksnumber is Ltitles+Ltexts,
factorial(Blocksnumber, Fact),
writeln(Stream, [Blocksnumber, Fact, M]).

%*****
%
% Checking the rules on the
% input for body text
%
%*****

text_analyze([B1|Rest], Old, Out, R):-
    append(Rest, Old, Checklist),
    text_check(B1, Checklist, Out2, R),
    append([B1], Old, Old2),
    text_analyze(Rest, Old2, Out3, R),
    append(Out3, Out2, Out).

text_analyze([], _, [], _).

% CHECK TITLE BLOCKS AGAINST BODY TEXT

text_check([Id1, T1, B1], [[Id2, T2, B2]|Rest], Out, R):-
    before_in_reading(B1, B2, R),
    text_check([Id1, T1, B1], Rest, Out2, R),
    append([[Id1, Id2]], Out2, Out).

text_check(B1, [_|Rest], Out, R):-
    text_check(B1, Rest, Out2, R),
    append([], Out2, Out).

text_check(_, [], [], _).

```

```

%*****
%
% Checking the rules on the
% input for title
%
%*****

merge_titles([[Id, T1, [X1,X2,Y1,Y2]]|TT],Texts,Merged):-
  title_check([Id, T1, [X1,X2,Y1,Y2]],Texts,Blocks),
  take_leftuppermost(Blocks,[Id2|_]),
  merge_titles(TT,Texts,Merged2),
  append([[Id,Id2]], Merged2, Merged).

merge_titles(_,_,[]).

take_leftuppermost([[Id, T1, [X1,Y1,X2,Y2]]|T],[Idb, T1b,
  [X1b,Y1b,X2b,Y2b]]):-
  take_leftuppermost(T,[Idb, T1b, [X1b,Y1b,X2b,Y2b]]),

take_leftuppermost([[Id, T1, [X1,Y1,X2,Y2]]|T],[Idb, T1b,
  [X1b,Y1b,X2b,Y2b]]):-
  take_leftuppermost(T,[Id, T1, [X1,Y1,X2,Y2]]),
  X1b#>=X1, Y1b#>=Y1.

take_leftuppermost([B],B).

% CHECK TITLE BLOCKS AGAINST BODY TEXT

title_check([Id1, T1, B1], [[Id2, T2, B2]|Rest], Out):-
  title_body(B1, B2),
  title_check([Id1, T1, B1], Rest, Out2),
  append([[Id2, T2, B2]],Out2, Out).

title_check(B1, [_|Rest], Out):-
  title_check(B1, Rest, Out2),
  append([], Out2, Out).

title_check(_, [], []).

%*****
%
% Encoding of Layout rules in
% rectangle model
%
%*****

```

```

% GENERAL %

before_in_reading(B1, B2, general):-
    precedes_X(B1, B2).

before_in_reading(B1, B2, general):-
    meets_X(B1, B2).

before_in_reading(B1, B2, general):-
    overlaps_X(B1, B2).

before_in_reading(B1, B2, general):-
    precedes_Y(B1, B2).

before_in_reading(B1, B2, general):-
    meets_Y(B1, B2).

before_in_reading(B1, B2, general):-
    overlaps_Y(B1, B2), precedes_X(B1,B2).

% VERTICAL COLUMNS %

before_in_reading(B1, B2, verticalColumns):-
    precedes_X(B1, B2).

before_in_reading(B1, B2, verticalColumns):-
    meets_X(B1, B2).

before_in_reading(B1, B2, verticalColumns):-
    overlaps_X(B1, B2),
    (precedes_Y(B1,B2); meets_Y(B1,B2); overlaps_Y(B1,B2)).

before_in_reading(B1, B2, verticalColumns):-
    (precedes_Y(B1, B2); meets_Y(B1,B2); overlaps_Y(B1,B2)),
    (precedes_X(B1,B2); meets_X(B1,B2); overlaps_X(B1,B2);
    starts_X(B1,B2); finishesi_X(B1,B2); equals_X(B1,B2);
    during_X(B1,B2); duringi_X(B1,B2); finishes_X(B1,B2);
    startsi_X(B1,B2); overlapsi_X(B1,B2)).

% HORIZONTAL COLUMNS %

before_in_reading(B1, B2, horizontalColumns):-
    precedes_Y(B1, B2).

before_in_reading(B1, B2, horizontalColumns):-
    meets_Y(B1, B2).

```



```
before_in_reading(B1, B2, horizontalColumns):-
    overlaps_Y(B1, B2),
    (precedes_X(B1,B2); meets_X(B1,B2); overlaps_X(B1,B2)).
```

```
before_in_reading(B1, B2, horizontalColumns):-
    (precedes_X(B1, B2);meets_X(B1,B2);overlaps_X(B1,B2)),
    (precedes_Y(B1,B2); meets_Y(B1,B2); overlaps_Y(B1,B2);
    starts_Y(B1,B2); finishes_i_Y(B1,B2); equals_Y(B1,B2);
    during_Y(B1,B2); duringi_Y(B1,B2); finishes_Y(B1,B2);
    startsi_Y(B1,B2); overlapsi_Y(B1,B2)).
```

```
%%%
%   RULES FOR TITLES
%%%
```

```
title_body(T,B):-
    (precedes_Y(T,B);meets_Y(T,B)).
```

```
%%%
%   RULES FOR FIGURES
%%%
```

```
% general
```

```
make_one_block(B1, B2):-
    rule_figure(general),
    figure(B1),
    caption(B2),
    precedes_Y(B1, B2).
```

```
make_one_block(B1, B2):-
    rule_figure(general),
    figure(B1),
    caption(B2),
    precedesi_Y(B1, B2).
```

```
% smallCaption
```

```
make_one_block(B1, B2):-
    rule_figure(smallCaption),
    figure(B1),
    caption(B2),
    precedes_Y(B1, B2),
    (startsi_X(B1, B2); duringi_X(B1, B2);
    finishesi_X(B1, B2); equals_X(B1, B2)).
```

```

make_one_block(B1, B2):-
    rule_figure(smallCaption),
    figure(B1),
    caption(B2),
    precedesi_Y(B1, B2),
    (startsi_X(B1, B2); duringi_X(B1, B2);
    finishesi_X(B1, B2); equals_X(B1, B2)).

%*****
%
% Allen's interval relations
%
%*****

% be careful that the threshold should never be
% bigger than half of the smallest document object!!!
% I'm not implementing this check.

precedes_X([_, _, Xf1, _], [Xo2, _, _, _]):-
    threshold(T),
    Xo2-Xf1 #>= T.

meets_X([_, _, Xf1, _], [Xo2, _, _, _]):-
    threshold(T),
    Xf1-Xo2 #<T, Xo2-Xf1 #< T.

overlaps_X([Xo1, _, Xf1, _], [Xo2, _, Xf2, _]):-
    threshold(T),
    Xf1-Xo2 #>= T,
    Xf2-Xf1 #>= T,
    Xo2-Xo1 #>=T.

startsi_X([Xo1, _, Xf1, _], [Xo2, _, Xf2, _]):-
    threshold(T),
    Xo1-Xo2#< T, Xo2-Xo1#< T,
    Xf1-Xf2#>= T.

during_X([Xo1, _, Xf1, _], [Xo2, _, Xf2, _]):-
    threshold(T),
    Xo1-Xo2 #>= T,
    Xf2-Xf1 #>= T.

finishes_X([Xo1, _, Xf1, _], [Xo2, _, Xf2, _]):-
    threshold(T),
    Xf1-Xf2 #< T, Xf2-Xf1 #<T,
    Xo2-Xo1 #>= T.

equals_X([Xo1, _, Xf1, _], [Xo2, _, Xf2, _]):-
    threshold(T),
    Xo1-Xo2 #< T, Xo2-Xo1 #< T,

```

178 • *Appendix C. IMPLEMENTATIONS*

```

Xf1-Xf2 #< T, Xf2-Xf1 #< T.

finishes_X(B1, B2):-
    finishes_X(B2, B1).

during_X(B1, B2):-
    during_X(B2, B1).

startsi_X(B1, B2):-
    starts_X(B2, B1).

overlaps_X(B1, B2):-
    overlaps_X(B2, B1).

meetsi_X(B1, B2):-
    meets_X(B2, B1).

precedesi_X(B1, B2):-
    precedes_X(B2, B1).

% AND ON THE Y AXIS

precedes_Y([_, _, _, Yf1], [_, Yo2, _, _]):-
    threshold(T),
    Yo2-Yf1 #>= T.

meets_Y([_, _, _, Yf1], [_, Yo2, _, _]):-
    threshold(T),
    Yf1-Yo2 #<T, Yo2-Yf1 #< T.

overlaps_Y([_, Yo1, _, Yf1], [_, Yo2, _, Yf2]):-
    threshold(T),
    Yf1-Yo2 #>= T,
    Yf2-Yf1 #>= T,
    Yo2-Yo1 #>=T.

starts_Y([_, Yo1, _, Yf1], [_, Yo2, _, Yf2]):-
    threshold(T),
    Yo1-Yo2#< T, Yo2-Yo1#< T,
    Yf1-Yf2#>= T.

during_Y([_, Yo1, _, Yf1], [_, Yo2, _, Yf2]):-
    threshold(T),
    Yo1-Yo2 #>= T,
    Yf2-Yf1 #>= T.

finishes_Y([_, Yo1, _, Yf1], [_, Yo2, _, Yf2]):-
    threshold(T),
    Yf1-Yf2 #< T, Yf2-Yf1 #<T,
    Yo2-Yo1 #>= T.

```

```
equals_Y([_, Yo1, _, Yf1], [_, Yo2, _, Yf2]):-  
    threshold(T),  
    Yo1-Yo2 #< T, Yo2-Yo1 #< T,  
    Yf1-Yf2 #< T, Yf2-Yf1 #< T.  
  
finishes_i_Y(B1, B2):-  
    finishes_Y(B2, B1).  
  
during_i_Y(B1, B2):-  
    during_Y(B2, B1).  
  
starts_i_Y(B1, B2):-  
    starts_Y(B2, B1).  
  
overlaps_i_Y(B1, B2):-  
    overlaps_Y(B2, B1).  
  
meets_i_Y(B1, B2):-  
    meets_Y(B2, B1).  
  
precedes_i_Y(B1, B2):-  
    precedes_Y(B2, B1).
```

BIBLIOGRAPHY

- [Adams, 1981] Adams, A. (1981). *The Negative*. Brown Little and Company.
- [Agostini and Aiello, 1999] Agostini, A. and Aiello, M. (1999). Teaching via the Web: A self-evaluation game using Java for learning logical equivalence. In Bra, P. D. and Leggett, J., editors, *Proc. of WebNet99*, pages 75–80. AACE.
- [Aiello, 2000] Aiello, M. (2000). Topo-distance: Measuring the difference between spatial patterns. In Ojeda-Aciego, M., P. de Guzmán, I., Brewka, G., and Moniz Pereira, L., editors, *Logics in Artificial Intelligence (JELIA 2000)*, LNAI, 1919, pages 73–86. Springer.
- [Aiello, 2001a] Aiello, M. (2001a). Computing spatial similarity by games. In Esposito, F., editor, *AI*IA 2001: Advances in Artificial Intelligence*, LNAI, 2175, pages 99–110. Springer.
- [Aiello, 2001b] Aiello, M. (2001b). Review of: Parts and Places, R. Casati and A. Varzi, MIT press. *Journal of Logic, Language and Information*, 10(2):269–272.
- [Aiello, 2002a] Aiello, M. (2002a). Document image analysis via model checking. *AI*IA Notizie*. Special issue on AI techniques in Pattern Recognition. To appear.
- [Aiello, 2002b] Aiello, M. (2002b). A spatial similarity based on games: Theory and practice. *Journal of the Interest Group in Pure and Applied Logic*. To appear.
- [Aiello et al., 1999] Aiello, M., Areces, C., and de Rijke, M. (1999). Spatial reasoning for image retrieval. In Lambrix, P., Borgida, A., Lenzerini, M., Möller, R., and Patel-Schneider, P., editors, *Proc. of the International Workshop on Description Logics (DL'99)*, pages 23–27.
- [Aiello et al., 2000] Aiello, M., Monz, C., and Todoran, L. (2000). Combining linguistic and spatial information for document analysis. In Mariani, J. and Harman, D.,

- editors, *Proc. of RIAO'2000 Content-Based Multimedia Information Access*, pages 266–275. CID.
- [Aiello and Smeulders, 1999] Aiello, M. and Smeulders, A. (1999). Where do logic and computer vision meet? In Gerbrandy, J., Marx, M., de Rijke, M., and Venema, Y., editors, *JFAK. Essays Dedicated to Johan van Benthem on the Occasion of his 50th Birthday*, pages 1–11. Vossiuspers, AUP.
- [Aiello and van Benthem, 1999] Aiello, M. and van Benthem, J. (1999). Logical patterns in space. Technical Report PP-1999-24, ILLC, University of Amsterdam.
- [Aiello and van Benthem, 2002a] Aiello, M. and van Benthem, J. (2002a). *Logical Patterns in Space*. In [Barker-Plummer et al., 2002]. To appear.
- [Aiello and van Benthem, 2002b] Aiello, M. and van Benthem, J. (2002b). A modal walk through space. Technical Report PP-2001-23, ILLC, University of Amsterdam. Submitted to the *Journal of Applied non-Classical Logics*.
- [Aiello et al., 2001] Aiello, M., van Benthem, J., and Bezhanishvili, G. (2001). Reasoning about space: the modal way. Technical Report PP-2001-18, ILLC, University of Amsterdam. Submitted to the *Journal of Logic and Computation*.
- [Allen, 1983] Allen, J. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26:832–843.
- [Altamura et al., 2001] Altamura, O., Esposito, F., and Malerba, D. (2001). Transforming paper documents into XML format with WISDOM++. *International Journal of Document Analysis and Recognition*, 4(1):2–17.
- [Areces, 2000] Areces, C. (2000). *Logic Engineering: The Case of Description and Hybrid Logics*. PhD thesis, University of Amsterdam.
- [Arlazarov et al., 1970] Arlazarov, V., Dinic, E., Kronrod, M., and Faradjev, I. (1970). On economical finding of transitive closure of a graph. *Soviet Math. Dokl.*, 11:1270–1272.
- [Asher and Vieu, 1995] Asher, N. and Vieu, L. (1995). Toward a geometry of common sense: a semantics and a complete axiomatization of mereotopology. In *IJCAI95*, pages 846–852. International Joint Conference on Artificial Intelligence.
- [Baayen et al., 1995] Baayen, R., Piepenbrock, R., and Gulikers, L. (1995). The CELEX lexical database (release 2). Distributed by the Linguistic Data Consortium, University of Pennsylvania.
- [Baeza-Yates and Ribeiro-Neto, 1999] Baeza-Yates, R. and Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. Addison Wesley.

- [Balbiani, 1998] Balbiani, P. (1998). The modal multilogic of geometry. *Journal of Applied Non-Classical Logics*, 8:259–281.
- [Balbiani et al., 1997] Balbiani, P., Fariñas del Cerro, L., Tinchev, T., and Vakarelov, D. (1997). Modal logics for incidence geometries. *Journal of Logic and Computation*, 7:59–78.
- [Barker-Plummer et al., 2002] Barker-Plummer, D., Beaver, D., van Benthem, J., and Scotto di Luzio, P., editors (2002). *Logic Unleashed: Language, Diagrams and Computation*. CSLI, Stanford. To appear.
- [Barwise and Moss, 1996] Barwise, J. and Moss, L. (1996). *Vicious Circles*. CSLI.
- [Bennett, 1995] Bennett, B. (1995). Modal logics for qualitative spatial reasoning. *Bulletin of the IGPL*, 3:1–22.
- [van Benthem, 1976] van Benthem, J. (1976). *Modal Correspondence Theory*. PhD thesis, University of Amsterdam.
- [van Benthem, 1983a] van Benthem, J. (1983a). Correspondence theory. In Gabbay, D. and Guenther, F., editors, *Handbook of Philosophical Logic*, volume II, pages 167–247. Reidel Publishing Company.
- [van Benthem, 1983b] van Benthem, J. (1983b). *The Logic of Time*, volume 156 of *Synthese Library*. Reidel, Dordrecht. [Revised and expanded, Kluwer, 1991].
- [van Benthem, 1985] van Benthem, J. (1985). *Modal Logic and Classical Logic*. Bibliopolis, Napoli. The Humanities Press, Atlantic Heights (N.J.).
- [van Benthem, 1991a] van Benthem, J. (1991a). *Language in Action. Categories, Lambdas and Dynamic Logic*. North-Holland, Amsterdam.
- [van Benthem, 1991b] van Benthem, J. (1991b). Logic and the flow of information. In Prawitz, D., Skyrms, B., and Westerståhl, D., editors, *Proc. 9th International Congress of Logic, Methodology and Philosophy of Science*, pages 693–724, Uppsala. Elsevier.
- [van Benthem, 1996] van Benthem, J. (1996). *Exploring Logical Dynamics*, volume 156. CSLI Publications, Stanford — Cambridge University Press.
- [van Benthem, 1998] van Benthem, J. (1998). Information transfer across Chu Spaces, a chapter in *Dynamic Odds and Evens*. Technical Report ML-98-08, ILLC, University of Amsterdam.
- [van Benthem, 1999] van Benthem, J. (1999). *Logic and Games*. Lecture notes Amsterdam and Stanford, <http://www.illc.uva.nl/~johan/teaching>.

- [van Benthem, 2000] van Benthem, J. (2000). Logical structures in mathematical morphology. Available at <http://www.science.uva.nl/johan/MM-LL.ps>.
- [de Berg et al., 2000] de Berg, M., van Kreveld, M., Overmars, M., and Schwarzkopf, O. (2000). *Computational Geometry: Algorithms and Applications*. Springer-Verlag. Second edition.
- [Bezhanishvili and Gehrke, 2001] Bezhanishvili, G. and Gehrke, M. (2001). A new proof of completeness of S4 with respect to the real line. Submitted.
- [Blackburn et al., 2001] Blackburn, P., de Rijke, M., and Venema, Y. (2001). *Modal Logic*. Cambridge University Press.
- [Bloch, 2000] Bloch, I. (2000). Using mathematical morphology operators as modal operators for spatial reasoning. In *ECAI 2000, Workshop on Spatio-Temporal Reasoning*, pages 73–79.
- [Blumenthal, 1961] Blumenthal, L. (1961). *A Modern View of Geometry*. Dover.
- [Burgess, 1984] Burgess, J. (1984). Basic tense logic. In Gabbay, D. and Guenther, F., editors, *Handbook of Philosophical Logic*, volume II, chapter 2, pages 89–133. Reidel.
- [Cartwright, 1975] Cartwright, R. (1975). Scattered objects. In Lehrer, K., editor, *Analysis and Metaphysics*, pages 153–171. Dordrecht: Reidel. Reprinted in *Philosophical Essays*: 171–186.
- [Casati and Varzi, 1999] Casati, R. and Varzi, A. (1999). *Parts and Places*. MIT Press.
- [Cesarini et al., 1999] Cesarini, F., Francesconi, E., Gori, M., and Soda, G. (1999). A two level knowledge approach for understanding documents of a multi-class domain. In [Hull et al., 1999], pages 135–138.
- [Cesarini et al., 1998] Cesarini, F., Gori, M., Marinai, S., and Soda, G. (1998). INFORMys: A flexible invoice-like form-reader system. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 20(7):730–746.
- [Chang and Liu, 1984] Chang, S. and Liu, S. (1984). Picture indexing and abstraction techniques for pictorial databases. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 6(7):475–484.
- [Chellas, 1980] Chellas, B. (1980). *Modal Logic: An Introduction*. Cambridge University Press.
- [Clarke, 1981] Clarke, B. (1981). A calculus of individuals based on ‘connection’. *Notre Dame Journal of Formal Logic*, 23(3):204–218.

- [Clarke, 1985] Clarke, B. (1985). Individuals and points. *Notre Dame Journal of Formal Logic*, 26(1):61–75.
- [Cohn and Varzi, 1998] Cohn, A. G. and Varzi, A. (1998). Connection relations in mereotopology. In Prade, H., editor, *Proc. 13th European Conf. on AI (ECAI98)*, pages 150–154. John Wiley.
- [Dabrowski et al., 1996] Dabrowski, A., Moss, A., and Parikh, R. (1996). Topological reasoning and the logic of knowledge. *Annals of Pure and Applied Logic*, 78:73–110.
- [Dastani, 1998] Dastani, M. (1998). *Languages of Perception*. PhD thesis, University of Amsterdam.
- [Dastani et al., 1997] Dastani, M., Scha, R., and Indurkha, B. (1997). An algebraic method for solving proportional analogy problems involving sequential patterns. In *Mind II: Computational Models of Creative Cognition*, pages 1–15.
- [Del Bimbo, 1999] Del Bimbo, A. (1999). *Visual Information Retrieval*. Morgan Kaufmann Publishers.
- [Del Bimbo et al., 1995] Del Bimbo, A., Vicario, E., and Zingoni, D. (1995). Symbolic description and visual querying of image sequences using spatio-temporal logic. *IEEE Trans. Knowledge and Data Engineering*, 7(4):609–622.
- [Delaunay, 1934] Delaunay, B. (1934). Sur la sphère vide. *Otdelenie Matematicheskii i Estestvennyka Nauk*, 7:793–800.
- [Di Sciascio et al., 2000] Di Sciascio, E., Donini, F., and Mongiello, M. (2000). Semantic indexing for image retrieval using description logics. In Laurini, R., editor, *Advances in Visual Information Systems (VISUAL 2000)*, volume 1929 of *Lecture Notes in Computer Science*. Springer.
- [Doets, 1987] Doets, K. (1987). *Completeness and Definability. Applications of the Ehrenfeucht Game in Second-Order and Intensional Logic*. PhD thesis, ILLC, University of Amsterdam.
- [Doets, 1996] Doets, K. (1996). *Basic Model Theory*. CSLI Publications, Stanford.
- [Dovier et al., 2001] Dovier, A., Piazza, C., and Policriti, A. (2001). A fast bisimulation algorithm. In Berry, G., Comon, H., and Finkel, A., editors, *Proc. of Computer Aided Verification (CAV'01)*, number 2102 in LNCS, pages 79–90. Springer-Verlag.
- [Egenhofer, 1991] Egenhofer, M. (1991). Reasoning about binary topological relations. In Gnther, O. and Schek, H., editors, *Advances in Spatial Databases*, volume 525 of *Lecture Notes in Computer Science*.

- [Egenhofer, 1997] Egenhofer, M. (1997). Query processing in spatial-query-by-sketch. *Journal of Visual Languages and Computing*, 8(4):403–424.
- [Egenhofer and Franzosa, 1991] Egenhofer, M. and Franzosa, R. (1991). Point set topological relations. *Journal of Geographical Information Systems*, 5:161–174.
- [Engelking, 1989] Engelking, R. (1989). *General Topology*. Heldermann Verlag.
- [Esposito et al., 2000] Esposito, F., Malerba, D., and Lisi, F. (2000). Machine learning for intelligent processing of printed documents. *Journal of Intelligent Information Systems*, 14(2/3):175–198.
- [Florack, 1997] Florack, L. (1997). *Image Structure*. Kluwer Academic Publishers.
- [Foley et al., 1990] Foley, J., van Dam, A., Feiner, S., and Hughes, J. (1990). *Computer Graphics: Principles and Practice*. Addison-Wesley.
- [Frakes and Baeza-Yates, 1992] Frakes, W. and Baeza-Yates, R., editors (1992). *Information Retrieval, Data Structures and Algorithms*. Prentice Hall.
- [Gabbay, 1981] Gabbay, D. M. (1981). An irreflexivity lemma with applications to axiomatizations of conditions on tense frames. In Mönnich, U., editor, *Aspects of Philosophical Logic: Some Logical Forays into Central Notions of Linguistics and Philosophy*, pages 67–89. Reidel.
- [Gabelaia, 2002] Gabelaia, D. (2002). Diamonds of space. Master’s thesis, ILLC, University of Amsterdam.
- [Gärdenfors, 2000] Gärdenfors, P. (2000). *Conceptual Spaces*. MIT Press.
- [Georgatos, 1993] Georgatos, K. (1993). *Modal Logics for Topological Spaces*. PhD thesis, City University of New York.
- [Girard, 1987] Girard, J.-Y. (1987). Linear logic. *Theoretical Computer Science*, 50:1–102.
- [Goldblatt, 1987] Goldblatt, R. (1987). *Orthogonality and Spacetime Geometry*. Springer-Verlag.
- [Goossens et al., 1994] Goossens, M., F., M., and Samarin, A. (1994). *The L^AT_EX Companion*. Addison-Wesley.
- [Goranko and Passy, 1992] Goranko, V. and Passy, S. (1992). Using the universal modality: gains and questions. *Journal of Logic and Computation*, 2:5–30.
- [Gurr, 1998] Gurr, C. (1998). Theories of visual and diagrammatic reasoning: Foundational issues. In *Formalizing Reasoning with Diagrammatic and Visual Representations*, Orlando, Florida. AAAI Fall Symposium.

- [Halpern and Shoham, 1991] Halpern, J. and Shoham, Y. (1991). A propositional modal logic of time intervals. *Journal of the ACM*, 38(4):935–962.
- [Hammer, 1995] Hammer, E. (1995). *Logic and Visual Information*. CSLI and FoLLi, Stanford.
- [Harel, 1983] Harel, D. (1983). Recurring dominoes: Making the highly undecidable highly understandable. In *Conference on Foundations of Computer Science*, volume LNCS 158, pages 177–194. Springer.
- [Heijmans, 1994] Heijmans, H. J. A. M. (1994). *Morphological Image Operators*. Academic Press, Boston.
- [Helly, 1923] Helly, E. (1923). Uber mengen konvexer k orper mit gemeinschaftlichen punkten. *Jahresber. Deutsch. Math.- Verein*, 32:175–176.
- [Hersh et al., 1994] Hersh, W., Buckley, C., Leone, T., and Hickam, D. (1994). OHSUMED: an interactive retrieval evaluation and new large test collection for research. In *Proc. of ACM-SIGIR conference on Research and development in information retrieval*, pages 192–201. Springer.
- [Herskovits, 1997] Herskovits, A. (1997). Language, spatial cognition, and vision. In Stock, O., editor, *Spatial and Temporal Reasoning*, pages 155–202. Kluwer Academic Publishers.
- [Hull et al., 1999] Hull, J., Lee, S., and Tombre, K., editors (1999). ICDAR, IEEE.
- [Johnstone, 1977] Johnstone, P. (1977). *Topos Theory*. Number 10 in Mathematical Society Monographs. Academic Press.
- [Johnstone, 1982] Johnstone, P. (1982). *Stone Spaces*. Number 3 in Cambridge Studies in Advanced Mathematics. Cambridge University Press.
- [Kamp, 1968] Kamp, J. (1968). *Tense logic and the theory of linear order*. PhD thesis, University of California, Los Angeles.
- [Kerdiles, 2001] Kerdiles, G. (2001). *Saying It with Pictures: a logical landscape of coneptual graphs*. PhD thesis, Univeristy of Amsterdam.
- [Klink et al., 2000] Klink, S., Dengel, A., and Kieninger, T. (2000). Document structure analysis based on layout and textual features. In Murshed, N. and Amin, A., editors, *Proc. of International Workshop on Document Analysis Systems, DAS2000*, pages 99–111. IAPR.
- [Klink and Kieninger, 2001] Klink, S. and Kieninger, T. (2001). Rule-based document structure understanding with a fuzzy combination of layout and textual features. *International Journal of Document Analysis and Recognition*, 4(1):18–26.

- [Knuth, 1984] Knuth, D. (1984). *The T_EXbook*. Addison–Wesley.
- [Knuth, 1968] Knuth, D. E. (1968). *The Art of Computer Programming*. Addison Wesley. Third edition 1998.
- [Knuth and Szwarcfiter, 1974] Knuth, D. E. and Szwarcfiter, J. L. (1974). A structured program to generate all topological sorting arrangements. *Information Processing Letters*, 2(6):153–157.
- [Kurtonina, 1995] Kurtonina, N. (1995). *Frames and Labels. A Modal Analysis of Categorical Inference*. PhD thesis, ILLC, Amsterdam.
- [Kurtonina and de Rijke, 1997] Kurtonina, N. and de Rijke, M. (1997). Bisimulations for temporal logic. *Journal of Logic, Language and Information*, 6:403–425.
- [Lee et al., 2000] Lee, K., Choy, Y., and Cho, S. (2000). Geometric structure analysis of document images: A knowledge approach. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22(11):1224–1240.
- [Lemon, 2002] Lemon, O. (2002). *Comparing the Efficacy of Visual Languages*. In [Barker-Plummer et al., 2002]. To appear.
- [Lemon and Pratt, 1998] Lemon, O. and Pratt, I. (1998). Complete logics for QSR: A guide to plane meretopology. *Journal of Visual Languages and Computing*, 9(1):5–21.
- [Lesniewski, 1983] Lesniewski, S. (1983). The foundation of mathematics. *Topoi*, 2:7–52. Translation of the paper of 1931.
- [Lewis, 1973] Lewis, D. (1973). *Counterfactuals*. Harvard University Press.
- [Li and Ng, 1999] Li, X. and Ng, P. (1999). A document classification an extraction system with learning ability. In [Hull et al., 1999], pages 197–200.
- [Marx and Venema, 1997] Marx, M. and Venema, Y. (1997). *Multi Dimensional Modal Logic*. Kluwer.
- [Matheron, 1967] Matheron, G. (1967). *Éléments pur une theorie des milieux poreux*. Masson.
- [McKinsey and Tarski, 1944] McKinsey, J. and Tarski, A. (1944). The algebra of topology. *Annals of Mathematics*, 45:141–191.
- [Mints, 1998] Mints, G. (1998). A completeness proof for propositional S4 in cantor space. In E. Orłowska, editor, *Logic at work : Essays dedicated to the memory of Helena Rasiowa*. Physica-Verlag, Heidelberg.

- [Munro, 1971] Munro, I. (1971). Efficient determination of the transitive closure of a directed graph. *Information Processing Letters*, 1(2):56–58.
- [Nagy, 2000] Nagy, G. (2000). Twenty Years of Document Image Analysis in PAMI. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(1):38–62.
- [Niyogi and Srihari, 1996] Niyogi, D. and Srihari, S. (1996). Using domain knowledge to derive logical structure of documents. In Alferov, Z., Gulyaev, I., and Pape, D., editors, *Proc. Document Recognition and Retrieval III.*, pages 114–125. SPIE.
- [Nute, 1983] Nute, D. (1983). Conditional logic. In Gabbay, D. and Guenther, F., editors, *Handbook of Philosophical Logic*, volume IV, pages 387–439. Reidel Publishing Company.
- [Park, 1981] Park, D. (1981). Concurrency and automata on infinite sequences. In *Proc. of the 5th GI Conference*, pages 167–183, Berlin. Springer Verlag.
- [Petraakis et al., 2001] Petraakis, E., Faloutsos, C., and King-Ip, L. (2001). ImageMap: An Image Indexing Method Based on Spatial Similarity. *IEEE Trans. on Knowledge and Data Engineering*. To appear.
- [von Plato, 1995] von Plato, J. (1995). The axioms of constructive geometry. *Annals of Pure and Applied Logic*, 76(2):169–200.
- [Pratt and Lemon, 1997] Pratt, I. and Lemon, O. (1997). Ontologies for Plane, Polygonal Mereotopology. *Notre Dame Journal of Formal Logic*, 38(2):225–245.
- [Pratt and Schoop, 1998] Pratt, I. and Schoop, D. (1998). A complete axiom system for polygonal mereotopology. *Journal of Philosophical Logic*, 27:621–658.
- [Preparata and Shamos, 1985] Preparata, F. and Shamos, M. (1985). *Computational Geometry: An Introduction*. Springer-Verlag.
- [Prior, 1967] Prior, A. (1967). *Past, Present and Future*. Oxford University Press.
- [Randell et al., 1992] Randell, D., Cui, Z., and Cohn, A. G. (1992). A Spatial Logic Based on Regions and Connection. In *Proc. of Int. Conf. on Principles of Knowledge Representation and Reasoning (KR'92)*, pages 165–176. San Mateo.
- [Randell et al., 2001] Randell, D., Witkowski, M., and Shanahan, M. (2001). From Images to Bodies: Modelling and Exploiting Occlusion and Motion Parallax. In *Proc. of Int. Joint Conference on Artificial Intelligence (IJCAI-01)*.
- [Rasiowa and Sikorski, 1963] Rasiowa, H. and Sikorski, R. (1963). *The Mathematics of Metamathematics*. Panstwowe Wydawnictwo naukowe.
- [Reynold, 1979] Reynold, L. (1979). *The Thames and Hudson Manual of Typography*. Thames & Hudson.

- [de Rijke, 1993] de Rijke, M. (1993). *Extended Modal Logic*. PhD thesis, ILLC, University of Amsterdam.
- [Robinson, 1971] Robinson, R. (1971). Undecidability and nonperiodicity for tilings of the plane. *Inventiones Math*, 12:177–209.
- [Rosenfeld, 1997] Rosenfeld, A. (1997). Eyes for Computers: How HAL Could “See”. In Stork, D., editor, *HAL’s Legacy*, pages 210–235. MIT Press.
- [Russ et al., 1996] Russ, T., MacGregor, R., Salemi, B., Price, K., and Nevatia, R. (1996). VEIL: Combining semantic knowledge with image understanding. In *ARPA Image Understanding Workshop*.
- [Sambin, 1987] Sambin, G. (1987). Intuitionistic formal spaces - a first communication. In Skordev, D., editor, *Mathematical logic and its applications*, pages 187–204. Plenum Press.
- [Sauvola and Kauniskangas, 2000] Sauvola, J. and Kauniskangas, H. (2000). Medi-aTeam Document Database II. CD-ROM collection of document images, University of Oulu, Finland. <http://www.mediateam.oulu.fi/MTDB/index.html>.
- [Schiirmann, 1997] Schiirmann, J., editor (1997). ICDAR, IEEE.
- [Segerberg, 1970] Segerberg, K. (1970). Modal Logics with linear alternative relations. *Theoria*, 36:301–322.
- [Segerberg, 1976] Segerberg, K. (1976). “Somewhere else” and “Some other time”. In *Wright and Wrong*, pages 61–64.
- [Serra, 1982] Serra, J. (1982). *Image Analysis and Mathematical Morphology*. Academic Press.
- [Shanahan, 1995] Shanahan, M. (1995). Default Reasoning about Spatial Occupancy. *Artificial Intelligence*, 74(1):147–163.
- [Shehtman, 1983] Shehtman, V. (1983). Modal Logics of Domains on the Real Plane. *Studia Logica*, 42:63–80.
- [Shehtman, 1999] Shehtman, V. (1999). “Everywhere” and “Here”. *Journal of Applied Non-Classical Logics*, 9(2-3):369–379.
- [Singer and Thorpe, 1967] Singer, I. M. and Thorpe, J. A. (1967). *Lecture Notes on Elementary Topology and Geometry*. Springer-Verlag.
- [Singh et al., 1999] Singh, R., Lahoti, A., and Mukerjee, A. (1999). Interval-algebra based block layout analysis and document template generation. In “*Workshop on Document Layout Interpretation and its Applications*”. Event connected with [Hull et al., 1999].

- [Smeulders et al., 2000] Smeulders, A., Worring, M., Santini, S., Gupta, A., and Jain, R. (2000). Content based image retrieval at the end of the early years. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(12).
- [Stebletsova, 2000] Stebletsova, V. (2000). *Algebras, relations and Geometries*. PhD thesis, University of Utrecht.
- [Tagare et al., 1995] Tagare, H., Vos, F., Jaffe, C., and Duncan, J. (1995). Arrangement: A spatial relation between part for evaluating similarity of tomographic section. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 17(9):880–893.
- [Tarski, 1938] Tarski, A. (1938). Der Aussagenkalkül und die Topologie. *Fund. Math.*, 31:103–134.
- [Tarski, 1959] Tarski, A. (1959). What is Elementary Geometry? In L. Henkin and P. Suppes and A. Tarski, editor, *The Axiomatic Method, with Special Reference to Geometry ad Physics*, pages 16–29. North-Holland.
- [Toda, 1990] Toda, S. (1990). On the complexity of topological sorting. *Information Processing Letters*, 35(5):229–233.
- [Todoran et al., 2001a] Todoran, L., Aiello, M., Monz, C., and Worring, M. (2001a). Document Understanding for a Broad Class of Documents. Technical Report ISIS-TR-2001-15, University of Amsterdam. Submitted to the *International Journal on Document Analysis and Recognition*.
- [Todoran et al., 2001b] Todoran, L., Aiello, M., Monz, C., and Worring, M. (2001b). Logical structure detection for heterogeneous document classes. In *Document Recognition and Retrieval VIII*, pages 99–110. SPIE.
- [Troelstra, 1992] Troelstra, A. S. (1992). *Lectures on Linear Logic*. CSLI.
- [Tsujimoto and Asada, 1992] Tsujimoto, S. and Asada, H. (1992). Major components of a complete text reading system. *Proc. of the IEEE*, 80(7):1133–1149.
- [Vardi, 1997] Vardi, M. (1997). Why is modal logic so robustly decidable. In Immerman, N. and Kolaitis, P., editors, *Descriptive Complexity and Finite Models*. American Mathematical Society.
- [Veltman, 1985] Veltman, F. (1985). *Logics for conditionals*. PhD thesis, University of Amsterdam.
- [Venema, 1991] Venema, Y. (1991). A modal logic for chopping intervals. *Journal of Logic and Computation*, 1(4):453–476.
- [Venema, 1992] Venema, Y. (1992). *Many-Dimensional Modal Logic*. PhD thesis, ILLC, University of Amsterdam.

- [Venema, 1996] Venema, Y. (1996). A crash course in arrow logic. In Marx, M., Masuch, M., and Pólos, L., editors, *Arrow Logic and Multimodal Logic*. CSLI.
- [Venema, 1999] Venema, Y. (1999). Points, Lines and Diamonds: a Two-Sorted Modal Logic for Projective Planes. *Journal of Logic and Computation*, 9(5):601–621.
- [Vickers, 1989] Vickers, S. (1989). *Topology via Logic*. Cambridge University Press.
- [Vickers, 1994] Vickers, S. (1994). Locales are not pointless. In Hankin, Mackie, and Nagarajan, editors, *Theory and Formal Methods of Computing*, pages 199–216. Imperial College Press.
- [Voronoi, 1908] Voronoi, G. (1908). Nouvelles applications des paramtres continus la thorie des forme quadratiques. *J. Reine Angew. Math.*, 134:198–287.
- [Walischewski, 1997] Walischewski, H. (1997). Automatic knowledge acquisition for spatial document interpretation. In [Schiirmann, 1997], pages 243–247.
- [Warshall, 1962] Warshall, S. (1962). A Theorem on Boolean Matrices. *Journal of the ACM*, 9(1):11–12.
- [Whitehead, 1929] Whitehead, A. (1929). *Process and Reality*. The MacMillan Company, New York.
- [Winter and Zwarts, 1997] Winter, Y. and Zwarts, J. (1997). A Semantic Characterization of Locative PPs. In *SALT7*.
- [Witten et al., 1999] Witten, I., Moffat, A., and Bell, T. (1999). *Managing Gigabytes: Compressing and Indexing Documents and Images*. Morgan Kaufmann Publishing, 2nd edition.
- [Worring and Smeulders, 1999] Worring, M. and Smeulders, A. (1999). Content based internet access to paper documents. *International Journal of Document Analysis and Recognition*, 1(4):209–220.

- \rightleftharpoons , *see* topological bisimulation
- $D\varphi$, *see* difference operator
- $E\varphi$, *see* modal operator
- K**, *see* modal logic
- $\text{\texttt{E}\text{T}\text{T}\text{X}}$, *see* document encoding rule
- $M_{\mathbb{R}^2}$, *see* Real plane
- \mathbb{R}^2 , *see* Real plane
- $\mathcal{S}(\mathbb{R})$, *see* serial set
- S4_w**, *see* modal logic
- $TG(X, X', n, x_1, x'_1)$, *see* topological game
- T_2 , *see* infinite binary tree
- T_i , *see* separation axioms
- tmd*, *see* topological distance
- topax*, *see* selective unraveling
- $TG(X, X', n)$, *see* topological game
- $U\varphi$, *see* modal operator
- \vec{r} , *see* topo-vector
- (X, O) , *see* topological space

- 2-fork frame, 44
- 2D strings, 101

- ACM, 129
 - Communications, 140
 - transaction class file, 129, 135
- admissible transition, 140
- affine bisimulation, 72
 - \rightleftharpoons , 72
 - reduction, 73
- affine geometry, 81
- Alexandroff space, 24, 28, 165
- Allen relations, 133
- Ansel Adams, 1
- arrow bisimulation, 97
- arrow logics, 4
- arrow model, 96
- Axiom
 - (4_@), 61
 - (4), 10
 - (Con), 54
 - (Dual_@), 61
 - (Dual.), 10
 - (Intersection), 61
 - (K_@), 61
 - (K), 10
 - (Label), 61
 - (N), 10
 - (R), 10
 - (Scope), 61
 - (T_@), 61
 - (T), 10
 - distribution, 28
 - Grzegorzcyk, 43
 - Pasch, 75
- axiomatization, 3
 - S4**, 10
 - S5**, 21, 54
 - elementary geometry, 68

- bag, *see* multi-set
- base, 155
- basic sets, 26
- betweenness, 3, 67, 138
 - Manhattan, 139
- bidimensional Allen relations, 177
- Borel Hierarchy, 48
- boundaries
 - inferred vs. detected, 153
- boundary, 155
- brittleness, 4

- canonical topological space, 26
- Cantor space, 3, 28, 156, 165, 167, 168
 - 2^ω , 33
- cantorization, 32
- categorical grammar, 95
- chequered, 49
- chop-stick, 11
- Chu morphism, 13
- closure, 155
- cluster, 30, 164–168
 - cycling, 33
 - initial, 30
 - proper, 30
 - representative, 33
 - simple, 30
 - stable, 33
- collinearity, 69
- column-wise, 136, 146, 149, 175
- compact, 156
- comparative nearness, *see* relative nearness
- conceptual spaces, 67
- conditional logic, 5, 89, 90
 - nested spheres, 90
- connected space, 156
- connected topological space, 13
- constructive geometry, 91
- continuity, 157
- continuous map, 13
- convex, 42, 48, 49
 - rectangular, 49

- correspondence theory, 25
- crack, 113
- Currying laws, 94

- decidable, 30
- Delaunay, 84–86
- dense, 156
- description logics, 102
- difference logic, 98
- difference operator
 - $D\varphi$, 53
- directed transitive cyclic graph, 159
- discrete topology, 156
- document encoding rule, 127–133, 135, 139–144, 146, 149, 174
 - L^AT_EX, 129, 131, 135
 - abstract, 133
 - formal, 129
 - generic, 129
 - informal, 129
 - SGML, 132
 - specific, 129
 - WYSIWYG, 131
- document image analysis, 4, 125–150
- document object, 125–129, 132–137, 139–143, 146–149
- document understanding, *see* document image analysis
- Duplicator, 15

- Eclipse, 142, 171
 - finite domain library, 171
- Ehrenfeucht-Fraïssé game, 5, 14, 107, 123, 153, 163
- elementary equivalence, 123
- elementary geometry, 3
- equidistance, 3, 82
- existential modal formula, 13

- finite model property, 29
- finite rank, 39
- float environment, 130
- flow formulas, 13
- formula

- flow, 13
- fractal structure, 168
- frame correspondence, 51
- frame correspondence techniques, 21
- frames, 9
- frontier, 155
- generated submodel, 30
- geometry
 - affine, 81
 - elementary, 68, 90
 - metric, 82
- Geometry of Compasses, 83
- HAL, 125, 150
- Hausdorff's residue, 39
- Helly's Theorem, 78
- Henkin construction, 26
- homeomorphism, 13, 33, 65, 157
- homogeneous coordinates, 79
- homotopy, 158
 - type, 158
- HTML, 127, 132
- image retrieval, 4, 101–123
- image similarity, 102
- indiscrete topology, 156
- inference rule
 - Modus Ponens, 10, 94, 98
 - Monotonicity, 10
 - Necessitation, 10
- infinitary modal logic, 12
- infinite binary tree, 32
 - T_2 , 32
- infinite rank, 39
- infinite topological game, 15
- interior, 155
- IRIS, 102, 119, 120, 123, 163, 169
- isosceles topo-distance, *see* topological distance
- Java, 119, 163
 - applet, 163
 - Swing, 119
- JEPD, 134
- jointly exhaustive and pairwise disjoint, *see* JEPD
- Kleene iteration, 77
- Kripke semantics, *see* possible worlds semantics
- Kripke topology, 28
- labeling, 123
- Lambek calculus, 93
- layout structure, 125, 126
- Lewis axioms, 90
- limit point, 155
- Lindenbaum Lemma, 28
- linear algebra, 91–100
- linear logic
 - multiplicative, 93
- locality, 25
- logical finiteness, 19
- logical labeling, 125, 126, 128, 133, 141, 146
- logical structure, 126–128, 149
- Manhattan betweenness, 139
- Mascheroni, 83, 84
- mathematical morphology, 4, 7, 91–100, 152
 - closing, 93
 - dilation, 93
 - erosion, 93
 - opening, 93
- maximally consistent, 26
- Media Team Data Base, *see* MTDB
- mereology, 7
- mereotopology, 101
 - connection, 103–105
 - parthood, 103, 104
- metric geometry, 82
- metric separable space, 23
- metric space, 156
- minimal normal modal logic, 28
- Minkowski operations, 4
- modal formula

- existential, 13
- modal logic, 5
 - K**, 7
 - S4_u**, 3, 5, 54–59, 61, 62, 102–107, 109–111, 113–115, 122, 123, 151–153, 194
 - S4**, 3, 7–11, 14, 19–21, 23, 25–32, 35, 38–40, 42–45, 48, 50, 51, 53–55, 58, 60–62, 109, 111, 114, 151–153, 194
 - S4_u**
 - normal form, 54, 57
 - arrow logics, 4, 82, 96–100
 - extended, 53
 - geometry, 67
 - hybrid reference, 60
 - infinitary, 12
 - possible worlds semantics, 9
 - sorted, 20
 - topological semantics, 8
 - universal, 54
- modal operator
 - $E\varphi$, 54
 - $U\varphi$, 54
 - difference, 53
 - dyadic, 3
 - existential, 69
 - hybrid, 60
 - universal, 3
 - versatile, 72, 87, 98
- modal rank, 18
 - multi-modal rank, 56
- model class
 - across vs. within, 153
- model comparison game, 14
 - Duplicator, 15
 - examples, 17
 - Spoiler, 15
- model comparison games, 101–123
- MTDB, 142–146, 149
- multi-set, 94
- neighborhood semantics, 26
- nominals, 99
- normal form, 54, 109
- one-step convexity, 67, 76
- ontology
 - intervals vs. instants, 152
 - regions vs. points, 152
- open, 155
- operator
 - Minkowski, 4, 92
- order preserving, 24
- orthonormal, 79
- p -morphism, 24
- Pappus property, 68
- path
 - real, 36
- Perl, 171
- plane
 - Pappian, 68
 - projective, 79
- polygonal ontology, 111
- possible worlds semantics, 9, 14, 23
- precision, 143
- projective plane, 79
- quantification
 - second-order, 25
- quantifier
 - two sorted, 52
- quasi-order, 23
- quasi-ordered set, 24
- RCC, 3, 82, 103–106, 121, 134, 135, 138
- reading order, 125–129, 136, 140, 142–150
 - spatially admissible, 140
- real line, 3, 165
- Real plane
 - \mathbb{R}^2 , 8
 - $M_{\mathbb{R}^2}$, 111, 114, 115, 117, 118
 - polygons, 111
- recall, 143

- region connection calculus, *see* RCC
- relative nearness, 82, 86
 - universal first-order theory, 83
- residuation law, 93
- resource logic, 93
- row-wise, 136, 137, 142, 144, 146, 148, 149, 175

- second-order logic, 65
- second-order quantification, 25
- selective unraveling, 32, 36
 - topax, 163
- separation axioms, 25, 156
 - T_i , 65
- serial, 42, 49
- serial set, 3
 - $S(\mathbb{R})$, 42
- SGML, 132
- similarity, 119
 - binary spatial, 119
 - spatial intricacy, 119
 - textual description, 119
- Since and Until, 3, 62
- sorts
 - first-order model, 25
 - points and lines, 67, 81
 - points and segments, 81
 - points and sets, 20, 65
- space, 7
 - Alexandroff, 24, 28
 - Cantor, 3, 156
 - discrete, 51
 - metric, 156
 - metric separable, 23
 - metric, compact, and dense-in-itself, 28
 - real line, 3
 - vector, 4, 97
- SpaRe, 127, 129, 130, 142–147, 149, 153, 163, 171
- spatial [bidimensional Allen] model, 133
- spatial coherence, 4
- spatial reasoning
 - avoid, 153
- spatially admissible reading order, 140
- specialization order, 24
- spike, 113
- Spoiler, 15
 - shortest possible win, 107
- spoon, 11
 - playing, 17
 - reduction, 17
 - two handles, 14
- spoons and forks, 8, 57
- Stone construction, 28
- Stone topology, 28

- tense logic, 5, 64
 - Future and Past, 76
- theory and practice, 5, 152
 - vicious circle, 5
- thick boundary, 127, 137, 138, 142, 147, 177
- thick boundary interpretation, *see* thick boundary
- thick boundary relations, *see* thick boundary
- threshold, 137, 138, 142, 144–148, 172
- tiling, 99
- Topo-approach, 151
- topo-approach, 151
- topo-bisimilar reduction, 14
- topo-bisimulation, *see* topological bisimulation
- topo-distance, *see* topological distance
- topo-frame, *see* topological frame
- topo-game, *see* topological game
- topo-model, *see* topological model
- topo-reduction, *see* topological reduction
- topo-vector, 114–119
 - \vec{r} , 114
 - size, 114
- topological bisimulation, 107, 166
 - \Leftrightarrow , 11, 55, 61
 - invariance, 25

- invariant, 12
- simulation, 11
- topological bisimulation for $\mathbf{S4}_u$, 55, 61
- topological bisimulation for $\mathbf{S4}$, 11
- total, 11
- topological distance, 107
 - tmd , 107
- topological frames, 13
- topological game, 14, 55, 107
 - $TG(X, X', n)$, 56
 - $TG(X, X', n, x_1, x'_1)$, 15
 - adequacy, 18
 - determined, 18
 - Duplicator, 55, 107
 - global round, 55
 - infinite, 15, 19, 57
 - local round, 55
 - Spoiler, 55, 107
 - strategies, 19, 57
 - winning strategy, 15
- topological interpretation, *see* topological semantics
- topological model, 9
- topological reduction, 14, 115
- topological semantics, 14, 23–51, 164
 - bisimulation, 11, 55, 61, 72, 97
 - normal, 28
 - topological semantics, 8
 - topological semantics of $\mathbf{S4}_u$, 54
 - topological semantics of $\mathbf{S4}_@$, 60
 - topological semantics of $\mathbf{S4}$, 9
- topological simulation, 11, 13
- topological sorting, 159
- topological space, 7, 9, 13, 155–158
 - (X, O) , 9
 - T_0 , 24, 25, 31
 - T_D , 31
 - base, 155
 - boundary, 62, 155
 - canonical, 26
 - Cantor, 156
 - closure, 155
 - compact, 156
 - connected, 13, 30, 58, 156
 - continuity, 157
 - dense, 156
 - discrete, 156
 - finite, 24
 - frontier, 155
 - Hausdorff, 66
 - homeomorphism, 157
 - homotopy, 158
 - indiscrete, 156
 - interior, 155
 - limit point, 155
 - metric, 156
 - open, 155
 - regular region, 57
 - well-connected, 30
- topology, *see* topological space
- total topological bisimulation, 11, 55
- triangle inequalities, 83, 88
- triangle inequality, 98
- truth lemma, 27
- universal modal operator, 3
- universal operator, 54
- upward closed, 23
- valuation
 - serial, 43
- vector space, 4, 97
- versatile modal operator, 72, 87, 98
- Voronoi, 84–86, 101, 149
- w.s., *see* winning strategy
- what you see is what you get, *see* WYSIWYG
- winning strategy, 15, 56
- WYSIWYG, 131, 132
- XML, 127, 132

SAMENVATTING

This is an abstract of the thesis in Dutch.²

Ruimtelijke structuren zijn essentieel voor perceptie en cognitie. Een groot deel van onze dagelijkse informatieuitwisselingen betreft de vraag wat er aan de hand is en waar. Daarnaast vormen ruimtelijke representaties een goede bron voor geometrische intuïties die een verklaring vormen voor algemene cognitieve taken. Hoe representeren we objecten die in de ruimte zijn gelocaliseerd? Hoe kunnen we over dit soort objecten redeneren? Bijvoorbeeld bij het opdekken van een tafel, wat zijn vanuit ruimtelijk oogpunt beschouwd de basis eigenschappen van, zeg, een lepel in relatie tot de rest van het bestek en de rest van de ruimte? Een ander basisaspect van perceptie is dat wij in staat zijn verschillende visuele scènes te vergelijken en eenzelfde object in deze verschillende scènes te identificeren. Zo kunnen we vaststellen welke feestelijk gedekte tafels ‘hetzelfde’ zijn. Logica verschaft middelen voor deze taak.

We moeten voorzichtig zijn als we het begrip ruimte in een logische theorie vatten en er vervolgens logische hulpmiddelen op loslaten. We kunnen namelijk niet verwachten dat de werkelijke ruimte in al zijn verscheidenheid zonder meer gevat is in onze formele theorie van deze ruimte. Zo zal onze theorie bepaalde natuurlijke, ruimtelijke aspecten niet kunnen behandelen, terwijl daarentegen sommige niet-natuurlijke, ruimtelijke fenomenen een rol zullen spelen. We zijn er echter ook niet op uit een volledige representatie van de ruimte te geven, maar we proberen de meest in het oog springende ruimtelijke fenomenen uit te drukken.

Onze bijdrage met deze dissertatie is tweeledig. In de eerste plaats onderzoeken wij nieuwe en bestaande ruimtelijke formalismen met het expliciete doel om logica's te identificeren met een redelijke uitdrukingskracht die tegelijkertijd mooie, meta-logische eigenschappen bezitten. In de tweede plaats onderzoeken we de haalbaarheid

²The samenvatting is mandatory for all thesis defended in the Netherlands which are written in English. Many thanks to Eva Hoogland for the translation.

van praktische toepassingen van dit soort kwalitatieve, ruimtelijke logica's. Hiertoe bestuderen we twee symbolische benaderingen van patroonherkenning.

Dit proefschrift bestaat uit zeven technische hoofdstukken, een introductie, een afsluitend hoofdstuk en drie appendices. De hoofdstukken 2 tot 5 vormen de theoretische kern van de dissertatie, de hoofdstukken 6 en 7 vormen de praktische component.

De eerste twee hoofdstukken geven de grenzen van onze benadering aan: Hoofdstuk 2 geeft aan wat we wel en niet kunnen uitdrukken, Hoofdstuk 3 behandelt welke axioma's we kunnen toestaan. Daarna analyseren we twee soorten uitbreidingen van deze benadering: logische (Hoofdstuk 4) en axiomatische uitbreidingen (Hoofdstuk 5).

In Hoofdstuk 2 brengen we de topologische interpretatie van modale logica's opnieuw tot leven door deze op te vatten als een algemene taal voor ruimtelijke patronen. Zo definiëren we een notie van bisimulatie voor topologische modellen aan de hand waarvan verschillende visuele scènes kunnen worden vergeleken. De resulterende notie van gelijkheid verfijnen we later door Ehrenfeucht-Fraïssé spelen te introduceren die op ruimtelijke structuren kunnen worden gespeeld.

In Hoofdstuk 3 onderzoeken we de topologische interpretatie van modale logica in moderne termen, waarbij we gebruik maken van de notie van bisimulatie die we zojuist hebben geïntroduceerd. We beschouwen modale logica's met een interessante topologische inhoud en presenteren ondermeer een nieuw bewijs van de volledigheid van $S4$ ten opzichte van de reële getallen (eerder bewezen door McKinsey en Tarski) en ook een volledigheidsbewijs van de logica van eindige verenigingen van convexe verzamelingen reële getallen.

In het volgende hoofdstuk beschouwen we logische uitbreidingen van de topologische modale benadering van ruimte. We introduceren universele en hybride modaliteiten en onderzoeken in hoeverre deze bijdragen aan de uitdrukkingskracht. Ook bekijken we een ruimtelijke versie van de tijdslogica van Since en Until. Een beknopte vergelijking met hogere-orde formalismen geeft een algemeen beeld van (uitgebreide) modale, ruimtelijke, logica's.

We vervolgen onze modale ruimtelijke onderzoeken in Hoofdstuk 5 door over te stappen op affine en metrische geometrieën, en op vectoralgebra. Dit levert een nieuwe onderverdeling in ruimtelijke patronen die analogieën suggereren tussen vernoemde wiskundige theorieën in termen van modale logica's, conditionele logica's en tijdslogica's. We onderzoeken de uitdrukkingskracht in termen van het ontwerp van de taal, bisimulaties en correspondentieverschijnselen. We leren verscheidene overeenkomsten tussen de verschillende gebieden, kennen, en stuiten op open vragen.

In Hoofdstuk 6 kijken we met andere ogen naar model-vergelijkende spelen ten einde een maat te ontwikkelen waarmee de gelijkheid van beelden bepaald kan worden. Dit soort spelen kunnen namelijk niet alleen gebruikt worden om te beslissen of twee gegeven modellen gelijk zijn, maar ook om een maat op te stellen die de verschillen binnen een *klasse* van modellen bepaalt. We laten zien hoe dit mogelijk is voor het geval van de ruimtelijke modale logica $S4_u$. Deze benadering geeft ons dus een maat voor ruimtelijke gelijkheid die gebaseerd is op topologische, model-vergelijkende spelen.

Als een toepassing geven we een algoritme dat effectief de gelijkheidsmaat berekent voor een klasse van modellen die volop gebruikt wordt in de informatica: polygonen van het reële vlak. Aan het eind van dit hoofdstuk geven we een overzicht van een geïmplementeerd systeem gebaseerd op onze gelijkheidsmaat.

In het laatste hoofdstuk gebruiken we een propositionele taal van kwalitatieve rechthoeken om de leesvolgorde van documenten te achterhalen. Hiertoe definiëren we eerst de notie van een ‘document-codeer-regel’ en analyseren we formalismen die deze regels zouden kunnen uitdrukken, zoals \LaTeX , SGML talen, etc. Met behulp van deze regel construeren we vervolgens een detector die de leesvolgorde van documenten achterhaalt. De document-codeer-regels die we bij deze constructie gebruiken zijn uitgedrukt in de propositionele taal van rechthoeken. Om te zorgen dat ons systeem de toets aan de realiteit doorstaat, introduceren we de notie van een *thick boundary interpretation* voor een kwalitatieve relatie. Als we het systeem testen op een collectie van heterogene documenten, zien we een mate van recall van 89%.

Tot besluit bevat het proefschrift drie appendices. Appendix A is een kort overzicht van basis topologische noties die gebruikt worden in de Hoofdstukken 2, 3 en 4. Appendix B geeft een algoritme dat gerichte, transitieve, cyclische graven sorteert volgens het systeem uit Hoofdstuk 7. In Appendix C komen drie implementaties aan bod die allen aan dit proefschrift zijn gerelateerd.

Titles in the ILLC Dissertation Series:

- ILLC DS-1996-01: **Lex Hendriks**
Computations in Propositional Logic
- ILLC DS-1996-02: **Angelo Montanari**
Metric and Layered Temporal Logic for Time Granularity
- ILLC DS-1996-03: **Martin H. van den Berg**
Some Aspects of the Internal Structure of Discourse: the Dynamics of Nominal Anaphora
- ILLC DS-1996-04: **Jeroen Bruggeman**
Formalizing Organizational Ecology
- ILLC DS-1997-01: **Ronald Cramer**
Modular Design of Secure yet Practical Cryptographic Protocols
- ILLC DS-1997-02: **Nataša Rakić**
Common Sense Time and Special Relativity
- ILLC DS-1997-03: **Arthur Nieuwendijk**
On Logic. Inquiries into the Justification of Deduction
- ILLC DS-1997-04: **Atocha Aliseda-LLera**
Seeking Explanations: Abduction in Logic, Philosophy of Science and Artificial Intelligence
- ILLC DS-1997-05: **Harry Stein**
The Fiber and the Fabric: An Inquiry into Wittgenstein's Views on Rule-Following and Linguistic Normativity
- ILLC DS-1997-06: **Leonie Bosveld-de Smet**
On Mass and Plural Quantification. The Case of French 'des'/'du'-NP's
- ILLC DS-1998-01: **Sebastiaan A. Terwijn**
Computability and Measure
- ILLC DS-1998-02: **Sjoerd D. Zwart**
Approach to the Truth: Verisimilitude and Truthlikeness
- ILLC DS-1998-03: **Peter Grunwald**
The Minimum Description Length Principle and Reasoning under Uncertainty
- ILLC DS-1998-04: **Giovanna d'Agostino**
Modal Logic and Non-Well-Founded Set Theory: Translation, Bisimulation, Interpolation

- ILLC DS-1998-05: **Mehdi Dastani**
Languages of Perception
- ILLC DS-1999-01: **Jelle Gerbrandy**
Bisimulations on Planet Kripke
- ILLC DS-1999-02: **Khalil Sima'an**
Learning efficient disambiguation
- ILLC DS-1999-03: **Jaap Maat**
Philosophical Languages in the Seventeenth Century: Dalgarno, Wilkins, Leibniz
- ILLC DS-1999-04: **Barbara Terhal**
Quantum Algorithms and Quantum Entanglement
- ILLC DS-2000-01: **Renata Wassermann**
Resource Bounded Belief Revision
- ILLC DS-2000-02: **Jaap Kamps**
A Logical Approach to Computational Theory Building (with applications to sociology)
- ILLC DS-2000-03: **Marco Vervoort**
Games, Walks and Grammars: Problems I've Worked On
- ILLC DS-2000-04: **Paul van Ulsen**
E.W. Beth als logicus
- ILLC DS-2000-05: **Carlos Areces**
Logic Engineering. The Case of Description and Hybrid Logics
- ILLC DS-2000-06: **Hans van Ditmarsch**
Knowledge Games
- ILLC DS-2000-07: **Egbert L.J. Fortuin**
Polysemy or monosemy: Interpretation of the imperative and the dative-infinitive construction in Russian
- ILLC DS-2001-01: **Maria Aloni**
Quantification under Conceptual Covers
- ILLC DS-2001-02: **Alexander van den Bosch**
Rationality in Discovery—a study of Logic, Cognition, Computation and Neuropharmacology
- ILLC DS-2001-03: **Erik de Haas**
Logics For OO Information Systems: a Semantic Study of Object Orientation from a Categorical Substructural Perspective

- ILLC DS-2001-04: **Rosalie Iemhoff**
Provability Logic and Admissible Rules
- ILLC DS-2001-05: **Eva Hoogland**
Definability and Interpolation: Model-theoretic investigations
- ILLC DS-2001-06: **Ronald de Wolf**
Quantum Computing and Communication Complexity
- ILLC DS-2001-07: **Katsumi Sasaki**
Logics and Provability
- ILLC DS-2001-08: **Allard Tamminga**
Belief Dynamics. (Epistemo)logical Investigations
- ILLC DS-2001-09: **Gwen Kerdiles**
Saying It with Pictures: a Logical Landscape of Conceptual Graphs
- ILLC DS-2001-10: **Marc Pauly**
Logic for Social Software
- ILLC DS-2002-01: **Nikos Massios**
Decision-Theoretic Robotic Surveillance
- ILLC DS-2002-2: **Marco Aiello**
Spatial Reasoning: Theory and Practice